This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Gorda, Basia

Title:
SynBac:

minimal synthetic baculovirus genomes

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

SynBac:

Minimal Synthetic Baculoviral Genomes

Barbara Gorda

PhD Thesis

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Life Sciences, School of Biochemistry.

Word count: 55,676

Abstract

The baculovirus expression vector system (BEVS) is among one of the favourites in protein expression systems due to the high protein yields attainable and relative ease of use and availability. The development of the system has been an ongoing task since its discovery as a successful protein production method. To realise the full potential of the system, key improvements are still necessary, notably to reduce DNA instability due to passage effects that impede industrial applications, as well as inherent proteolysis of produced recombinant target proteins. Through data mining and phylogenetic analysis, a putative synthetic baculoviral genome, SynBac, has been designed. In an ultimate embodiment, this SynBac genome will be devoid of all elements, which negatively affect or serve no obvious purpose in target protein expression and virus replication and infection in a laboratory environment. The hypothesis is that elimination of some or all of these presumably non-essential elements will improve the characteristics of the baculovirus, as a production tool for important biomolecules in academic and industrial research.

The aim of this thesis is to iteratively create a minimal synthetic baculovirus genome, SynBac, and validate SynBac for genomic stability and protein expression properties. The first approach involved the generation of hybrid baculoviral genomes by deletion of wild-type DNA and replacement with synthetic fragments devoid of non-essential genes. Rewiring of a first fragment provided a functional initial genome, SynBac1. Addition of further synthetic fragments resulted in virus with impaired virus replication. Thus, necessitating careful identification of individual genes that were necessary to restore virus function. To streamline genome engineering, different recombinases were tested resulting in an efficient and marker-less protocol to engineer the baculoviral genome. 67 single gene deletions of the SynBac1 genome were generated, providing critical information about the requirement of individual genes. Unlike previously thought, only several genes, ac7, ac20, ac32, ac45, ac56, ac87, ac110 and ac112/113, could be deleted without negatively affecting the replication and protein expression of the virus. The Tn7 attachment site was relocated and VloxP and RoxP sites inserted into identified non-essential genes, resulting in functional viruses, thus further confirming that these genes were dispensable. The stability of the SynBac variants was analysed and suggested the SynBac genomes are more stable compared to EMBacY. The viruses appeared to develop lower titers during amplification, however this data remains to be corroborated by obtaining precise virus titer information. Notwithstanding, the protein expression capacity of the SynBac variants proved to be equal to EMBacY, suggesting the SynBac baculoviruses may exhibit superior protein productivity, on a per virion basis.

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Imre Berger, for enabling me to carry out this PhD research project and provide support and guidance during the four and a half years of my PhD. Secondly, I would like to thank my industrial collaborators at AstraZeneca, specifically Robert Roth, Jenny Gunnarsson and Marcello Maresca, for assisting me with the project and enabling me to spend 3 months at AstraZeneca, Sweden. Particular thanks go to Robert Roth, who suggested me as a suitable candidate for this PhD project, as I just completed my industrial placement year within the group at AstraZeneca.

Furthermore, huge thanks to all the past and present members of the B Floor lab, which created a great atmosphere to work in and always offered a helping hand. Special thanks go to Martin, who provided all the experimental guidance at the beginning of the project, and Kapil, Francesco, Christine, Gunjita and Julien who supported me throughout. Particularly big thanks go to Hannah, with whom I shared the first part of the PhD project, I undoubtedly would have found the project more daunting if it were not for the shared effort. Thank you for the continued support and for agreeing to read through my thesis.

I would like to reserve the biggest thanks to my partner and fellow lab member, Julien. Apart from providing continuous advice in the lab, I am so very grateful for all your assistance, tips and tricks and recommendations about the project, especially listening to my rants about my results outside of working hours and providing excellent suggestions about the figures I was producing. I would have not been able to complete this project without you, thank you for being my number one person I could rely on.

Lastly, but certainly not least, especially big thanks go out to all my family and close friends. To my Mum and Dad, thank you so much for providing endless support from the start of school to the end of my PhD, as well as both my Grandmas for their praise and interest, even though they do not understand a single thing about what I am doing. And a huge thank you to my wonderful friends, Laura and Nia, who believed in me every step of the way and always provided words of encouragement.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: Barbara Gorda DATE: 17/03/2021

Table of Contents

Abstract I

Acknowledgements II

Author's declaration III

List of Figures XIII

List of Tables XVI

List of Abbreviations XVII

Chapter 1: Introduction 1

1.1 Introduction to the Baculoviruses 2

1.1.1 Discovery and Diversity 2

1.1.2 Baculovirus Phenotypes 3

1.1.3 Infection Cycle 6

1.1.4 DNA Replication 9

1.1.4.1 Origins of DNA Replication: hrs 9

1.1.4.2 Non-hr oris 10

1.1.4.3 Trans-acting Sequences and DNA Replication Mechanism 11

1.2 Baculovirus Applications 12

1.2.1 The Baculovirus Expression Vector System (BEVS) 12

1.2.1.1 Homologous Recombination-based Systems 13

1.2.1.2 Transposition-based Systems 14

1.2.1.3 Hybrid System 17

1.2.1.4 The MultiBac™ System 19

1.2.1.5 Improving the Genome– Generation of SynBac 22

1.2.2 Further uses 23

1.2.2.1 Bioinsecticide 23

1.2.2.2 Mammalian Transduction and Gene Therapy 24

1.2.2.3 Large-scale Protein Production 25
1.3 Baculovirus Genome Instability 26
1.3.1 Baculoviruses with Genetic Alterations 26
1.3.2 Defective Interfering Viruses and the Passaging Effect 27
1.3.3 Alleviating the Accumulation of Defective Interfering Viruses 29
1.4 Engineering Approaches 30
1.4.1 Homologous Recombination 30
1.4.2 Site-specific Recombination 32
1.4.3 Synthetic Biology and Other Engineering Methods 33
1.5 Thesis Aims 38
Chapter 2: Materials and Methods 39
2.1 Materials 40
2.1.1 Commercial Products 40
2.1.2 Cell Strains 40
2.1.3 Antibiotics 40
2.1.4 Standard Buffers 41
2.2 Bacterial Growth and Molecular Biology Techniques 41
2.2.1 Cultivation of *Escherichia coli* 41
2.2.2 Transformation of Chemically Competent Cells 41
2.2.3 Transformation of Electro-competent Cells 42
2.2.4 DNA Preparation 42
2.2.5 Restriction Digestion 43
2.2.6 DNA Ligation 43
2.2.7 Gibson Assembly 43
2.2.8 Standard Polymerase Chain Reaction 43
2.2.9 Colony Polymerase Chain Reaction 44
2.2.10 Agarose Gel Electrophoresis 44

2.2.11 Tn7 Transposition 44
2.2.12 Red/ET Homologous Recombination 44
2.2.13 Site-specific Recombination 45
2.2.13.1 Fusion Reactions 45
2.2.13.2 Extraction Reactions 45
2.3 Insect Cell Culture 45
2.3.1 Cultivation of Sf21 Cells 45
2.3.2 Transfections and V0 Virus Production 45
2.3.3 V1 Virus Amplification 46
2.3.4 Viral Genome Stability Assessment 46
2.4 Fluorescence Quantification 47
2.4.1 Sample Preparation 47
2.4.2 Fluorimetry and Data Processing 47
2.5 Protein Biochemistry 47
2.5.1 SDS-PAGE 47
2.5.2 Western Blotting 48
2.5.3 Expression and Purification of Recombinases 48
2.6 Live Cell Imagining 49
2.7 Flow Cytometry 50
Chapter 3: Minimal Baculoviral Genome: The Primary Iterative Approach 51
3.1 Introduction 52
3.1.1 SynBac Design 52
3.1.2 Baculovirus Engineering 54
3.1.3 First Genomic Modification 56
3.1.4 Generation of Tools and Methods Required for Fragment Integration 57
3.2 Chapter Aims 59
3.3 Results 60

3.3.1 Iterative Fragment Removal 60
3.3.1.1Confirmation of Starting Material 60
3.3.1.2 SynBac16 61
3.3.1.2.1 Helper Plasmid and SynBac1Δ6 Isolation 61
3.3.1.2.2 SynBac16 Generation and Validation 63
3.3.1.3 Generation of Tools and Methods Required for Further Fragment Integration 64
3.3.1.3.1 Recombinase Purification and Validation 65
3.3.1.3.2 Generation of Fluorescent Donor and Acceptor Plasmids 67
3.3.1.3.3 Validation of Donor Plasmids 69
3.3.1.4 SynBac1654 70
3.3.1.4.1 Fragment 4 Modification 70
3.3.1.4.2 SynBac1654 Generation and Validation 71
3.3.2 Individual Fragment Removal 74
3.3.2.1 SynBac14 74
3.3.2.2 SynBac13 76
3.4 Discussion 78
3.4.1 Iterative Fragment Removal 78
3.4.1.1Confirmation of Starting Material 78
3.4.1.2 SynBac16 79
3.4.1.3 Generation of Tools and Methods Required for Further Fragment Integration 80
3.4.1.4 SynBac1654 81
3.4.2 Individual Fragment Removal 83
3.4.2.1 SynBac14 83
3.4.2.2 SynBac13 83
3.4.3 Concluding Remarks and Future Directions 84
Chapter 4: Generation and Testing of Single Gene Deletions in the Baculoviral Genome 87
4.1 Introduction 88

4.1.1 Single Gene Knockout Virus Generation and Testing 88
4.2 Chapter Aims 91
4.3 Results 92
4.3.1 Blasticidin Cassette Amplification and Recombineering 92
4.3.2 Blasticidin Integration Validation 92
4.3.3 Single Gene Knockout Virus Testing 95
4.3.3.1 First Round - All Genes 95
4.3.3.1.1 Fragment 2A 96
4.3.3.1.2 Fragment 2B 97
4.3.3.1.3 Fragment 3 98
4.3.3.1.4 Fragment 4 99
4.3.3.1.5 Fragment 5 100
4.3.3.1.6 Fragment 6 101
4.3.3.1.7 Other Genes 102
4.3.3.2 Second Round - Verification of Single Gene Knockout Candidates 104
4.3.3.2.1 Single Gene Deletions - Medium Performance 104
4.3.3.2.2 Single Gene Deletions - Highest Performance 105
4.3.3.3 Third Round - Retesting Highest Performing Single Gene Deletions 107
4.4 Discussion 109
4.4.1 Generation of Single Gene Knockout Viruses 109
4.4.2 Initial Testing of Single Gene Knock-out Viruses 110
4.4.3 Retesting the Single Gene Knockout Viruses 112
4.4.4 Concluding Remarks and Future Directions 114
Chapter 5: Tn7 Site Relocation and VloxP/RoxP and EYFP Insertion 117
5.1 Introduction 118
5.1.1 Tn7 Transposition Site Relocation 118
5.1.2 VloxP and RoxP Insertion 121

5.1.3 EYFP Insertion 121
5.2 Chapter Aims 122
5.3 Results 123
5.3.1 VloxP, RoxP and Tn7 Fragment Preparation 123
5.3.2 VloxP and RoxP Insertion into SynBac1 with Original Tn7 Site 124
5.3.2.1 Ac20VloxP Insertion 124
5.3.2.2 Ac87RoxP Insertion 125
5.3.3 Tn7 Site Relocation 128
5.3.3.1 Ac32Tn7 128
5.3.3.2 Ac87Tn7 129
5.3.3.3 Ac110Tn7 130
5.3.3.4 Ac112/113Tn7 132
5.3.3.5 Reassessment of Best Clones for New Tn7 Sites 133
5.3.4 VloxP Insertion into SynBac1 containing New Tn7 Sites 135
5.3.4.1 Ac32Tn7 Ac20VloxP 135
5.3.4.2 Ac87Tn7 Ac20VloxP 136
5.3.4.3 Ac110Tn7 Ac20VloxP 137
5.3.4.4 Ac112/113Tn7 Ac20VloxP 139
5.3.4.5 Reassessment of Best Clones for VloxP Insertion into New Tn7 Site Viruses 141
5.3.5 EYFP Insertion 142
5.3.5.1 SynBac1 Original Tn7 143
5.3.5.2 SynBac1 Original Tn7 Ac20VloxP Ac87RoxP 144
5.3.5.3 SynBac1 Ac32Tn7 Ac20VloxP 145
5.3.5.4 SynBac1 Ac87Tn7 Ac20VloxP 146
5.3.5.5 SynBac1 Ac110Tn7 Ac20VloxP 147
5.3.5.6 SynBac1 Ac112/113Tn7 Ac20VloxP 148
5.3.5.7 Best EYFP Insertion Clones 149

5.4 Discussion....152
5.4.1 Generation of SynBac1 Variants....152
5.4.2 Tn7 and VloxP/RoxP Localisation....152
5.4.3 EYFP Insertion - an Optimised Approach....153
5.4.4 Variations between Fluorescence Readout....154
5.4.5 Concluding Remarks and Future Directions....155
Chapter 6: SynBac Variants Validation....157
6.1 Introduction....158
6.1.1 RoxP and VloxP Site Validation....158
6.1.2 Genome Stability....159
6.1.3 Protein Expression....160
6.2 Chapter Aims....161
6.3 Results....162
6.3.1 RoxP and VloxP Site Validation....162
6.3.1.1 Redesign of VloxP Donor Plasmid....162
6.3.1.2 Integration Site Validation....164
6.3.2 Genome Stability....166
6.3.2.1 P10 Absence Confirmation....166
6.3.2.2 Microscopy Analysis during Serial Passaging....167
6.3.2.3 Fluorescence Quantification during Serial Passaging....172
6.3.2.4 Cytometry Analysis of Genome Stability....174
6.3.3 Heterologous Target Protein Expression....177
6.3.3.1 Single Protein Expression Tests....177
6.3.3.2 Single Protein Expression Retest....180
6.3.3.3 Protein Expression at V2 Amplification....182
6.3.3.3 TAF569 – Multi-subunit Protein Complex Expression....184
6.4 Discussion....186

6.4.1 VloxP and RoxP Sites are Functional 186
6.4.2 Determination of Genome Stabilities 187
6.4.3 SynBac Variants Protein Yields Match EMBacY 189
6.4.4 Concluding Remarks and Future Directions 189
Chapter 7: Conclusions and Outlook 191
7.1 Thesis Overview 192
7.2 Concluding Remarks and Future Directions 195
7.2.1 Baculovirus Genome Minimisation Strategy: SynBac1 195
7.2.2 Different Recombination Systems can be Simultaneously Utilised for Baculovirus Engineering 197
7.2.3 Targeted Single Gene Deletions Identify Truly Dispensable DNA 198
7.2.4 Tn7 Site Relocation and VloxP/RoxP Insertion Confirmed Results of the Single Gene Deletion Trial 199
7.2.5 SynBac1 Variants: Target Protein Expression and Genome Stability 200
7.2.6 Alternative Strategies for Baculovirus Improvement 201
7.2.6.1 Transfer Plasmid Modifications 202
7.2.6.2 Tn7 Transposition Improvement 202
7.2.6.3 Delaying Cell Lysis and Enhancing Protein Production 203
References 207
Chapter 8: Appendix 232
8.1 List of plasmids 232
8.2 Sequences of interest 234
8.2.1 Chapter 3 234
8.2.1.1 pDS 234
8.2.1.2 pACEBac1 (Change LoxP site for RoxP, VloxP, SloxP or Vox variants) 234
8.2.1.3 pUADM_RoxP_AmCyan 234
8.2.1.4 pUBDM_SloxP_mTagBFP 235
8.2.1.5 pUPDM_VloxP_EGFP 235

8.2.1.6 pUZDM_Vox_mCherry 236
8.2.1.7 Dre recombinase coding sequence 236
8.2.1.8 VCre recombinase coding sequence 236
8.2.1.9 SCre recombinase coding sequence 237
8.2.1.10 Vika recombinase coding sequence 237
8.2.1.11 HA-Gent-HA fragment 4 (removal of native fragment 4) 237
8.2.1.12 HA-Frag4-HA fragment with AmpR 237
8.2.1.13 HA-Gent-HA fragment 3 (removal of native fragment 3) 238
8.2.1.14 HA-Frag3-HA fragment with AmpR 238
8.2.2 Chapter 4 239
8.2.2.1 BSD cassette 239
8.2.3 Chapter 5 239
8.2.3.1 ac20VloxP fragment 239
8.2.3.2 ac87RoxP fragment 239
8.2.3.3 ac32RoxPTn7 fragment 239
8.2.3.4 ac87RoxPTn7 fragment 240
8.2.3.5 ac110RoxPTn7 fragment 240
8.2.3.6 ac112/113RoxPTn7 fragment 240
8.2.3.7 pUCDM_YFP_polh 241
8.2.4 Chapter 6 241
8.2.4.1 pUPDM_VloxP_mCherry 241
8.2.4.2 pKL_NLSmTagBFP x pUCDM_SNAPTag_GP64 x pSPL_mCherry 242
8.2.5 Baculovirus Genomes 245
8.3 List of primers 276
8.4 AcMNPV genes 291
8.5 Expected PCR sizes for validation of blasticidin integration 301

List of Figures

Figure 1.1 – Classification of the *Baculoviridae* family. 2
Figure 1.2 - Structural composition of the occlusion derived virus (ODV) and the budded virus (BV). 6
Figure 1.3 – Infection cycle of a multicapsid nucleopolyhedrovirus. 8
Figure 1.4 – Bac-to-Bac bacmid system. 16
Figure 1.5- flashBAC™ baculovirus expression system. 18
Figure 1.6 – MultiBac transfer vectors. 21
Figure 1.7 – Defective interfering particle formation during serial passaging of the baculovirus. 28
Figure 1.8 – Bacteriophage λ recombination system. 31
Figure 1.9 – Cre/LoxP recombination reactions. 33
Figure 3.1– Assigned deletion fragments within the AcMNPV genome. 53
Figure 3.2 – Experimental procedure for the modification of the baculovirus genome. 55
Figure 3.3 – Fragment 1 validation. 56
Figure 3.4 – Inspection of starting baculovirus genomes. 61
Figure 3.5 - SynBac1Δ6 and pHelper isolation. 62
Figure 3.6 - SynBac16 generation and validation. 64
Figure 3.7 - Purified recombinase enzymes and their activity. 66
Figure 3.8 - Fluorescent donor plasmids and their fusions. 68
Figure 3.9 - Fluorescent protein expression in EMBacY transfected Sf21 cells. 69
Figure 3.10 - Fragment 4 modification. 71
Figure 3.11 - SynBac1654 generation and validation. 73
Figure 3.12 - SynBac14 generation and validation. 75
Figure 3.13 - SynBac13 generation and validation. 77
Figure 4.1 - Generation of single gene deletion baculovirus genomes by Red/ET homologous recombination. 90
Figure 4.2 - Blasticidin integration validation by PCR on SynBac1 single gene deletion variants. 94
Figure 4.3 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 2A. 96
Figure 4.4 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 2B. 97

Figure 4.5 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 3. 98
Figure 4.6 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 4. 99
Figure 4.7 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 5. 100
Figure 4.8 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 6. 101
Figure 4.9 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located outside of original fragment design. 103
Figure 4.10 – EYFP and AmCyan production at 72 and 96 hpi from the middle rank of SynBac1 single gene deletion candidates. 105
Figure 4.11 – EYFP and AmCyan production at 72 and 96 hpi from the highest rank of SynBac1 single gene deletion candidates. 106
Figure 4.12 – EYFP and AmCyan production at 72 and 96 hpi from the retest of highest ranked SynBac1 single gene deletion candidates. 108
Figure 5.1 – Spatial distribution of hr regions and potential genes for deletion and insertion of the Tn7, VloxP and RoxP sites. 120
Figure 5.2 VloxP/RoxP and Tn7 site cassette design and preparation. 124
Figure 5.3 – Ac20VloxP insertion. 125
Figure 5.4 – Ac87RoxP insertion. 126
Figure 5.5 – Reassessment of SynBac1 ac20VloxP ac87RoxP. 127
Figure 5.6 – Ac32Tn7 insertion. 129
Figure 5.7 – Ac87Tn7 insertion. 130
Figure 5.8 – Ac110Tn7 insertion. 131
Figure 5.9 – Ac112/113Tn7 insertion. 133
Figure 5.10 – Reassessment of the different SynBac1 Tn7 genomes. 134
Figure 5.11 – Ac20VloxP insertion into SynBac1 Ac32Tn7. 136
Figure 5.12 – Ac20VloxP insertion into SynBac1 Ac87Tn7. 137
Figure 5.13 – Ac20VloxP insertion into SynBac1 Ac110Tn7. 138
Figure 5.14 – Ac20VloxP insertion into SynBac1 Ac112/113Tn7. 140
Figure 5.15 – Reassessment of the different SynBac1 Tn7 ac20VloxP genomes. 141
Figure 5.16– EYFP insertion into SynBac1. 143
Figure 5.17– EYFP insertion into SynBac1 ac20VloxP ac87RoxP. 144

Figure 5.18– EYFP insertion into SynBac1 ac32Tn7 ac20VloxP. 145
Figure 5.19– EYFP insertion into SynBac1 ac87Tn7 ac20VloxP. 146
Figure 5.20– EYFP insertion into SynBac1 ac110Tn7 ac20VloxP. 147
Figure 5.21– EYFP insertion into SynBac1 ac112/113Tn7 ac20VloxP. 148
Figure 5.22 – Reassessment of the different SynBac1 genomes. 151
Figure 6.1 - Fluorophore location in genome stability analysis. 160
Figure 6.2 - Fluorescent donor plasmids and protein expression in EMBacY transfected Sf21 cells. 163
Figure 6.3 – VloxP and RoxP site validation. 165
Figure 6.4 – P10 absence confirmation. 167
Figure 6.5 – Microscopy analysis during serial passaging. 171
Figure 6.6 – Quantification of fluorescence during serial passaging. 173
Figure 6.7 – Gating applied to flow cytometry data for fluorescence determination. 174
Figure 6.8 – Genome stability analysis. 176
Figure 6.9 – Rate of fluorescence decrease. 177
Figure 6.10 – Single protein expression test. 179
Figure 6.11 – Single protein expression retest. 181
Figure 6.12 – Protein expression at V2 amplification. 183
Figure 6.13 – TAF569 – multi-subunit protein complex expression. 185

List of Tables

Table 1.1 – Baculoviruses with genetic alterations........27
Table 2.1 – Cell strains utilised throughout this thesis. 40
Table 2.2 – Antibiotics, their working concentration and growth medium utilised throughout this thesis........40
Table 2.3 – Standard buffers utilised throughout this thesis. 41
Table 2.4 Excitation and emission wavelength values used for fluorimetry analysis. 47
Table 3.1 Site specific recombinases and their recognition sequences........57
Table 3.2 Colour donor plasmids and their assigned antibiotic marker, site-specific recombination system and fluorescent marker. 58
Table 8.1 –List of used plasmids in this thesis summarizing important details, antibiotic resistance, and source........232
Table 8.2 –List of used primers in this thesis displaying sequence and use. Sequences highlighted in blue represent primer overhangs........276
Table 8.3 – Planned gene deletions/retentions within Fragment 1. 291
Table 8.4 – Planned gene deletions/retentions within Fragment 2A. 292
Table 8.5 – Planned gene deletions/retentions within Fragment 2B. 293
Table 8.6 – Planned gene deletions/retentions within Fragment 3. 293
Table 8.7 – Planned gene deletions/retentions within Fragment 4. 295
Table 8.8 – Planned gene deletions/retentions within Fragment 5. 296
Table 8.9 – Planned gene deletions/retentions within Fragment 6. 296
Table 8.10 – Remaining genes of the AcMNPV baculovirus that were not included within the initial planned fragment design. 297
Table 8.11 PCR validation fragment sizes of blasticidin integration. 301

List of Abbreviations

AcMNPV	*Autographa californica* multiple nucleopolyhedrovirus
AcNPV	*Autographa californica* nucleopolyhedrovirus
AgMNPV	*Anticarsia gemmatalis* multiple nucleopolyhedrovirus
AmpR	β-lactamase/ Ampicillin resistance
BEVS	Baculovirus expression vector system
BIIC	Baculovirus infected insect cells
BmNPV	*Bombyx mori* nucleopolyhedrovirus
bp	Base pair
BP	Band pass filter
BSA	Bovine serum albumin
BV	Budded virus
CBP	Calmodulin binding protein
CuniNPV	*Culex nigripalpus* nucleopolyhedrovirus
CpGV	*Cydia pomonella* granulovirus
CRISPR/Cas9	Clustered regularly interspaced short palindromic repeats associated Cas9 endonuclease
crRNA	CRISPR RNA
CV	Column volume
DIP	Defective interfering particle
DNA	Deoxyribonucleic acid
DPA	Day after proliferation arrest
ds	Double stranded
E. coli	*Escherichia coli*
EGFP	Enhanced green fluorescent protein
EYFP	Enhanced yellow fluorescent protein
GentR	Gentamicin acetyltransferase
gRNA	guide RNA
GV	Granulosis viruses
HA	Homology arm
HearSNPV	*Helicoverpa armigera* single nucleopolyhedrovirus
hpi	Hours post infection
HR	Homologous recombination

HRP	Horseradish peroxidase
HSV-1	*Herpes simplex* virus
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilobase pair
LB	Luria-bertani
LdMNPV	*Lymantria dispar* multiple nucleopolyhedrovirus
Mb	Megabase pair
MNPV	Multiple nucleopolyhedrovirus
mTagBFP	Monomeric blue fluorescent protein
NeleNPV	*Neodiprion lecontei* nucleopolyhedrovirus
NeseNPV	*Neodiprion sertifer* nucleopolyhedrovirus
NPV	Nucleopolyhedrovirus
OB	Occlusion body
OD	Optical density
ODV	Occluded-derived virus
OpMNPV	*Orgyia pseudotsugata* multiple nucleopolyhedrovirus
ORF	Open reading frame
PAM	Protospacer adjacent motif
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PlxyGV	*Plutella xylostella* granulovirus
PlxyNPV	*Plutella xylostella* nucleopolyhedrovirus
polh	Polyhedrin
RoNPV	*Rachiplusia ou* nucleopolyhedrovirus
RNA	Ribonucleic acid
SB	SynBac
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
SEM	Standard error of mean
SeMNPV	*Spodoptera exigua* multiple nucleopolyhedrovirus
Sf21	*Spodoptera frugiperda* (IPLB-Sf21-AE)
SpliNPV	*Spodoptera littoralis* nucleopolyhedrovirus
SpltNPV	*Spodoptera litura* nucleopolyhedrovirus

ss	Single stranded
SSR	Site-specific recombinase
STDEV	Standard deviation
TB	Terrific broth
TBST	Tris-buffered saline with Tween-20
TEMED	N,N,N',N'-Tetramethylethylenediamine
TnGV	*Trichoplusia ni* granulovirus
tracrRNA	Transactivating CRISPR RNA
WB	Western blot
wt	Wild type

Chapter 1: Introduction

1.1 Introduction to the Baculoviruses

1.1.1 Discovery and Diversity

Baculoviruses are a diverse group of viruses, from the *Baculoviridae* family, containing a DNA genome that is double stranded, circular, and supercoiled. Baculovirus genomes range from 80-180 kb in size and are packed in rod-shaped nucleocapsids, ranging from 230-385 nm in length and 40-60 nm in diameter.[1] The *Baculoviridae* family is split into two genera: the nucleopolyhedroviruses (NPVs) and granulosis viruses (GVs). The genera is determined by presence of polyhedra for NPVs and granules or capsules for GVs in the infecting virion.[2,3] Their main hosts are various arthropod insects of the Lepidoptera, Hymenoptera and Diptera orders, which further classify the *Baculoviridae* into four groups depending on the host the virus infects. Refer to Figure 1.1 for the different classifications of the *Baculoviridae* family and examples of viruses within the groups.[4]

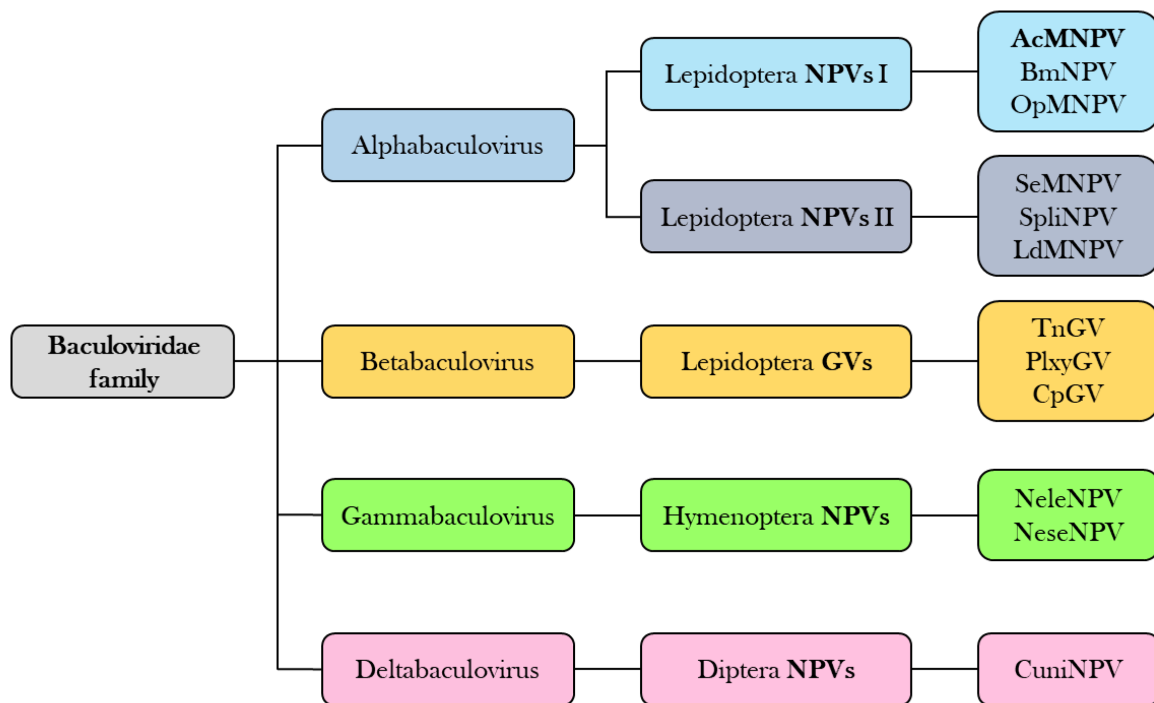


Figure 1.1 – Classification of the Baculoviridae family.

The Baculoviridae family is split into four groups depending on the insect host the virus infects. Alphabaculoviruses are nucleopolyhedroviruses that infect Lepidoptera and are further split into two groups based on which protein is used for virion fusion to host cell. Betabaculoviruses are granuloviruses that infect Lepidoptera. Gammabaculoviruses and Deltabaculoviruses are nucleopolyhedroviruses that infect Hymenoptera and Diptera, respectively.

Alpha- and Betabaculoviruses comprise NPVs and GVs, respectively, which infect Lepidoptera insects, and Gamma- and Deltabaculoviruses which infect Hymenoptera and Diptera, respectively.[5-8] The Lepidopteran NPVs are further divided into Group I and II, depending on what protein is utilised in the fusion of the virions to the host cell.[9,10] Additionally, NPVs are divided based on the presence of single or multiple nucleocapsids within polyhedra, with multiple nucleopolyhedrovirus (MNPVs) containing predominantly 5-15 nucleocapsids per envelope, and single nucleopolyhedrovirus (SNPVs) which seldom display more than one nucleocapsid per envelope.[11] One of the most studied baculovirus is the *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV), which is the viral backbone for the development of SynBac, the topic of this thesis.

A defining feature of baculoviruses is their virally encoded RNA polymerase. Similarly to other DNA viruses, baculoviruses utilise the host's transcription machinery for early gene expression and subsequently rely on their own RNA polymerase for the production of late and very late genes.[12,13] Despite the huge variation of gene content between the baculovirus genomes, a well characterised set of around 30 genes are present in all sequenced baculoviruses. This gives an insight into the key roles of the proteins responsible for infection of insect gut cells, virion structure and the expression of late genes from the virally encoded RNA polymerase.[14]

Since the initial discovery of baculoviruses as early as mid-1800s and the rod-shaped virions entirely validated using the electron microscope in the late 1940s, numerous reports have identified major functions of proteins and key components of the infection and DNA replication cycles.

1.1.2 Baculovirus Phenotypes

Baculoviruses have evolved to produce two phenotypically distinct virions to infect their insect hosts. Occlusion derived viruses (ODV) facilitate the initial infection of the host, and budded viruses (BVs) spread the infection to other cells within the host.[6] The key proteins that make up the two baculovirus phenotypes will be exemplified in the following section. Firstly, the proteins of the nucleocapsid core that is common in the two virions will be described. A protein named p6.9 (ac100) condenses the baculoviral DNA into a nucleocapsid, which is enveloped by vp39 (ac89), a capsid protein.[15-19] The nucleocapsids display a rod like shape with distinctive basal and apical structures, (Figure 1.2). The pp78/83 (ac9) protein has been shown to localise to the basal end of the nucleocapsid and plays a key role in the assembly of nuclear actin and thus, the movement of virions into the cytoplasm of the infected cell.[20,21] Another protein involved in nuclei

actin scaffolds is vp80 (ac104), although the exact association at the basal end has not been absolutely demonstrated.[22,23] FP25 (ac61) and ODV-EC27 (ac144) have been shown to function in the envelopment of the nucleocapsids in the nuclei, as FP25 deletion mutants displayed less ODVs and more BVs and no nucleocapsids were detected within the capsid in the mutant lacking ODV-EC27.[24-28] BV/ODV-C42 (ac101) encodes a capsid-associated protein that is reported to interact with the pp78/83 protein and transport it into the nuclei since nucleocapsid formation was altered in the deletion mutant.[22,29] The capsid protein (ac54), also known as VP1054, is another essential component for nucleocapsids as it's deletion caused tube-like structures rather than the expected nucleocapsids and the virus was no longer infectious.[30,31]

The nucleocapsids are enveloped in the nuclei to form ODVs, the initiators of the infection that targets the insect's midgut epithelial cells. The AcMNPV ODV contain multiple nucleocapsids surrounded by ODV-E66 (ac48), the main component of the ODV envelope, (Figure 1.2).[32] Between the envelope and capsid protein is the tegument protein, GP41 (ac80), required for the egress of nucleocapsids from the nuclei.[33,34] Another ODV associated protein is ODV-E25 (ac94), which has been shown to function in virion occlusion, considering it's deletion results in no evident ODV.[35-37] Likewise, the per *os* infectivity factors (PIFs), are also associated with the ODV envelope. PIFs are proteins that are responsible for the oral infectivity of the virus and its entry into the insect midgut epithelial cells.[38] Specifically, PIFs 1-3 (ac119, ac22, ac115) are at the core of the PIF complex and their presence is important for the assembly of the remaining PIFs.[39,40] A larger complex of PIFs 0, 4, 6, 7 and 9 (ac138, ac96, ac68, ac110, ac108) rely on the Pif 1-3 complex for binding, with Pif 8 (ac83) binding to the larger complex.[41-44] Pif 5 (ac148) is similarly associated to the ODV envelope, however it is not part of the PIF complex. Rather it is classified as a Pif due to the loss of infectivity in insects upon ac148 deletion.[40,45,46] For a more in depth review of the protein composition of ODVs, see Braunagel et al. (2003).[22]

The ODVs are packed into occlusion bodies (OB) (Figure 1.2). Here several ODVs are embedded into a matrix of polyhedrin (ac8), which naturally forms a crystal structure and serves to stabilise the virions in the environment.[47,48] The OBs are additionally sealed by polyhedron envelope protein (PEP), also known as the calyx. These form a smooth and seamless surface around the polyhedra that further enhances the stability of the OBs. The protective layer of polyhedrin and PEP means that the ODVs can persist outside of the insect indefinitely, as the crystal structure of the proteins protect the virions against UV rays and temperature changes, thus increasing their chance of transmission to another host.[49,50] Although highly stable, the OBs are

alkali-soluble and thus release the ODVs into the insect midgut (an alkaline environment), where they can initiate the infection via the epithelial cells.

Upon entry of the ODVs and during the initial infection, BVs are propagated inside the cell. The BV usually contains a single nucleocapsid (Figure 1.2) whose envelopes are substantially different from ODVs, as they acquire their envelope from budding out of the host cell's plasma membrane. Prior to budding out, the plasma membrane is modified by viral proteins, including GP64 (ac128), which is an envelope fusion protein responsible for the transfer of BVs between cells.[10,51] Some viruses lack a GP64 homolog and instead utilise a homolog of the F protein (ac23). Although ac23 is also present in GP64 containing baculoviruses, the F protein is mostly inactive, suggesting the acquisition of GP64 later in evolution and displacement of the F protein's function.[52] Another BV envelope associated protein is BV/ODV-E26 (ac16). This protein is also present in ODV, interacts with the nucleocapsid associated protein FP25 and forms a complex with cellular actin.[53] Upon deletion of BV/ODV-E26 in AcMNPV and BmNPV, a delay in the production of BV and an inhibition of OB assembly was recorded.[54] BVs are delicate in nature compared to the stable structure of the ODVs, which emphasises their role in the infection cycle. BVs are protected inside the host whereas ODVs need to survive in the external environment. A key point to highlight is the shape of the BV envelope, which resembles more of a bag, with more space around the nucleocapsid, than the previously imaged BVs which were artifacts of the negative staining procedure due to their fragile structure.[55]

The two structurally different virions demonstrate the demand for separate infection units for horizontal and vertical transmission. The ODVs and BVs are highly efficient at infecting different tissues of the host and have evolved specifically to facilitate the biphasic infection cycle of the baculoviruses.[56]

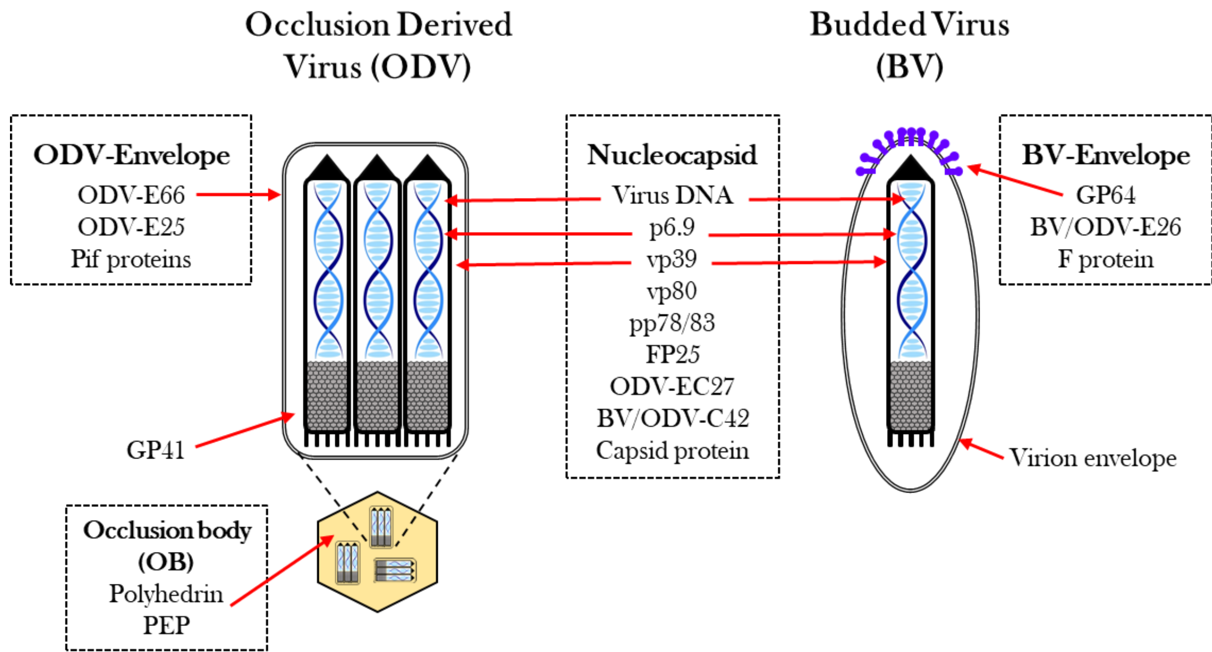


Figure 1.2 - Structural composition of the occlusion derived virus (ODV) and the budded virus (BV).

Illustration of ODV and BV, displaying the baculoviral DNA packed with the help of the p6.9 basic DNA binding protein and within the nucleocapsid made of the major capsid protein, vp39. Multiple nucleocapsids are usually packed to form the ODV, which is enveloped by ODV-E66, the component of the ODV envelope and GP41, the tegument protein between the envelope and vp39, which is required for nuclear egress of the nucleocapsids. The ODVs are further packaged into occlusion bodies (OBs), which are comprised of the polyhedrin and PEP proteins. A BV contains a single nucleocapsid enveloped in the host cell's membrane upon exit from the cell and contain the GP64 protein responsible for the entry and exit of BVs into and from cells, respectively. The key proteins associated with the nucleocapsids are listed in the middle, and proteins correlated to ODV and BV envelopes are indicated on the left and right, respectively (not an all-encompassing list of all associated proteins). For details of the proteins' functions see the text above. Image adapted from Funk et al. (1997) and Shelly et al. (2013).[57,58]

1.1.3 Infection Cycle

The biphasic infection cycle of baculoviruses relies on the generation of the two distinct virion phenotypes, ODVs and BVs (Figure 1.3). The primary infection of insects commences when the larvae ingest food contaminated with OBs. The alkaline environment of the insect midgut subsequently causes the polyhedrin matrix to dissolve and thus release the ODV (1).[59] The ODVs, now present in the midgut lumen cross the peritrophic membrane (2), before directly fusing to the microvilli plasma membrane of the midgut epithelial cells. Here, the nucleocapsids are then released into the cytoplasm (3).[60] The nucleocapsids migrate to the nucleus through

actin-based transport, the capsid is uncoated at the nuclear membrane and the naked viral DNA is released (4).[21] The initial cycles of viral replication take place through the expression of early genes, replication of genomic DNA and then the expression of late genes which generate the structural proteins of the nucleocapsids (5).[61] The virus replication takes place in the virogenic stroma located in the nucleus of the infected cell and at late stages of the infection the newly formed nucleocapsids leave the nucleus and acquire a temporary membrane (6).[62,63] During transport of the nucleocapsids from the nucleus to the basal plasma membrane of the midgut cell, this temporary membrane is lost. When the nucleocapsids bud out from the cell, they acquire a modified plasma membrane envelope that contains virus-encoded proteins such as GP64 and F protein (7).[64] The progeny BVs subsequently spread the infection to neighbouring cells by penetrating the hemocoel and either directly crossing the basal lamina of the midgut, or through the conduit of tracheal cells.[65] During primary infection of the midgut columnar cells, only a few cells produce OBs, which are usually small and lack virions.[66]

The secondary infection of various susceptible tissues is initiated through the attachment of the BVs to the cell surface and subsequent internalisation by endocytosis (8). Unlike ODVs, BVs use receptor-mediated adsorptive endocytosis to enter the cell, by utilising the envelope fusion proteins, GP64 and F protein.[67,68] The BVs are internalised into endosomes, which migrate in the direction of the nucleus.[69] The BV envelope and the endosomal membrane fuse together due to low pH and the nucleocapsids are released into the cytoplasm (9).[70] The nucleocapsid enters the nucleus and the uncoated DNA undergoes viral replication (10,11), as described earlier (4,5). More BV progeny is produced, which continues the spread of the infection throughout the insect host. In the very late phases of infection, nucleocapsids are encased in the *de novo* formed envelope to form ODVs (12). The ODVs are then embedded into the matrix of the OBs polyhedrin protein (13), which remain in the nucleus until the cell ruptures.[71] Upon death of the infected larvae by liquefaction due to virally encoded chitinase and cathepsin genes, the OBs are released into the environment, where the infection cycle can commence once again. Hence, the ODV virion is responsible for the initiation and spread of infection from insect to insect, whereas the BV virion plays a role in the viral transmission from cell to cell within an infected insect.

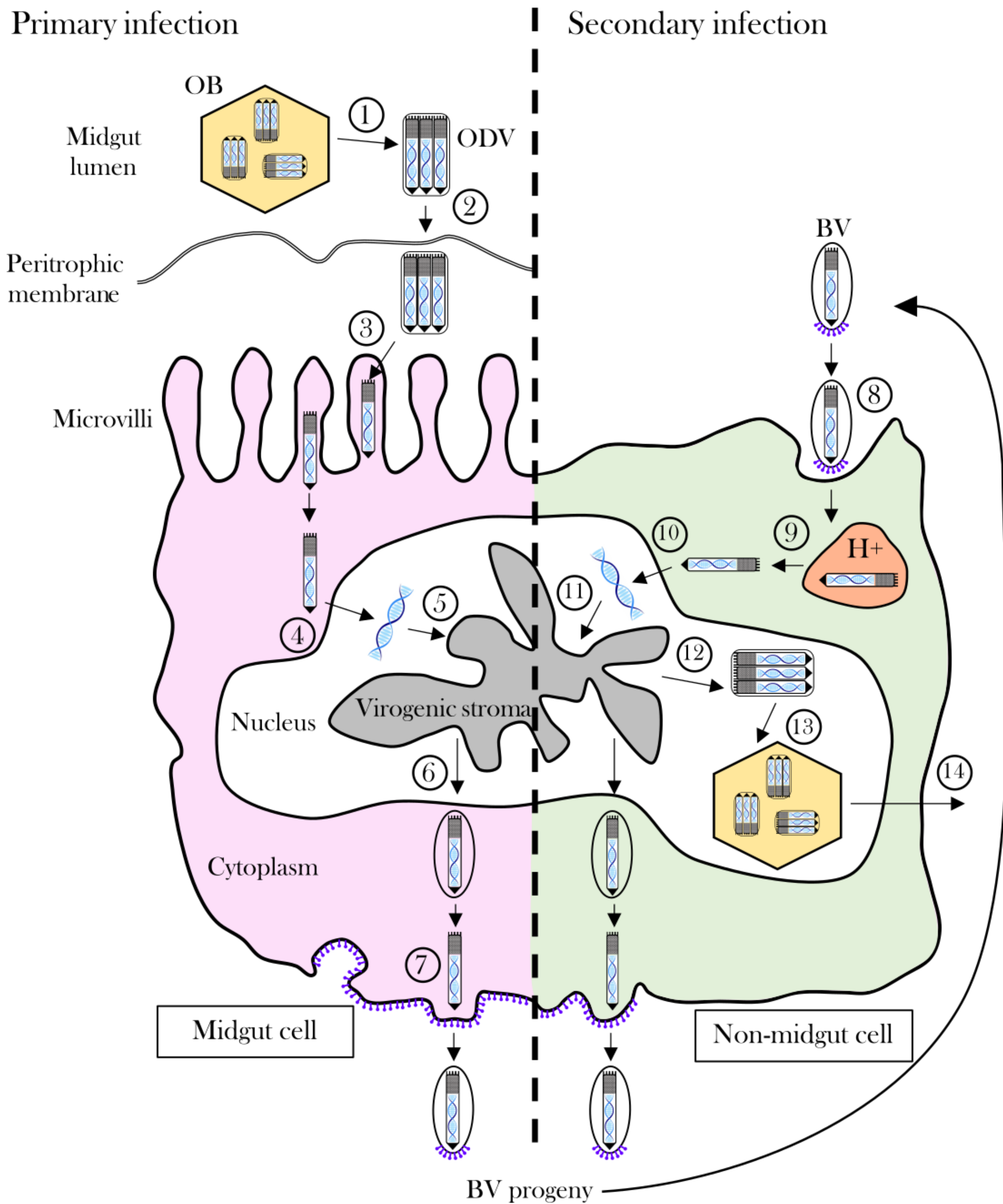


Figure 1.3 – Infection cycle of a multicapsid nucleopolyhedrovirus.

The sequential steps of primary infection of larval midgut cells and the secondary infection of non-midgut cells are showed on the left and right, respectively. The information for the numbered steps is found in the text above. Image adapted from Ikeda et al (2015).[11]

1.1.4 DNA Replication

Similarly to other large DNA viruses, baculoviruses initiate their infection cycle by utilising transcription activators and enhancers to manipulate the host transcriptional machinery. Employment of enhancer sequences and transcription factors is important for the virus to compete for the control of the strictly regulated host cell's transcription apparatus. Without these tools, the viral genome would struggle to replicate due to the overwhelming number of cellular genes undergoing transcription. Early in the infection, the virus employs transcriptional enhancers to use the host RNA polymerase to transcribe the first set of genes, including the virally encoded RNA polymerase, which later takes precedence.

1.1.4.1 Origins of DNA Replication: hrs

Viral origins of replication (oris) have been identified using two methods: transient replication assays and the analysis of defective viral genomes after serial passaging of virus at high multiplicity of infection. Transient replication assays are based on transfection of a plasmid containing viral DNA in virus infected cells. The newly replicated DNA from the plasmid is distinguished from cellular DNA by restriction digestion by DpnI. This enzyme can only cut methylated DNA, i.e. plasmid DNA derived from bacteria. Thus, no restriction digestion occurs if the plasmid is replicated in the insect cells, as the DNA sequence contains no methylation sites.[72-74] The alternative approach of identifying oris, involved serial passaging of the virus. This results in the accumulation of defective genomes that contain deletions and thus replicate faster due to their smaller size. The defective genomes rely on a helper genome containing the complete set of genes to enable their amplification. The defective genomes frequently show an enrichment of replication origins, as these sequences are favoured for amplification if the virus genome is to be replicated again.[75-77] A more comprehensive review of defective genomes will be presented in Section 1.3.

Utilising the two methods, ori activity has been identified in the baculovirus and was located to homologous regions (hrs).[78] Hrs in the AcMNPV genome are regions of DNA that contain repeats of around 70 bp units, encompass an imperfect palindromic sequence of 30 bp and have an EcoRI site (GAATTC) near their centre. The hrs in the AcMNPV genome (hr1, hr1a, hr2, hr3, hr4a, hr4b, hr4c, hr5) are interspersed at eight locations and contain 2-8 units of repeat sequence. Although, the presence of hrs is common among baculoviruses, they show great variability in terms of their homology.[79] The hrs have been classified as origins of DNA replication and transcription enhancers, as they bind the transcription activator IEI (ac147),[80] which enhances

the IEI transactivation levels by up to 1000- fold.[81,82] Additionally, hrs show a high percentage of sequences similar to cAMP and 12-O-tetradecanoylphorbol 13-acetate response elements (CRE and TRE) between the palindromes. The CRE and TRE bind to cellular transcription factors and stimulate RNA polymerase II-dependant transcription in other systems. In AcMNPV, they have been found to activate transcription in transient replication assays by binding to proteins from the host cell.[83]

The importance of hrs in the AcMNPV has been demonstrated through their deletion and subsequent analysis of the capability of the virus to replicate. Although deletion of one, or a combination of two hrs did not affect DNA replication,[84] the sequential deletion of hrs proved to reduce virus replication. A 10-fold reduction in the amount of BV produced was recorded when five hrs were removed, and this rose to 10,000-fold when all eight regions were deleted. Furthermore, removal of all eight hrs resulted in decreased levels of expression of some of the key proteins (Lef-3, GP64 and VP39) required for successful virus amplification.[85] The results from these investigations clearly highlight the significance of the hrs as origins of DNA replication and transcriptional enhancers.

1.1.4.2 Non-hr oris

As previously mentioned, hrs have been implicated in ori activity, however deletion of just one or two of the hr sequences did not affect DNA replication. The occurrence of other sequences was therefore proposed and identified via transient replication assays in different species, including AcMNPV,[84,86] OpMNPV,[87,88] SeMNPV[89-91] and SpliMNPV.[92] These sequences, named non-hr oris, lack a typical repeat or palindrome sequence that is normally present in hrs. The non-hr oris were found to be located both inside and outside of open reading frames, with a strong disposition to be near core viral genes, enabling their transmission during replication. Although the non-hr oris show no structural homology to hr oris, or each other within the same genome, they contain multiple direct and inverted repeats, AT-rich sequences and palindromes, all of which are key aspects of eukaryotic oris.[93]

Similarly to hr oris, non-hr oris were enriched in the genomes of defective baculoviruses, delivering additional evidence for their role in DNA replication.[77,94] The activity of non-hr oris during DNA replication of the baculovirus was demonstrated with quantitative PCR studies of the AcMNPV,[95] and the interaction of viral DNA polymerase with a non-hr ori in SpliMNPV.[96] An example of a non-hr ori in AcMNPV is within the coding region of the p94 gene (ac134), which showed reduced BV titers due to the spontaneous deletion of p94 upon serial passaging

of the virus.[97] Another non-hr ori is located within the helicase, also known as p143 (ac95), whereby sequence analysis exposed both a unique distribution of imperfect palindromes and a high AT nucleotide content. The p143 ori was also found to be active as an ori in mammalian cells, even without the presence of other viral proteins that are normally necessary for precise replication.[98] Recently, 12 different non-hr oris were identified in a genome of AgMNPV, with sequences found both outside and within intragenic regions. Unlike the primary sequence itself, the secondary structure of the non-hr oris may be more important in initiating replication, through the binding of transcription factors like IE-1.[99]

1.1.4.3 Trans-acting Sequences and DNA Replication Mechanism

Apart from the requirement of the hr and non-hr oris to initiate DNA replication, trans-acting elements are necessary for successful DNA replication. These include DNA polymerase (ac65), helicase (ac95), lef-1 (ac14) and lef-2 (ac6), which function as a DNA primase and primase accessory factor. These four genes are found in all sequenced baculovirus genomes to date.[100] Two other essential genes are required for ori-dependent DNA replication in transient replications assays, IE-1 (ac147) and lef-3 (ac67). IE-1 functions as a transcription factor by binding to oris, whereas lef-3 is a single stranded DNA binding protein. Dependent on the circumstances, other elements are essential for stimulating DNA replication, including IE-2 (ac151), lef-7 (ac125), p35 (ac135) and pe38 (ac153).[101]

The initiation of DNA replication in baculoviruses occurs in a similar way to other well-characterized systems, such as *E. coli*. An origin binding protein interacts with an origin of replication to unravel the region, thus enabling helicase to induce further unwinding to allow DNA polymerase binding. Three genome replication mechanisms have been described to occur in baculoviruses: rolling-circle replication,[102] theta-like replication and recombination-dependent replication.[103] The rolling-circle DNA replication mechanism was suggested due to the discovery that plasmids containing hr regions replicated as high molecular mass concatemers (long-continuous DNA molecule of multiple copies of the same genome).[73] This was also true for pUC-based plasmids lacking viral oris replicating in infected insect cells.[104] However, Kool et al (1995),[105] suggested that the replication could encompass multiple mechanism, comparable to herperviruses.[106,107] From the several findings, it has been proposed DNA replication is likely a complex combination of the three different mechanisms; with theta-like at initiation, rolling-circle during synthesis and recombination-dependent replication at later stages.

The numerous discoveries about the baculovirus infection cycle, different types of virions and DNA replication have allowed the generation of various genome modifications. These have enabled baculoviruses to be used prolifically in many and diverse fields of application. A notable example includes its use, as a protein expression tool, which will be explored and further investigated within this thesis.

1.2 Baculovirus Applications

1.2.1 The Baculovirus Expression Vector System (BEVS)

Research of the virus replication and infection cycle in larvae as well as the establishment of continuous cell lines enabled the baculoviruses, especially AcMNPV, to be utilised in a laboratory setting. A significant development that enabled this use was the identification of the highly expressed protein, polyhedrin, which forms the protective layer of the ODV. As the spread of infection within cell culture is mitigated through BV, polyhedrin was deemed as non-essential in this setting. Thus, the gene expressing polyhedrin was deleted and replaced by a gene of interest, which utilised the strong promoter to produce high levels of expression.[108] This first modification to the baculoviral genome signifies the inception of the baculovirus expression system, which has become one of the most commonly used protein expression tools.

The baculovirus is an excellent choice for use as a protein expression system due to several key aspects. Among these, is the ability of the virus to cause systemic infections. Other eukaryotic systems demonstrate high levels of protein expression, an example includes the manufacture of ovalbumin in the avian oviduct and numerous milk proteins in mammals. However, the proteins produced in these systems are limited to these specific tissues.[109] The ability of the baculovirus to cease host gene transcription allows for the efficient expression of viral genes.[110] Additionally, the termination of early and late baculovirus genes in the viral cycle allows high expression of very late genes, i.e. your gene of interest.[111-113] The baculovirus exhibits elevated levels of DNA replication, thus a high concentration of unpackaged DNA is accessible for transcription. Likewise, the baculovirus RNA polymerase is extremely efficient, which enables high levels of mRNA production and RNA capping.[114] Lastly, tremendous levels of gene expression is achievable due to the biosynthetic capacity of insect cells.

Among the key aspects of the baculovirus, additional advantages deem the baculovirus expression system as the optimum choice for heterologous protein expression, depending on the characteristics of the expressed protein. Firstly, the expression of proteins in eukaryotic insect

cells allows for more authentic folding, post-translational modifications (e.g. glycosylation, phosphorylation, internal cleavage) and oligomerization of the protein than its prokaryotic counterparts.[115-120] Although protein expression systems based on lower organisms like bacteria, yeast and fungi provide high levels of target protein expression and are simple to use, they lack the complexity to perform the above tasks.[121] Additionally, insect cell lines are comparatively easy to maintain compared to mammalian cells, as they can be cultivated in both suspension or adherent cultures and can also be grown without the addition of serum. The maintenance of insect cell lines is also relatively cheap when compared to mammalian cell lines, which require more complex media formulations. Furthermore, the deletion of the polyhedrin gene prevents ODV production, which decreases the virus survival chances in the natural environment, and insect cell lines do not support other viruses that affect humans.[122,123] These factors ensure the biosafety of the system for production of protein therapeutics and recombinant subunit vaccines.[124] Lastly, the DNA insert capacity into the baculovirus genome is virtually unlimited, as the nucleocapsid envelope shows a high degree of flexibility. Thus large DNA inserts can be successfully incorporated and expressed,[75] such as the expression of INO80 complex by Eustermann at el. (2018).[125]

1.2.1.1 Homologous Recombination-based Systems

The recognition that *in vitro* protein production requires only the presence of BVs was realised by Smith and colleagues in 1983. They not only successfully demonstrated that disruption of the polyhedrin gene was non-fatal, but also that it could be replaced by a gene of interest. Smith et all. showed that the insertion of the coding sequence for human IFN-β resulted in high-level expression of the active interferon, attributable to the strong polyhedrin promoter in *Spodoptera frugiperda* Sf21 cells (IPLB-Sf21-AE).[126] The recombinant baculovirus was generated by utilising homologous recombination between the viral DNA and a co-transfected plasmid that contained the gene to be expressed. The foreign gene was flanked with sequences corresponding to the polyhedrin region in the baculoviral genome, which allowed the homologous recombination to take place *in vivo*. The selection of recombinant baculoviruses was a time-consuming procedure, requiring skill and precision, and resulting in very low recombination efficiencies of 0.1 – 1%.[127] The process relied on identifying occlusion negative plaques in cell monolayers, which required extensive screening to select a correctly recombined virus.

A variety of modifications were undertaken to improve the efficiency and selection of the recombinant viruses. Initial engineering methods involved linearization of the virus. This aimed to favour the amplification of the recombinant genome through its re-circulation upon the

insertion of the foreign gene.[128] Furthermore, the lacZ gene was incorporated at the polyhedrin locus to aid in the selection of plaques containing the recombinant virus. The lacZ gene expresses β-galactosidase, responsible for the breakdown of a supplementary compound (X-gal or BluOGal) which generates a visible blue colour. Upon recombination the lacZ gene would be disrupted and thus the desired recombinant plaques would appear white, in comparison to the blue parental plaques. An example of the first commercial system is the BacPAK6/BaculoGold (Clontech) where three Bsu36I restriction sites were added.[129] This improved the process of homologous recombination, since the triple digested linearized genome was missing an essential gene (orf1629, ac9) that assists in the replication of the virus.[130] Recombination with a transfer vector restored the gene, which also contained the recombinant gene to be expressed.

The next generation, BacVector (Novagen), was based on the previously triple cut viruses but incorporated deletions of five additional non-essential genes. These deletions were shown to improve the expression of the foreign gene because of the decreased metabolic load in the insect cells. Two further genes, v-cath (ac127) and chiA (ac126) were later deleted which, decreased proteolysis of target proteins and enhanced secretion, respectively. Despite the improvements of the homologous recombination-based systems, these methods still required labour-intensive plaque purification. Therefore, the most commonly used systems are based on a transposition-based bacmid, as described below.

1.2.1.2 Transposition-based Systems

The undesirable plaque purification step was eliminated by the generation of a bacterial artificial chromosome (BAC), also known as a bacmid, by Luckow et al in 1993.[131] The system, Bac-to-Bac™, was patented and is currently one of the most utilised baculovirus technologies on the market (Invitrogen). The bacmid (bMON14272) contains a mini-F replicon that allows the viral genome to be maintained as episomes in bacterial cells.[132-134] This greatly simplifies the production of recombinant baculoviruses as they can be propagated in *E. coli* cells and no plaque purifications are required. The bacmid additionally contains a kanamycin resistance marker for selection and a bacterial transposon sequence, mini-attTn7, derived from an *E. coli* glucosamine synthetase (glmS) C-terminal region. The attTn7 sequence is the target site for the transposition of inserts encompassed by the left and right Tn7 recognitions sites, Tn7L and Tn7R. Five genes (TnsA-E) are required for the transposition reaction into the attTn7 site.[135-138] The TnsA-E genes are supplied in trans by the helper plasmid, pMON7124, which contains a tetracycline resistance marker for selection, but lacks the Tn7L sequence. Therefore, it does not integrate itself into

other sequences. The foreign gene to be inserted is cloned into a customisable transfer vector, pFastBac, which contains the Tn7L and Tn7R sequences.

A summary of the key components of the Bac-to-Bac™ system, including the transfer vector, the bacmid and the resulting recombinant baculovirus is illustrated in Figure 1.4. Briefly, the transposition occurs in bacteria upon transformation of the transfer vector into bacmid containing cells, DH10β. The foreign gene is inserted under the control of the strong polyhedrin promoter (polh), with a terminating polyadenylation sequence (polyA), and a gentamicin resistance marker (GentR), which can be utilised for selection in insect cells. The region between the Tn7L and Tn7R sites is inserted with the aid of the helper plasmid that encodes the transposition functions. The insertion occurs at the attTn7 site, located within the lacZα gene, which is situated between the kanamycin marker (KanR) and the mini-F replicon. Successful transposition of the Tn7L and Tn7R encompassing region disturbs the open reading frame of the lacZα gene, thereby allowing blue/white selection for identification of the recombinant bacmid. The bacteria are then amplified and the bacmid DNA is extracted through alkaline lysis. The purified bacmid is subsequently transfected into insect cells, which initiates the production of infectious BVs. The recombinant virus is further amplified and can then be utilised for the expression of the foreign gene. The entire process from recombinant bacmid generation to the analysis of expression takes around 2 weeks. This is reduced from the 4-6 weeks required for the co-transfection and plaque purification of the homologous recombination-based system.[118,139,140]

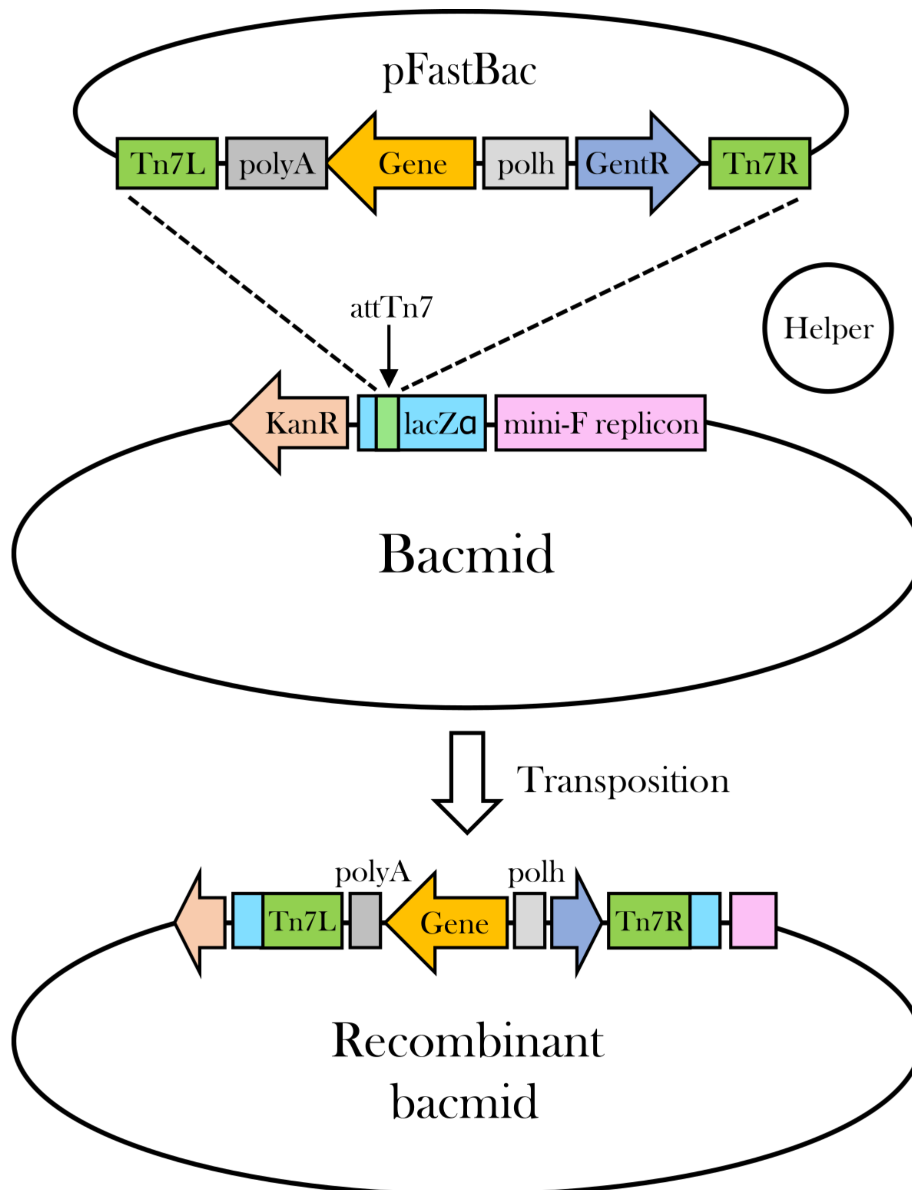


Figure 1.4 – Bac-to-Bac bacmid system.

The pFastBac transfer vector contains the foreign gene to be expressed (Gene), polh promoter and polyA signal sequence, gentamicin resistance (GentR) and flanking Tn7 recognition sites, Tn7R and Tn7L. The bacmid genome contains a kanamycin resistance marker (KanR), the mini-attTn7 transposition site located within the lacZα open reading frame and the mini-F replicon. The Helper plasmid found together with the bacmid in the DH10β cells is responsible for the expression of the Tn7 transposase, which enables the transposition of the fragment located between the Tn7L and Tn7R of the transfer vector. The resulting recombinant bacmid contains the region between Tn7L and Tn7R inserted within the lacZα gene.

1.2.1.3 Hybrid System

Subsequent modifications combined the *in vivo* recombination in insect cells, with the advantages of bacmid technology in the flashBAC™ system, (Figure 1.5).[141] The baculoviral genome was further diversified through the deletion of the C-terminus of an essential gene, orf1629 (ac9), and containing the bacterial artificial chromosome at the locus of the polyhedrin gene. The presence of the KanR and mini-F replicon permitted the amplification of the bacmid in bacterial cells. The absence of the full length orf1629 gene resulted in an inactive virus that was incapable of replicating in insect cells. However, virus replication was restored following the successful homologous recombination between a transfer vector containing the complete orf1629 gene and the defective viral genome.[142] The approach offers numerous advantages over other baculovirus systems, as there is no need for antibiotic or blue/white screening. Likewise, the necessity for a helper plasmid is removed. Plaque purifications are also not required as the non-recombinant viruses are unable to replicate. The yields and quality of recombinant proteins were shown to be improved following further deletions of non-essential genes. The chiA (ac126) and v-cath (ac127) genes were deleted in the original flashBAC constructs, as they were found to restrict the secretory pathway and induce proteolysis of expressed proteins, respectively.[143,144] In an improved version of this system, flashBACULTRA, three additional genes were deleted, p26 (ac136), p10 (ac137) and p74 (ac138). The deletion of the highly expressed p10 gene was shown to reduce the competition between the promoters for RNA polymerase.[145,146] The p74 protein is thought to be non-essential for the spread of virus in cell culture, as it is known to be involved in insect midgut infection.[147] Finally, although the function of p26 has not been well established, its deletion did not affect virus replication in cell culture.[148]

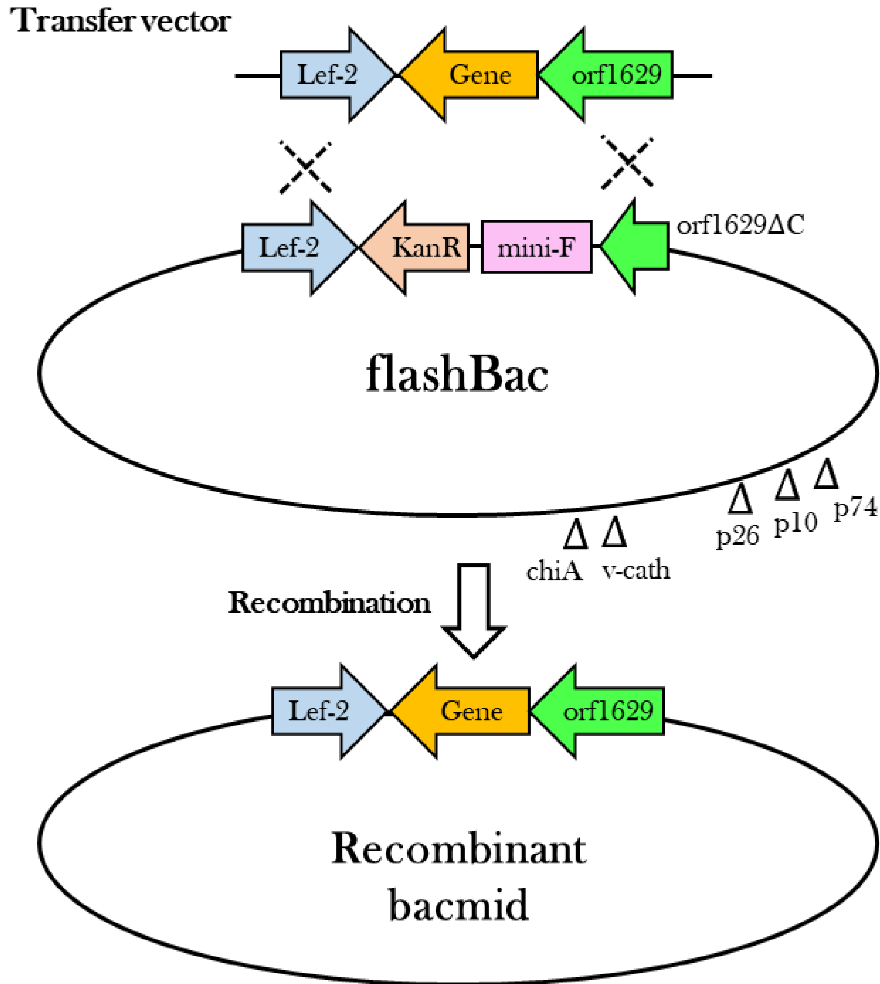


Figure 1.5- flashBAC™ baculovirus expression system.

The transfer vector contains the foreign gene to be expressed (Gene), flanked by the lef-2 and orf1629genes. The flashBAC™ bacmid contains the mini-F replicon (mini-F) and kanamycin resistance gene (KanR) for bacmid replication in bacterial cells. The orf1629 gene is truncated at the C-terminus, which prevents replication of the bacmid in insect cells. The location of the deleted inessential genes (p74, p10, 026, v-cath and chiA) is indicated (Δ). Upon recombination with the transfer vector, the full orf1629 gene is restored and the KanR and mini-F replicon is replaced by the foreign gene. Image adapted from Hitchman et al (2012). [149]

1.2.1.4 The MultiBac™ System

Up to this point, most of the systems only allowed the expression of single proteins, however expression of multi-protein complexes was facilitated by the establishment of the MultiBac system.[150–152] The technology is based on the bacmid derived by Luckow[131] with additional engineering and a set of dual promoter (polh and p10) transfer vectors. Similarly to the FlashBAC system, the MultiBac baculoviral genome was modified by deleting the chiA and v-cath (ac126 and ac127) genes for improved protein expression. In place of the two genes, a LoxP integration site was added to facilitate insertion of additional DNA cassettes, with the help of Cre recombinase, the LoxP corresponding enzyme. The designed transfer vectors contain multiplication modules and recombination elements, which facilitate the assembly of multigene cassettes.

The vectors are split into two groups, acceptors and donors, (Figure 1.6 (A)). The acceptors contain standard origins of replication (Ori) for cloning strains of *E. coli*, whereas the donors comprise a conditional origin of replication, R6Kγ, and therefore can only be amplified in pir positive *E. coli* strains.[153] Both types of vectors contain a multiple cloning site (MCS) for the insertion of expression cassettes under the control of the very late promoters, polh or p10. The vectors contain a LoxP site and different antibiotic resistance markers to enable the assembly of the fusion plasmid. The acceptor vectors contain the Tn7L and Tn7R sites for the insertion of the cassettes into the baculovirus genome by Tn7 transposition. Multiple expression cassettes can be integrated into one vector through the LoxP sites, where an acceptor and donor result in a dimer with four cassettes, and an acceptor and two donors generate a trimer with six cassettes (Figure 1.6 (B)). The assembly of the fusion plasmids is facilitated by the presence of the R6Kγ oris on the donor plasmids, as the plasmids alone are unable to replicate in cloning strain of *E. coli*, thus eliminating a possibility of unassembled plasmids persisting.

Gene insertion into the MultiBac genome can be performed through two independent mechanisms: Tn7 transposition as described in Section 1.2.1.2 for Bac-to-Bac™ and Cre/LoxP mediated recombination. The genes of interest in a donor plasmid can be inserted directly into the viral genome with Cre recombinase, through utilising the LoxP sites present in both the vector and the viral genome. The use of acceptor and donor plasmids has greatly enhanced the performance of the system when compared to previously used techniques such as co-infection or co-transfection approaches. The expression of multi-protein subunits has therefore been made uncomplicated and flexible in the MultiBac system. One example includes the expression

of the 223 kDa mediator head complex, comprising of seven subunits, for X-ray crystallography structural determination.[154]

The LoxP site in MultiBac was utilised for the creation of the EMBacY system, a real-time protein expression reporter system. A donor plasmid comprising an EYFP expression cassette, under the control of the polh promoter, was integrated at the LoxP site with Cre recombinase. The expression of the EYFP allows the user to track virus amplification and protein expression through fluorescent measurements. As the EYFP is produced under the control of the very late promoter, as are the genes of interest, the trend of EYFP expression can be directly correlated to the foreign gene production. The ability to track the fluorescence greatly assists in determining optimum time to harvest for maximal protein yield, although this will be dependent on the protein target expressed and the proteins inherent degradation pathway.

Further modifications to the MultiBac system included the addition of mammalian glycosylation patterns (SweetBac),[155,156] ability to engineer multiprotein complexes with noncanonical amino acids (MultiBacTAG),[157] as well as systems for the transduction of mammalian cells (MultiMam and MultiPrime).[158] In addition, a system (OmniBac) that allows both Tn7 transposition and homologous recombination has been developed.[159]

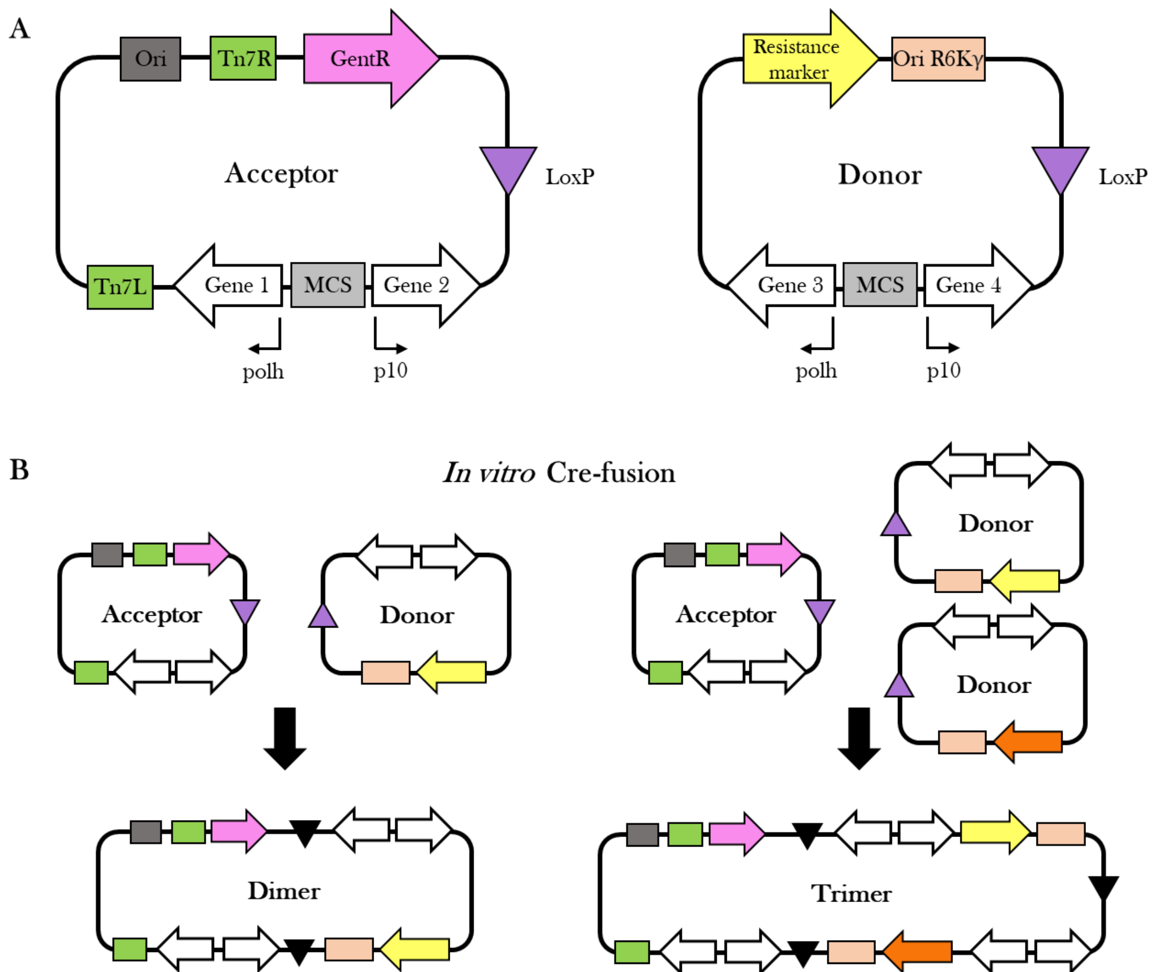


Figure 1.6 – MultiBac transfer vectors.

A) Representation of an acceptor and donor plasmid. Both vectors contain a multiple cloning site (MCS)) for the insertion of two expression cassettes driven by polh and p10 promoters (white arrows), and a LoxP site (purple triangle) for assembly of plasmids by Cre/LoxP fusion. The acceptor contains a standard origin of replication (Ori), the Tn7L and Tn7R sites for Tn7 transposition and most commonly a gentamicin resistance marker (GentR) (additional resistance markers can be found in alternative acceptor vectors). The donor contains a conditional origin of replication (Ori R6Kγ) and a range of resistance markers dependent on the donor of choice. **B)** *In vitro* Cre-fusion of acceptor and donor(s) to create multigene transfer vectors. An acceptor and one or two donors (different resistance markers, yellow and orange arrows) can be fused through Cre/LoxP mediated recombination to generate dimer or trimers, respectively. The fused expression cassettes are located between the Tn7L and Tn7R sites for integration into the MultiBac genome. Image adapted from Fitzgerald et al. (2006).[160]

1.2.1.5 Improving the Genome– Generation of SynBac

Since the discovery of baculoviruses and their application as protein expression systems, major improvements have allowed the BEVS to become one of the most popular expression systems. As presented in Section 1.2.1, engineering of the baculoviral genome, such as the insertion of a bacterial origin of replication and a lacZ-attTn7 cassette, greatly simplified the generation of recombinant baculoviruses. The generation of specific transfer vectors, as in MultiBac, further improved the system for production of multi-protein complexes.

As well as modifications to the baculoviral genome itself, numerous improvements concerning the insect cell culture were made. Among these, the optimisation of cell culture media components and the implementation of recent equipment, significantly increased the ease of culturing the insect cells.[161-163] The protocols utilised by each of the different systems on the market have been highly adapted. The methods aim to reduce the time frame for the production of protein, set out the best conditions for virus amplification and indicate the optimal point of harvest for desired proteins.[164]

Although the baculovirus genome has been considerably improved through both genome engineering and the optimisation of technical factors, nonetheless the baculovirus expression system experiences a major limitation due to the lytic nature of the infection cycle. The cell death and lysis limit the duration of protein expression and the secretory pathway is restricted at later stages of infection, compromising the folding and secretion of extracellular proteins.[165] The secretory pathway problem has been partly reduced following the deletion of the chiA and v-cath genes (ac126 and ac127), as described in FlashBAC and MultiBac. However, a decrease in cell lysis and prolonging of protein expression is desirable and such investigations are still on-going. One study performed random mutagenesis on the baculoviral vector and isolated viruses that caused decreased cell lysis. The delayed lytic baculovirus resulted in better protein folding and less protein degradation.[166] Others have engineered novel expression cassettes containing enhancer sequences and trans activator elements, such as IE1 and IE0 (ac147 and ac151), which lead to improved cell viability and higher protein expression yields.[167] Similar results were observed in a study where a stable insect cell line was made to constitutively express the apoptotic inhibitor P35 (ac135), however establishment of stable cell lines takes time and variations in yields is common between cell clones.[168]

Three additional genes, p26 (ac136), p10 (ac137) and p74 (ac138) were deleted in the FlashBAC system, without deleterious effects on the virus amplification and improved heterologous target

protein expression. Numerous baculoviral genes are specifically required for viral infection of larvae, such as the Pifs, but should be dispensable in a cell culture environment. Plentiful investigations have been carried out to determine the functions of each gene within both the AcMNPV genome and their homologs in other baculoviruses (Tables 8.2 - 8.10 in appendix). The information from these studies has been collated together and a blueprint for a synthetic baculoviral genome, SynBac, has been designed.[159] The genome minimisation aims to uncover the minimal operational genome of AcMNPV without affecting the virus amplification and protein expression functions. The creation and validation of SynBac variants, as well as a study examining viruses with single gene deletions will be presented in Chapters 3 and 4.

1.2.2 Further uses

1.2.2.1 Bioinsecticide

Although baculoviruses are most commonly known and utilised as a protein expression tool, other applications of these diverse viruses exist. Baculoviruses in the natural environment play a key role in regulating the populations of diverse insects and once this was understood, numerous developments of baculoviruses as biological pest control agents were described.[169,170] Despite many baculoviruses registered as biological pesticides, the use of viruses for insect control received little acceptance. This was due to their limited host range and the relatively slow speed in which they kill the insects.[171] To accelerate the speed of kill, baculoviruses have been engineered by insertion of insect specific neurotoxins, such as AaIT,[172] and hormones or enzymes, such as the insect juvenile esterase.[173] Additionally, the deletion of the egt gene (ac15), which encoded an ecdysteroid UDP-glucosyltransferase was found to increase the speed of insect death.[174] The generation of these recombinant viruses greatly improved their efficiency and thus decreased the level of vegetation consumption by the infected insects.

However, the in vivo production of recombinant baculoviruses for biocontrol is labour intensive. Nevertheless, this system does provide an option for high value crops that may already be resistant to standard chemical insecticides, or as potential pest controls in developing countries without access to industrial technologies.[175] The *in vitro* generation of baculoviruses as bioinsecticides at large scales may be less costly and assure less variation, however, due to *in vitro* replication, their ability to infect insect larvae may be decreased.[176] An example of baculovirus use as pest control is the HearNPV that targets the *Helicoverpa armigera* pest of cotton, which has developed numerous resistances to standard insecticides. In China, the addition of HearNPV

to insecticidal products is prevalent, with 17 products on the market produced by 10 different companies in 2014, with 40 more biocontrol products produced with 10 other baculoviruses.[177]

1.2.2.2 Mammalian Transduction and Gene Therapy

The need for multiple gene insertion into mammalian cells is crucial for modern research, such as simultaneous analysis of various parameters in a living cell or co-expression of multiple proteins for cellular reprogramming. In the past, the methods utilised included co-transfection of several plasmids, transfection of IRES-based polycistronic vectors, cleavage of precursor fusion proteins or co-infection of cells with several viruses (for example, lentivirus), each carrying a single gene. However, these approaches present numerous difficulties: the need for multiple selection markers for co-transfection of plasmids; problematic fragment cloning for polycistronic constructs; compromised protein functions due to N and C termini cleavage of precursors; and need for high biosafety laboratories for lentivirus use. Additionally, controlling the expression levels of the individual constructs and transfection/co-infection of each cell in a population is problematic and unrealistic.[178]

The delivery of multiple genes into mammalian cells has been improved by the development of the baculovirus system, MultiMam. Based on the ease of assembly of multi-construct plasmids from the MultiBac system, the acceptor and donor vectors were modified to include mammalian specific promoters (CMV or CAG).[179] The recombinant baculoviruses are first amplified in insect cells, following standard procedures and then the virus is concentrated from the supernatant using ultra-centrifugation. The virus can be subsequently used for the transduction of mammalian cells, resulting in protein expression driven by the mammalian promoters. The system can also be used for the generation of stable cell lines, thus demonstrating the diverse applications of baculoviruses in modification of mammalian cells.[120,180] An example of baculovirus-mediated multigene delivery in mammalian cells, includes the development of the MultiPrime system by Mansouri et al. (2016). This system is capable of genome editing in primary cells by delivering the CRISPR/Cas9 system.[181-183] Baculovirus based gene delivery for CRISPR/Cas9 and RNAi applications is becoming more widespread in recent years, with further studies destined to improve the system.[184,185]

In the future, baculoviruses may be used in gene therapy due to their ease of use and their large DNA cargo capacity. However, their success in this field highly depends on their efficiency of gene delivery *in vivo,* and their ability to target specific cells or tissues.[186] Baculovirus-based gene editing delivery may be improved by pseudotyping the baculovirus, such as display of vesicular

stomatitis virus glycoprotein (VSVG) on the BV membrane.[187] Furthermore, the aspect concerning the safety of the viral vector is of upmost importance. Although the virus is limited to replication in insect cells and therefore avoids the risk of persistence in the mammalian cells, further studies are required to investigate virus genome integration and the immune and cytotoxic effects.[188]

1.2.2.3 Large-scale Protein Production

Sufficient protein production for general research purposes like structural and functional studies is easily performed in small scale suspension or tissue cultures. The generation of protein therapeutics and vaccine candidates, however, requires larger yields of protein. Increasing the total volume of cell culture using small scale systems would be costly and labour intensive, thus the development of large-scale culture methods in bioreactors was required.[189] Various bioreactors for insect cell protein expression have been developed, including airlift reactors, bubble columns, packed-bed reactors, stirred tank reactors, rotating wall vessels, wave bioreactors and perfusion cultures.[190-196] As the baculovirus undergoes a lytic infection, single batch and perfusion reactors are limited in terms of production time due to decline of cell culture viability. Continuous bioreactor systems enable the cellular amplification and the virus production to be separated, hence the protein production time can be elongated.[197] However, continuous systems pose a great disadvantage as a large distribution in cell and virus residence time is observed. A large proportion of cells in the infection reactor may not yet be expressing the heterologous protein by the time they are harvested, resulting in decreased protein yields. The problem was partially fixed by exchanging the single infection reactor for two or more reactors, thus keeping the residence time the same.[198] However, although initial protein yields were higher, the overall protein production was reduced. The results suggested an accumulation of defective interfering (DI) baculoviruses, caused by an increase in the passaging effect,[199] first modelled by De Gooijer et al. (1992).[200] Numerous modifications to the methods and equipment have been performed to mitigate the passage effect,[161,201,202] however, it remains a prevalent issue for baculovirus applications. In order to optimise the protein production, which is particularly important for protein therapeutic manufacture, understanding the generation of DI viruses is crucial.

1.3 Baculovirus Genome Instability

1.3.1 Baculoviruses with Genetic Alterations

Genome plasticity is a key aspect of all biological systems. In particular, viruses display a great deal of genome plasticity due to their quick replication cycles and the large number of progeny viruses generated at each virus passage. The generation of various genotypes and phenotypes provides greater flexibility and adaptability, allowing their survival in adverse circumstances.[203] As with all viruses, baculoviruses display genome instability and the genomic alterations may exhibit themselves as deletions, insertions, point mutations, inversions, frameshifts or duplications. The generation of such alterations is true for both the natural replication of the virus in an insect host and within insect cell cultures. The phenotype of the virus mutant is dependent on the type of genetic alterations, and these can produce defective and non-defective viruses. A defective virus requires a helper virus for its replication, where the helper virus provides the necessary elements that are missing from the defective virus. Independent of the virus being defective or non-defective, the mutant viruses may be interfering and non-interfering in terms of the amplification of the parental virus. (Table 1.1) provides a brief summary of various examples of mutant baculoviruses.

	Type of mutant	Description
1.	*Infectious with no apparent change in phenotype*	Common mutations within a natural baculovirus pool that result in no phenotypical change:[204] • Single nucleotide polymorphisms[205] • Insertion of transposons (TED transposon in ac134)[206] • Hr duplications[207] • Non-hr ori expansions[208]
2.	*Infectious with different phenotype*	Deletion mutants resulting in non-defective virus. In cell culture the mutants usually lack oral infectivity or ability to liquify larvae. Also present in natural baculovirus pool.[176,209]
3.	*Infectious interfering*	Deletion/insertion mutants. Predominate upon serial passaging (replication advantage over parental virus). Insertions commonly found in the fp-25k locus (ac61) resulting in higher BV production and fewer polyhedra.[210,211]
4.	*Defective without replication advantage*	Generated during normal virus amplification and require helper virus for replication. Usually disappear after virus passaging or present at submolar quantities.
5.	*Defective interfering - Deletion mutant*	Smaller size and thus higher density of oris that interfere with the replication of parental virus. Eventually predominate in cell culture.[212,213]
6.	*Defective interfering- Duplicated oris*	Reiteration of oris during replication resulting in mutant viruses with a replicative advantage. Eventually predominate in cell culture.[76,94]

Table 1.1 – Baculoviruses with genetic alterations.

1.3.2 Defective Interfering Viruses and the Passaging Effect

Particularly applicable to standard cell culture and large-scale protein production is the generation of defective interfering viruses due to deletions, also known as defective interfering particles (DIPs). The existence of defective interfering viruses was first discovered by von Magnus (1954), where incomplete genomes of influenza virus formed during serial passage of the virus in chickens.[214] Since then, DIPs have been documented in a wide range of viruses and the impacts on virus replication and infection have been reviewed.[215] As briefly stated in the section above,

defective interfering viruses are deletion mutants resulting in a smaller size due to spontaneous deletions occurring during virus replication. The replication-defective viruses interfere with the replication of the parental virus as the density of the oris is higher and the smaller genome can be processed faster.[216] The accumulation of DIPs happens during serial passaging of the virus (Figure 1.7). Upon continuous virus amplification, deletions, particularly in the region of the inserted protein of interest, result in a decrease in number of cells containing the complete baculoviral genome, thus a decrease in overall protein yield. The serial passaging effect is especially prominent in large scale industrial protein expression where a high multiplicity of infection (MOI) and continuous production is common. The ratio of DIPs to parental virus highly depends on the MOI, as well as the virus species and host cell utilised. The generation of DIPs in AcMNPV infected insect cells and bioreactors has been examined by numerous groups.[76,77,213,217,218] Lee and Krell (1992) discovered the mean size of the virus genome at passage 65 was around 50 kb, which is approximately 40% of the size of the parental genome. A small sequence of only 2.8 kb was retained in defective genomes at passage 81, where it was subsequently shown to exclusively contain cis-acting sequences and non-hr oris, both of which are required for viral replication.[94]

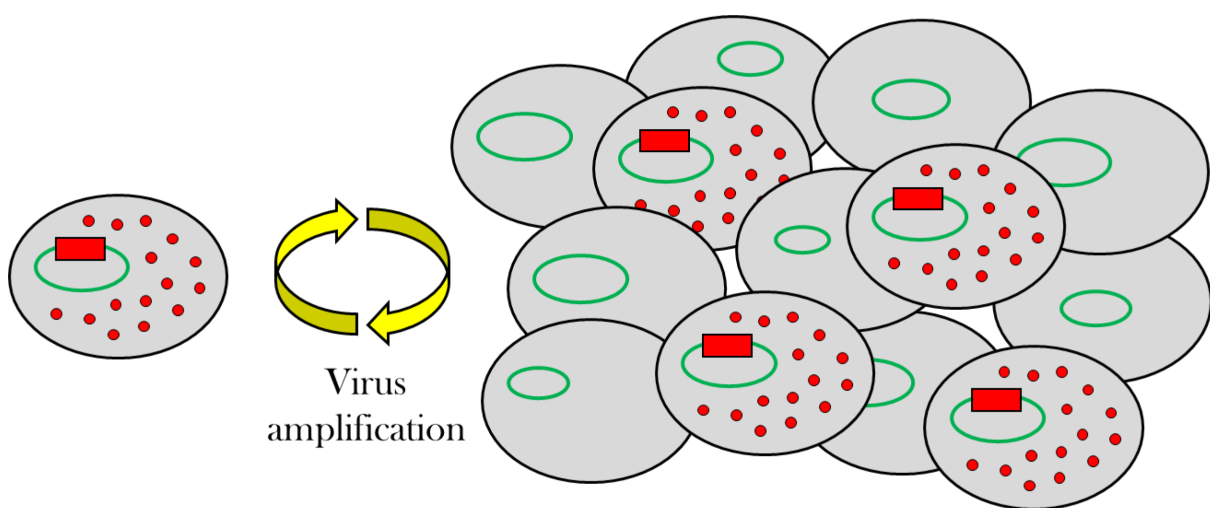


Figure 1.7 – Defective interfering particle formation during serial passaging of the baculovirus.

Illustration displaying the formation of defective interfering particles upon serial passaging of the virus. The baculovirus genome (green ovals) is susceptible to DNA deletion. In particular, the segment containing the expression cassette (red rectangle) for the protein of interest (red circles). Upon serial virus amplification the number of defective interfering particles increases due to the replication advantage over parental genome. The reduction in the number of cells containing the parental genome with the expression cassette leads to an overall decrease of protein yields.

1.3.3 Alleviating the Accumulation of Defective Interfering Viruses

The mechanisms of baculovirus defective interference was studied in detail by Gorben Pijlman (2003), utilising a sensitive PCR-based method to detect DIP formation.[219] The study confirmed that DIPs appear intrinsically with baculovirus replication in cell culture, even at the first virus passage. The generation of DIPs was a result of heterologous and homologous recombination, which can happen sequentially.[220] In addition to DNA deletions, insertions of host transposable elements into the virus genome, particularly multiple copies, can also lead to homologous recombination.[221] This can result in inversions or deletions, which further contributes to the generation of DIPs. The fp25k and egt/da26 loci in the baculovirus genome are specifically linked to this phenomenon.[210]

A further investigation of DIPs and their composition revealed the region most susceptible to deletions was the bacterially derived DNA sequence containing the mini-F replication origin, kanamycin resistance and Tn7 insertion site.[222] The baculovirus expression system is not the only system to experience instability in DNA sequences derived from bacteria. The pseudorabies virus also experienced spontaneous deletions in the bacterially derived DNA segment when it was reconstituted in mammalian cell culture. However, the instability was resolved by relocating the DNA sequence to a different locus.[223,224] The flashBAC system (presented in Section 1.2.1.3) alleviated the potential for DIP formation through the removal of the mini-F and kanamycin resistance marker by homologous recombination of the cassette containing the foreign gene.[149]

Similar to the pseudorabies virus example, the bacterial DNA region in the SynBac genome will be relocated in prospects of improving the stability of the genome. Specifically, the lacZ-attTn7 cassette will be moved but the mini-F and kanamycin resistance marker will remain at present location. The lacZ-attTn7 cassette is located in the middle of the bacterially derived region. Separating the Tn7 transposition site from the flanking mini-F and kanamycin resistance marker and inserting it within the native baculoviral genome may reduce the chances of heterologous gene deletion upon serial passaging of the virus. Together with the planned genome minimization, the relocation of the Tn7 transposition site aims to improve the genome stability and protein expression yields of the baculovirus expression system.

1.4 Engineering Approaches

Efficient genome minimization strongly relies on the right choice of engineering methods. The techniques need to be effective and easily implemented on a baculovirus genome. While the baculovirus genome is large and standard cloning procedures cannot be performed, the use of homologous and site-specific recombination provides excellent methods for easy and seamless engineering. The following section describes the two approaches, which will be utilised in the creation of SynBac. Additional techniques and examples of generated minimal genomes in the synthetic biology field will also be summarized.

1.4.1 Homologous Recombination

The baculoviral genome will be modified through recombination-mediated genetic engineering, known as recombineering. The method has become very popular in the engineering of large constructs as it removes the need for suitable restriction sites used in conventional cloning.[225,226] Examples of utilising homologous recombination to engineer large constructs includes: the generation of deletion mutants in the cos2 cosmid in Streptomyces albus and the equine herpesvirus-1 Ab4p bac.[227,228] The technique relies on linear DNA substrates that can be single or double-stranded and contain approximately 50 base pairs of homology on each side that corresponds to the target site. The linear DNA is commonly made using high-fidelity PCR, where the primer design determines the exact junction of the recombination event.[229] Using double-stranded DNA, one can produce deletions, insertions, gene replacements and inversions. In the case of modifying the baculoviral genome, a native DNA region can be deleted by recombination with a PCR fragment. This fragment contains the exact homology arms found in the target region and an antibiotic resistance gene for selection of recombinants. The same procedure can also be used for the insertion of new fragments, however the antibiotic resistance gene in these fragments should be flanked by enzyme recognition sites such as LoxP, to enable its subsequent removal.

A specific example of a system that can be used for this approach includes the Red/ET recombination method provided by GeneBridges. The recombineering occurs by utilising three Red proteins (Exo, Beta and Gam) of the λ phage, or the corresponding RecET proteins of the rac prophage (Figure 1.8). Linear DNA is normally degraded when introduced into *E. coli* by the RecBCD nuclease, but the λ Gam protein prevents this by inhibiting the nuclease activity of the RecBCD. The double-stranded linear DNA is processed by the λ Exo protein, which binds to one strand and degrades the other strand in a 5' - 3' direction. The resulting double-stranded DNA containing a 3' overhang is the substrate for the λ Beta protein, which promotes the

annealing of the complementary DNA strands at the recombination site.[230] The genes encoding these proteins need to be expressed at suitable times and levels to achieve recombination at high efficiencies. The proteins are thus expressed from an L-arabinose inducible cassette in a temperature-sensitive plasmid, usually maintained at 30 °C. However, upon induction and incubation at 37 °C, the proteins are expressed, and the plasmid is lost after several rounds of replication. This incubation at 37 °C is not only optimal for the protein expression, but also crucial for the removal of the plasmid after the recombination event where it is no longer needed.[231]

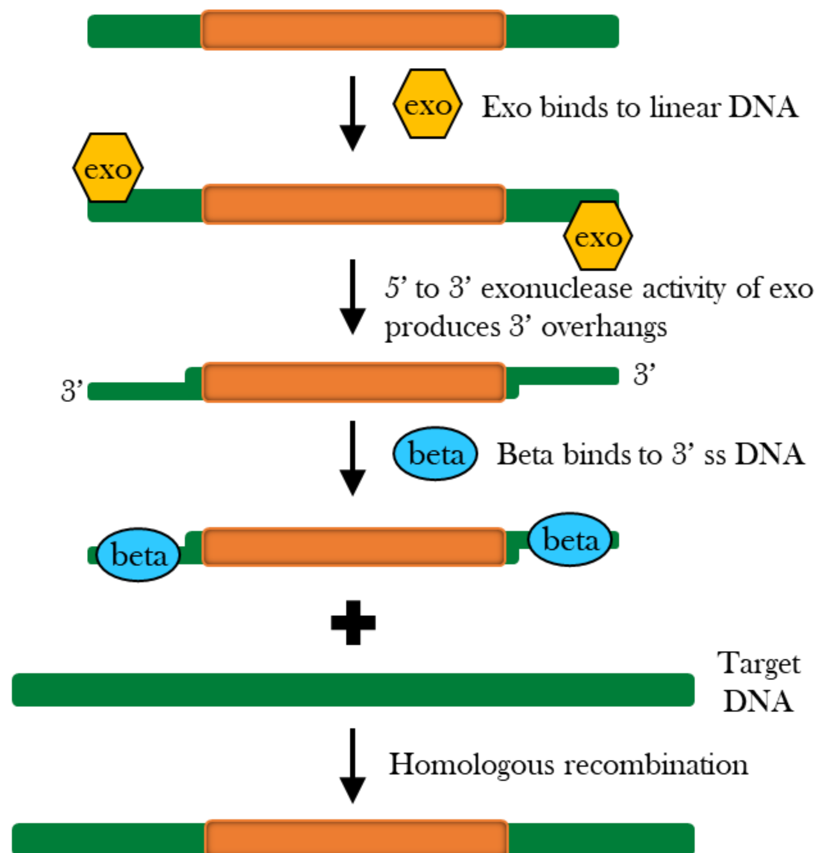


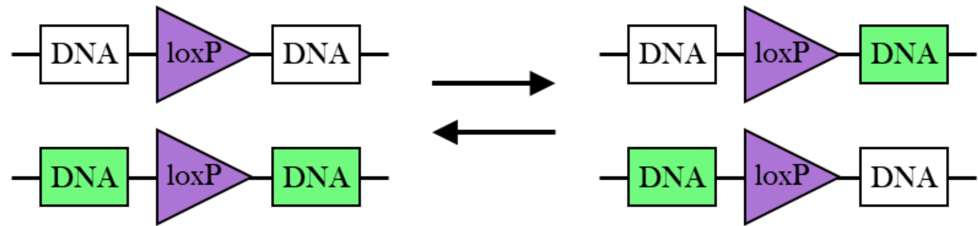
Figure 1.8 – Bacteriophage λ recombination system.

The exo protein (orange hexagon) has exonuclease activity when bound to linear dsDNA and degrades one strand in the 5' – 3' direction. The 3' overhangs are bound by the beta protein (blue oval), which promotes annealing of the homology arms in the target DNA, resulting in homologous recombination. The gam protein (not shown) counteracts the degradation of ds linear DNA by the RecBCD nuclease. Image adapted from Sharan et al. (2009).[232]

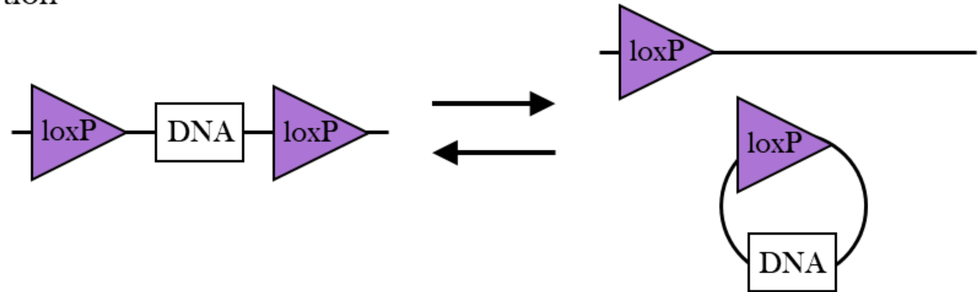
1.4.2 Site-specific Recombination

The cleavage and ligation of two specific sites on a DNA strand by a recombinase enzyme is known as site-specific recombination. Numerous examples have been identified, all of which act upon uniquely conserved asymmetric DNA sequences. A commonly used system for genome engineering includes the tyrosine site-specific recombinase Cre, which recognizes LoxP sites. Depending on the orientation and localisation of these sequences one can produce translocations, deletions, insertions and inversions (Figure 1.9). Translocations are achieved by placing the LoxP sites on different DNA strands, where the DNA sequence after the LoxP site on one strand exchanges for the DNA sequence found after the LoxP site on the other strand. Deletions are achieved by placement of LoxP sites on the same strand and with same orientation. Inversions are a result of LoxP sites placed on the same strands but in opposite directions, resulting in the DNA sequence or gene reversing its directionality. The Cre/LoxP system is originally derived from the bacteriophage P1 and does not require any additional protein or sequences in addition to the Cre and its recognition site.[233,234] The mechanism of action involves a nucleophilic reaction by a tyrosine hydroxyl group on the phosphodiester backbone to form the recombination intermediate. Cre recombinase mediates recombination between two LoxP sites, which are 34 bp long and include an 8 bp spacer and 13 bp palindromes on either side.[235] These sites are small enough to be introduced easily in genome engineering, but large enough to reduce cryptic occurrence and off targets due to their well-defined sequences. The Cre recombinase binds to the two palindromes and forms a dimer, which then forms a tetramer with the dimer present on the second site. Due to the directionality of the LoxP sites, they are parallel in orientation upon tetramer formation. The Cre protein cuts the double stranded DNA at both LoxP sites and the strands are efficiently re-joined by a DNA ligase. The high efficiency of the Cre recombinase has been displayed previously by mediating translocations between non-homologous chromosomes and the deletions of large regions.[236–238] For our purpose, the system will be used for excision of antibiotic resistance markers during genome modifications of SynBac.

Translocation



Deletion/Insertion



Inversion

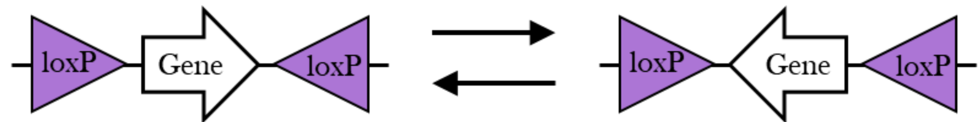


Figure 1.9 – Cre/LoxP recombination reactions.

The bacteriophage P1 Cre/LoxP system catalyses reversible site-specific recombination reactions of sequences flanked by LoxP sites (purple triangle). Translocations are achieved when the LoxP sites are found on separate DNA strands, resulting in the exchange of the DNA sequences (green and white rectangles) found after the LoxP sites between the two strands. Deletions of DNA sequences can be achieved by placing the LoxP sites, in the same direction, on either side of the DNA sequence. The reverse reaction achieves the insertion of the circular fragment containing one LoxP site back into the original DNA fragment/plasmid via the LoxP site. Placing two LoxP sites in the opposite direction on either side of a DNA sequence or gene (white arrow), results in the inversion of the DNA sequence. Image adapted from Kratochwil and Rijli (2014).[239]

1.4.3 Synthetic Biology and Other Engineering Methods

The field of synthetic biology is expanding as DNA synthesis technologies rapidly evolve. The design and creation of synthetic genomes is enabling us to study the structure, function and evolution of genomes, with the goal of improving the genome application for research. In particular, the design of minimal genomes provides insight into gene essentiality and functions, which can be translated into the generation of optimised genomes.

1.4.3.1 CRISPR/Cas9

Since the discovery of the CRISPR/Cas9 system and its ability to precisely and efficiently engineer genomes, the explosion of this genome editing tool has widely facilitated research. The most commonly known CRISPR/Cas9 system originated from a type II CRISPR-Cas system in *S. pyogenes* bacteria, with a variety of other systems and subtypes available from different organisms. In bacteria, the CRISPR-Cas system acts as a defence mechanism against invading foreign DNA, such as viruses and plasmids.[240] The system incorporates foreign DNA sequences between CRISPR repeat sequences, known as arrays, in the bacterial host genome. This generates CRISPR RNA (crRNA) transcripts, which contain a 'protospacer' region of variable sequences from the foreign DNA. Each crRNA hybridizes to a transactivating CRISPR RNA (tracrRNA), and the pair associates with the Cas9 nuclease.[241,242] The protospacer-encoded section of the crRNA recognises the foreign DNA and directs the Cas9 complex to cleave the target sequence found 20 nucleotides 5' to protospacer adjacent motifs (PAMs).[243] The CRISPR/Cas system widely used for engineering utilizes a single guide RNA (gRNA), which is a fusion between the crRNA and a section of the tracrRNA. The gRNA forms a complex with the Cas9 and guides it to cleave target DNA sites lying adjacent to the 5'-NGG PAM sequences. The cleavage generates a double stranded break. This leads to insertions or deletions (indels) when repaired by non-homologous end joining, or accurate repair by homologous recombination in the presence of a donor template DNA. With the precise target recognition and presence of a multitude of types of CRISPR/Cas systems, ability to efficiently edit, modify or regulate specific loci in an extensive range of cells and organisms has been made possible.[244]

Although, the CRISPR/Cas system has been mainly developed for use in mammalian cells and organisms, engineering examples have also been presented in prokaryotic cells and viruses. The CRISPR/Cas9 system was exploited in editing the large 152 kb genome of the *Herpes simplex virus* (HSV-1). When suitable target sequences were utilised, the editing efficiency was more than 50%. This circumvented the construction of bacterial artificial chromosomes and enabled the use of the CRISPR/Cas9 system for the functional evaluation of viral genes.[245] Another example of engineering viruses with CRISPR/Cas9 includes the poxvirus. Although the repair of breaks was inefficient, Cas9 cleavage resulted in DNA replication inhibition, thus suppressing the virus spread. This unforeseen result can be utilised for selection of conventionally produced poxvirus recombinants, where the separation of recombinants from the parental virus is otherwise unfeasible without the presence of selection markers. As a result, the CRISPR/Cas9 application has potential for significantly increasing the speed of poxvirus-based vaccine manufacture.[246]

Recently, the use of CRISPR/Cas9 for baculovirus engineering in Sf21 cells was reported.[247] The Cas9 and gRNA complex was delivered by lipofection and generated knockouts in ODV-E26, F-Protein, p74 and Ac18. The editing efficiency ranged from 10-40% depending on the gene, with a range of deletions, insertions and frameshift mutations. The targeted genes are known to be non-essential in the baculovirus expression vector system and no significant differences were seen in the production of budded virions or protein expression. Likewise, CRISPR/Cas9 was applied on the wild-type virus for the production of non-genetically modified viruses for use as insecticides. The truncation of the egt gene was previously shown to improve the insecticidal properties of the virus, and the knockout generated displayed comparable improvements to prior studies.[248] The potential for knock-ins was also explored, confirming the ability of the CRISPR/Cas9 system for baculovirus engineering. The transfection of the gRNA and Cas9 complex into the Sf21 cells and subsequent plaque purifications of edited viruses is, however, time-consuming. Carrying out the editing on the bacmid propagated in *E. coli* cells instead, avoids the need for *in-vitro* assembly of the gRNA/Cas9 complex and virus purification by plaque assays. Nevertheless, the overall design and prerequisite for PAM sequences makes the CRISPR/Cas9 system less suitable for precise gene deletion compared to the Red/ET homologous recombination method, as utilised in this thesis.

1.4.3.2 Chemical Synthesis

Template-independent chemical synthesis paved the way for synthetic biology in the early 2000s. The complete poliovirus complementary DNA (7.4 kb) and infectious genome of bacteriophage φX174 (5.4 kb) were synthesised by oligonucleotide assembly.[249,250] The key steps of chemical synthesis are gel purification of pooled oligonucleotides and their ligation under strict annealing conditions. The last stage involves assembly of the full-length genomes from the ligation products by polymerase cycling. The successful synthesis of these viral genomes edged the field towards the synthesis of larger and more complex organisms. The Craig Venter Institute described the synthesis of the *Mycoplasma genitalium* (583 kb) and *Mycoplasma mycoides* (1.1 Mb) genomes in 2008 and 2010, respectively.[251,252] Assembled cassettes of chemically synthesised oligonucleotides (5 to 7 kb) were combined using *in vitro* recombination. The intermediates (24, 72 and 144 kb) were then cloned into *E. coli* cells as bacterial artificial chromosomes. These were sequenced before the complete genome was built by transformation-associated recombination cloning in yeast *S. cerevisiae*. The endeavours reported by the Craig Venter Institute enabled chemical synthesis in the megabase scale and immensely progressed *in vitro* DNA assembly procedures, as well as genome transplantation techniques. The work on

Mycoplasma mycoides was further pursued for the generation of a reduced genome in order to obtain information about the minimal gene set required for cellular life. Using a primary design centred around molecular biology knowledge and transposon mutagenesis data, the genome was put through a cycle of designing, building and testing. The genome, JCVI-Syn3.0, was reduced from 1.1 Mb to 531 kb, with a total of 473 genes, and resulted in the smallest known genome capable of autonomous replication.[253]

1.4.3.3 Computational Tools for Genome Design

The availability of computational tools for genome design is limited, however algorithms utilising minimal genomes are emerging. Minimal genomes provide a simple readout for algorithms by assessing the cell replication ability. Recently, the whole-cell model of the *Mycoplasma genitalium* was put through computational cycles with the goal of uncovering new minimal in silico genomes. The results indicated the possibility of smaller genomes than the already reduced JCVI-Syn3.0, as well as the detection of 10 genes with low essentiality.[254] The process of laboratory testing based on current knowledge of individual genes is laborious and expensive. Thus, combining laboratory findings and computational tools may serve as an attractive approach for improved genome engineering.

1.4.3.4 Yeast Engineering

Whole-genome synthesis and minimization was extended into eukaryotic genomes with the organisation of an international consortium with the goal of reconstructing all 16 chromosomes of *S. cerevisiae*. By 2014 the first artificial yeast chromosome synIII (273 kb) was synthesised.[255] Five further chromosomes were produced, with approximately 40% of the total yeast genome redesigned by 2018.[256-260] Among one of the engineering approaches utilised for exploring the generated S. cerevisiae chromosomes was synthetic chromosome rearrangement and modification by LoxP-mediated evolution (SCRaMbLE). The method aims to generate combinatorial genomic diversity of segments by rearrangement of the sections flanked by designed recombinase sites. The project, named Sc2.0 specifically examined the genome of the S. cerevisiae yeast, where 43 site-specific recombination sites were introduced into the synthesized synIXR chromosome arm.[255] Based on the Cre/LoxP system, an inducible form of the recombinase, Cre-EBD was designed to act on the modified sites, LoxPsym. The deep sequencing results of 64 unique isolates validated the capability of SCRaMbLE as a tool for exploring genomic diversity. In total 156 deletions, 94 duplications, 89 inversions and an additional 55 complex rearrangements were discovered. All strains retained segments containing

essential genes and genes conferring a fast growth phenotype.[261] The data from the minimal genomes acquired could be used for identifying dispensable genes, thus streamlining the genome for industrial applications. Alternatively, the duplications potentially highlight genes that may confer advantage to the genome if multiple copies are present. Additionally, the engineering of vectors and development of exogenous pathways have been combined into one SCRaMbLE method for improvement of heterologous protein expression. Examples of the engineered pathways include β-carotene and violacein, where continuous SCRaMbLEing of the genome resulted in further optimisation of the pathways.[262] Overall, the SCRaMbLE technique offers valuable information about the diversity of a genome and in combination with phenotype-based selections, presents an excellent choice for acquiring designer genomes.

1.4.3.5 Rewriting of Genetic Code and Incorporation of Artificial Amino Acids

An amino acid can be encoded by multiple codons, offering the possibility of the redundant codons to be erased or reassigned. This can be utilised to generate organisms with modified genetic codes capable of incorporating artificial amino acids. These unnatural amino acids can confer a multitude of properties from post-translational modifications including: acetylation and ubiquitination, increasing the activity or stability of therapeutic proteins, or enabling the study of protein folding through the use of fluorescent amino acids, such as p-cyanophenylalanine.[263–265] Until recently, most of the genome-wide rewriting attempts were performed in viral genomes, as well as few examples including *E. coli*, *Salmonella* and *S. cerevisiae*.[266,267] By the process of oligo-mediated recombineering, every single TAG stop codon in *E. coli* (321 total) was changed to TAA, validating the redundancy of the stop codon.[268,269] Subsequently, 13 sense codons of ribosomal genes and 123 occurrences of the rare arginine codons (AGA and AGG) were altered in *E. coli*.[270,271] Genetic code expansion was implemented into the MultiBac system to combine the advantages of high expression of protein complexes with site-specific incorporation of artificial amino acids. The resulting MultiBacTAG system was utilised for fluorescent labelling of target proteins using click chemistries, glycoengineering of antibodies such as Herceptin, as well as structure-function studies of transcription factor protein assemblies.[157] With chemical synthesis methods developed, more complex rewriting of gene cassettes and genomic substitutions have been possible with the goal of creating new organisms with tailored functions.[272]

The range of engineering approaches introduced above represent a handful of methods utilised for the creation of designer genomes, with new techniques constantly developed. However, the ease of use and precision of Red/ET homologous recombination and site-specific recombination perfectly complies with the objective of this thesis.

1.5 Thesis Aims

The work described within this thesis is based on the hypothesis that the currently utilised AcMNPV baculovirus genome can be minimised by deletion of non-essential genes. The removal of these genes and relocation of the Tn7 transposition site are believed to increase protein production and aid the genome stability of the virus. Thus, the fundamental aim of this thesis is to:

- Iteratively create a minimal synthetic baculovirus genome, SynBac.

The key aim will be accomplished through the following objectives:

- Generation of minimal baculoviral genome by deletion of fragments consisting of multiple non-essential genes.
- Development of tools and methods required for engineering the baculoviral genome.
- Provision of information about the essentiality of individual genes.
- Relocation of the Tn7 transposition site.
- Validation of the generated SynBac variants.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Commercial Products

If not otherwise stated, all chemicals were purchased from Sigma Aldrich. DNA preparations, clean-up and gel extraction kits were purchased from Qiagen and New England Biolabs. All enzymes and their corresponding buffers were purchased from New England Biolabs. Primers were purchased from Eurofins and Sigma Aldrich.

2.1.2 Cell Strains

Strain	Growth Medium	Source
DH10β	LB	Invitrogen
BL21 Star (DE3)	LB	Invitrogen
Top10	LB	Invitrogen
BW23473	LB	Berger Group
IPLB-Sf-21-AE (Sf21)	Sf-900 II Serum free medium	Berger Group

Table 2.1 – Cell strains utilised throughout this thesis.

2.1.3 Antibiotics

All purchased from Sigma Aldrich, apart from blasticidin, puromycin and zeocin that were purchased from Thermo Fisher. For blasticidin selection of work carried out in Chapter 4, Fast-Media® Blas TB sachets from InvivoGen were used.

Antibiotic	Concentration (µg/mL)	Growth Medium
Ampicillin	100	LB
Blasticidin	100	Low Salt LB pH 8
Chloramphenicol	30	LB
Gentamicin	10	LB
Kanamycin	50	LB
Puromycin	100	LB
Spectinomycin	50	LB
Streptomycin	30	LB
Tetracycline	10	LB
Zeocin	50	Low Salt LB pH 7.5

Table 2.2 – Antibiotics, their working concentration and growth medium utilised throughout this thesis.

2.1.4 Standard Buffers

Buffer	Composition
Luria-Bertani (LB) medium	1% [w/v] tryptone, 0.5% [w/v] yeast extract, 1% [w/v] NaCl
LB agar	1.5% [w/v] agar in addition to LB
Low Salt LB	1% [w/v] tryptone, 0.5% [w/v] yeast extract, 0.1% [w/v] NaCl
TAE	40 mM Tris pH 8, 20 mM acetic acid, 1 mM EDTA
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
SDS-PAGE running buffer	25 mM Tris/HCl pH 8.2, 192 mM glycine, 0.1% [w/v] SDS
Protein loading buffer (4x)	200 mM Tris pH 6.8, 8% [w/v] SDS, 40% [v/v] glycerol, 4% [v/v] β-mercaptoethanol, 50 mM EDTA, 0.08% [w/v] bromophenol blue
Coomassie staining	40% [v/v] ethanol, 10% [v/v] acetic acid, 0.2% [w/v] Coomassie brilliant blue
SDS-PAGE de-stain	10% [v/v] ethanol, 7.5% [v/v] acetic acid
TBST	20 mM Tris pH 7.5, 150 mM NaCl, 0.1% [v/v] Tween-20
PBST	PBS as above with the addition of 0.1% [v/v] Tween-20
BCIP/NBT	100 mM Tris/HCl pH 8.2, 100 mM NaCl, 5mM $MgCl_2$

Table 2.3 – Standard buffers utilised throughout this thesis.

2.2 Bacterial Growth and Molecular Biology Techniques

2.2.1 Cultivation of *Escherichia coli*

E. coli cells were cultivated in LB medium and LB agar plates supplemented with appropriate antibiotics, unless selecting for Blasticidin, where TB media and the Fast-Media® Blas TB sachets were used. Grown at 37°C, unless otherwise stated for Red/ET reactions at 30°C and bacterial protein expression.

2.2.2 Transformation of Chemically Competent Cells

100 ng – 1 µg plasmid DNA and 5 – 10 µl ligation reactions were placed into 50-100 µl chemically competent *E. coli* cells and incubated on ice for 30 minutes. A heat shock was applied by placing the cell mixture at 42°C for 45 seconds, followed by 2 minutes on ice. 1 ml LB medium was immediately added, and cells recovered at 37°C (unless otherwise stated) for 1 – 16 hours. Cells were plated onto LB agar plates containing appropriate antibiotics, with the addition of IPTG

(0.1 mM final concentration) and BluOGal (100 μg/ml final concentration) for baculoviral transformations.

2.2.3 Transformation of Electro-competent Cells

100 ng – 1 μg plasmid DNA of larger size and up to 1/5 volume of baculoviral DNA were firstly buffer exchanged into MQ water using a Millipore 0.22 μm membrane (30 minutes – 1 hour) or the Monarch DNA clean-up kit. The DNA was added to 50 μl electrocompetent *E. coli* cells and transferred to 0.1 cm electroporation cuvettes (BioRad) kept on ice. The cuvettes were wiped dry and electroporated using the MicroPulser™ Electroporator (BioRad) at 1.80 kV. 1 ml LB medium was added to the cuvette and transferred to an eppendorf for recovery at 37°C (unless otherwise stated) for 1 – 16 hours. Cells were plated onto LB agar plates containing appropriate antibiotics, with the addition of IPTG (0.1 mM final concentration) and BluOGal (100 μg/ml final concentration) for baculoviral transformations.

2.2.4 DNA Preparation

LB medium supplemented with appropriate antibiotics were prepared and transferred into 14 ml polypropylene round-bottom tubes (Falcon) or 25-100 ml autoclaved conical flasks. Colonies were transferred to the LB medium using inoculation loops and cultures incubated at 37°C (unless otherwise stated), 180 rpm for 12 – 16 hours The DNA was prepared using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. Culture volumes of >10 ml used double the volume of P1, P2 and N3 buffers. For bacmid preparation, 3 ml cultures were centrifuged at 4,000 x g for 10 minutes and supernatant removed. Bacmid was prepared using buffers from Qiagen Bacmid prep kit. The pellet was resuspended in 300 μl of Buffer P1 and transferred to an eppendorf tube. 300 μl of Buffer P2 was added, gently inverted and left for no longer than 5 minutes. 300 μl of Buffer N3 was added, gently inverted and centrifuged at 16,000 x g for 10 minutes. The supernatant was transferred to a new eppendorf and centrifuged for an additional 5 minutes. The supernatant was transferred again to a new eppendorf tube and precipitated using 100% isopropanol (40% v/v final), tubes inverted and centrifuged for 10 minutes. The supernatant was carefully removed and 200 μl of 70% ethanol was added and centrifuged for 5 minutes. The ethanol was removed and additional 50 μl ethanol added. Under a sterile hood, the ethanol was removed, and pellet left to air dry. The DNA was resuspended by adding 30 μl of filter sterilized H_2O. 20 μl used for transfection of Sf21 cells and 10 μl taken for analysis by agarose gel electrophoresis.

2.2.5 Restriction Digestion

Restriction digests were used to validate plasmids by restriction mapping and for cutting of vector and inserts according to enzyme manufacturer's manuals. Typically, 1-2 µl DNA (100-300 ng total concentration), 1 µl CutSmart buffer (NEB), 0.5 µl enzyme was used and filled up to 10 µl with MQ water. The reactions were incubated for 2-3 hours at 37 ˚C, unless otherwise stated, and restriction mapping reactions analysed by electrophoresis (0.8-1% agarose gel).

2.2.6 DNA Ligation

100 – 300 ng of vectors and inserts with compatible sticky ends were mixed in 1:1 molar ratio with 2 µl T4 DNA ligase and 2 µl T4 DNA ligase buffer in a 20 µl reaction. The reaction was incubated at room temperature for 1-2 hours, following transformation into appropriate competent cells.

2.2.7 Gibson Assembly

Inserts and vectors of 2:1 molar ratio respectively with total 200 ng vector were added to 15 µl Gibson assembly master mix (New England BioLabs) and filled with MQ water up to 20 µl total volume. The reactions were incubated for 1 hour at 50 ˚C and transformed into appropriate competent cells.

2.2.8 Standard Polymerase Chain Reaction

PCR was used to amplify fragments of interest using the SimpliAmp Thermal Cycler (Life Technologies) and Thermo Fisher reagents. In general, 30 ng of template DNA was mixed with 10 µl 5X HF or GC buffer, 2.5 µl of forward primer and reverse primer at 10 µM concentration, 1 µl of 10 mM dNTPs mix, 0.5 µl Phusion High-Fidelity DNA Polymerase and filled up to 50 µl total volume with MQ water. Typically, the template was initially denatured at 98˚C for 30 seconds, following a cycling through a denaturing step at 98 ˚C for 10 seconds, an annealing step at primer specific temperatures and an elongation step at 72 ˚C for 30 seconds per kb of gene to be amplified. The cycle of denaturing, annealing and elongation was repeated 30 - 35x, followed by a final elongation step at 72 ˚C for 10 minutes. The PCR reactions were stored at 4 ˚C prior to utilisation. PCR fragments were purified using PCR cleaning and gel extraction kits from New England Biolabs and Qiagen, respectively.

2.2.9 Colony Polymerase Chain Reaction

Colony PCR was used to check for the correct modifications to the baculoviral genome. Colonies to be checked were mixed in 35 µl of water. 2.5 µl of the diluted colony mix was added to 4 µl 5X HF buffer, 1 µl of forward primer and reverse primer at 10 µM concentration, 0.4 µl of 10 mM dNTPs mix, 0.2 µl Phusion High-Fidelity DNA Polymerase and filled up to 20 µl total volume with MQ water. The reactions were initially denatured at 98˚C for 10 minutes to allow cell lysis, following a cycling through a denaturing step at 98 ˚C for 30 seconds, an annealing step at primer specific temperatures for 30 seconds and an elongation step at 72 ˚C for 30 seconds per kb of gene to be amplified. The cycle of denaturing, annealing and elongation was repeated 35x, followed by a final elongation step at 72 ˚C for 10 minutes.

2.2.10 Agarose Gel Electrophoresis

DNA was separated using Thermo Scientific OWL Easycast B1-BP equipment. DNA was mixed with Purple Loading Dye (6X, NEB) and resolved on 0.8 – 1% agarose gel containing GelRed (10,000X, Biotium) in TAE buffer. Fragments were sized against 1 kb and 1 kb Plus DNA ladders (NEB). Electrophoresis was performed at 100 – 150 V for 30 – 60 minutes and imaged using a myECL™ imager (Thermo Scientific). If required, following DNA extractions were performed using the Monarch® Gel Extraction Kit (NEB) as per manufacturers protocol.

2.2.11 Tn7 Transposition

Recombinant baculoviruses were generated by transforming an acceptor or acceptor-donor fusion containing the heterologous gene(s) into chemically competent cells comprising a baculovirus genome and the pHelper (pMON7124). Reactions were incubated at 37°C for 16 hours to allow for the Tn7 transposition to occur.

2.2.12 Red/ET Homologous Recombination

E. coli cells containing a bacmid to be modified were transformed with the pREDET or pKD46 and overnight cultures grown at 30 °C. 10 ml of LB was inoculated with 300-500 µl of the overnight culture and incubated at 30°C until OD_{600}: 0.3. 1.4 ml cultures were induced with 50 µl L-arabinose (10%) and incubated at 37°C for 1 hour. The cells were prepared for electroporation by centrifugation at 10,000 x g for 30 seconds at 2°C, supernatant discarded, and cell pellet re-suspended in 1 ml 10% sterile glycerol. This was repeated twice, and the cell pellet re0suspended in the remaining ~50 µl supernatant. 100 - 200 ng of DNA fragment (prepared by

PCR) was added and electroporated. The transformations were recovered for 2-4 hours before plating onto LB-agar containing the appropriate antibiotics and incubated overnight at 37°C.

2.2.13 Site-specific Recombination

2.2.13.1 Fusion Reactions

Fusions reactions between acceptor and donor plasmids of the Cre/LoxP, Dre/RoxP, SCre/SLoxP, VCre/VLoxP and Vika/Vox systems were set up in 25 µl reactions. 300 - 500 ng total DNA with equal molar concentrations were mixed with 2.5 µl of Cre buffer for Cre/LoxP and 3.1 NEB buffer for remainder and 1 - 3 µl of recombinase enzyme (Cre (NEB), remainder purified in-house). For fusion reactions with the baculoviral genome no more than 5 µg of bacmid was used. The reactions were incubated at 37 °C for 30 minutes to 1 hour, heat inactivated at 70°C, dialysed against MQ water for 30 – 60 minutes and transformed into appropriate competent cells.

2.2.13.2 Extraction Reactions

Resistance markers were removed from 24 µl of prepared baculovirus DNA mixed with 3 µl of the specific recombinase (Dre, VCre, SCre and Vika) based on the recognition sites (RoxP, VloxP, SloxP and Vox) present and 3 µl of 3.1 Buffer (10x, NEB). The reactions were incubated at 37°C for 1 hour, heat inactivated at 70°C for 10 minutes, dialysed against MQ water for 30 – 60 minutes and electroporated into DH10β cells.

2.3 Insect Cell Culture

2.3.1 Cultivation of Sf21 Cells

IPLB-Sf-21-AE (Sf21, Life Technologies) cell line was routinely maintained between 0.4 x 10^6 and 2 x 10^6 cells/ml in Sf900 II SFM media (Life Technologies). The cultures were grown at 27°C in 250-1000 ml glass Erlenmeyer flasks (Pyrex) with screw caps at 100 rpm on MaxQ 3000 shakers (Thermo Scientific) and in 50 ml conical centrifuge tubes (Falcon) at 370 rpm on Compact Digital microplate shaker (Thermo Scientific). The cell densities were determined using a Zeiss Primovert microscope and a Neubauer counting chamber.

2.3.2 Transfections and V0 Virus Production

Recombinant baculoviral genomes were extracted as previously described in Section 2.2.4. 200 µl of Sf900 II SFM media was added to each tube containing 20 µl bacmid DNA. Transfections

were performed on 6 well plates (VWR) in duplicates. 2 ml of medium and 0.8-1 x 10^6 cells in 1 ml were transferred into the wells and left for 15 minutes at 27 ˚C. For X tubes of DNA, X00 µl of medium and 2*X0 µl of XtremeGene transfection reagent (Roche) were mixed in a tube. 100 µl of the transfection reagent mixture was added to each DNA containing tube. 2 x 150 µl of the DNA-transfection reagent mix was added drop by drop to each duplicate well. A cell only and medium control were included. The plates were incubated at 27°C for 72 hours, then the V0 virus harvested from the 6 well plates by combining the supernatant from the duplicate wells.

2.3.3 V1 Virus Amplification

1.4 ml of V0 virus was added to 10 ml cell culture at 0.7-0.8 x 10^6 cell/ml in 50 ml falcon tubes and the concentration maintained below 2.0 x 10^6 cells/ml. The cells were counted every 24 hours to identify DPA (day after proliferation arrest) and V1 virus (supernatant) harvested at DPA+24h hours by centrifuging at 4,500 g for 10 minutes and stored at 4°C protected from light. For comparison experiments of the different SynBac variants, 0.5 x 10^6 cells were taken out every day, centrifuged at 16,000 x g for 3 minutes and supernatant discarded for later analysis of protein production (fluorescence and SDS-PAGE).

2.3.4 Viral Genome Stability Assessment

V0 and V1 virus amplification occurred as previously described above. V2-V5 cultures were similarly set up in 50 ml falcon tubes with 10 ml cell culture at 0.7-0.8 x 10^6 cell/ml and virus harvested by centrifugation as previously. 120 µl of V1 was added to start V2 amplification, and 85 µl of V2, V3 and V4 was added to start V3, V4 and V5 amplification, respectively. 0.5 x 10^6 cells were taken at 48 hpi, centrifuged at 16,000 x g for 3 minutes and supernatant discarded for fluorimetry analysis. 1 x 10^6 cells were transferred to a 6 well plate with 2 ml medium prior to analysis by fluorescence microscopy. A Draq7 (BD Biosciences) dilution mixture (0.3 µM) was prepared as per manufactures protocol and added to cells in 5 ml round-bottom tubes (Falcon) for flow cytometry analysis.

2.4 Fluorescence Quantification

2.4.1 Sample Preparation

Protein production in the mutant baculoviral genomes was quantified by fluorescence assays. The 0.5 x 10^6 cells samples collected at V1 amplification were resuspended in 250 µl PBS and sonicated for 10 seconds at 70% amplitude. The samples were centrifuged at 16,000 x g for 3 minutes and 100 µl of the supernatant was placed into black 96-well BD Falcon micro-plates. A known value of EYFP standard and PBS were included as controls.

2.4.2 Fluorimetry and Data Processing

Fluorescence data was measured using a Tecan infinite M200 Pro spectrophotometer microplate reader using appropriate excitation wavelengths and emission scans for each fluorophore. The excitation wavelength and the highest emission peak wavelength values, which were taken for plotting data, can be visualised in (Table 2.4). The EYFP values were adjusted according to the EYFP standard and all triplicates averaged. For comparison experiments of the different SynBac variants, the data was normalised to the control baculoviral genome (SynBac1 or EmBacY) = 1.

Fluorescent protein	Excitation (nm)	Emission (nm)
EYFP	488	528
AmCyan	430	488
mTagBFP	402	452
mCherry	570	606
EGFP	460	510

Table 2.4 Excitation and emission wavelength values used for fluorimetry analysis.

2.5 Protein Biochemistry

2.5.1 SDS-PAGE

Protein samples were resolved by SDS-PAGE using manually cast 12 – 18% Tris-Glycine or NuPAGE™ 4 – 12% Bis-Tris Protein gels depending on size of protein and resolution range required. The resolving gel layer was composed of 30% (37.5:1) bis acrylamide (Bio-Rad) (12 – 18% total), 375 mM Tris-HCl pH 8.8, 0.1% [w/v] SDS, 0.4% [v/v] APS and 0.05% [v/v] TEMED, and the stacking gel layer on top containing 6% of the (37.5:1) bis acrylamide, 125 mM Tris-HCl pH 6.8, 0.1% [w/v] SDS, 0.4% [v/v] APS and 0.05% [v/v] TEMED. Protein samples were mixed with 4x Protein Gel Loading Buffer (Table 2.3) and boiled for 5 minutes at 95°C prior to loading.

Protein ladder used was the PageRuler™ Plus Prestained Protein Ladder (ThermoFisher Scientific). The samples were resolved at 100 V in SDS-PAGE running buffer (Table 2.3) or MES buffer (ThermoFisher Scientific) until proteins passed through the stacking gel and then at 200 V until the gel front reached the bottom of the gel. Gels were stained in Coomassie Blue (Table 2.3) for 30 minutes, following destaining in SDS-PAGE de-stain until desired level of band visibility was achieved.

2.5.2 Western Blotting

Resolved proteins on the SDS-PAGE gels were transferred to PVDF membranes using the Trans-Blot Turbo Transfer System (BioRad) using the mixed molecular weight setting (7 minutes, 2.5 A, 25 V). The membranes were blocked with TBST (Table 2.3) with 3 % bovine serum albumin (BSA) for 1 hour at room temperature and incubated with either a 1:5000 Anti poly Histidine Alkaline Phosphatase Mouse Monoclonal Antibody (Sigma A5588), 1:5000 Anti Penta Histidine HRP conjugate (Qiagen 163023544) or 1:2000 Rabbit Polyclonal CBP tag antibody (Abcam 119488) dilution in TBST with 0.2% BSA at room temperature for 1 hour or 4°C overnight. The membranes were washed with three successive 8-minute TBST washes. For the CBP detection the membrane was incubated with a 1:4000 secondary Goat anti-Rabbit HRP antibody (Thermo Fisher 31463) for 1 hour at room temperature followed TBST washes as above. The Alkaline Phosphatase antibody detection utilised the NBT/BCIP solution (Roche) diluted in NBT/BCIP buffer (Table 2.3) according to manufacturer's protocols. The HRP antibody detection utilised the SuperSignal™ West Pico PLUS Chemiluminescence Substrate (Thermo Fisher) according to manufacturer's protocols and visualised on a myECL™ (Thermo Fisher).

2.5.3 Expression and Purification of Recombinases

BL21-Star *E. coli* cultures containing the plasmids encoding for the four different recombinases, pET-28a_(Dre/SCre/VCre/Vika) 6xHis were used to inoculate LB medium with starting $O.D_{600}$ of ~0.08 and grown shaking at 180 rpm, 37ºC until $O.D_{600}$ of ~0.5. The cultures were transferred to 18ºC, induced with 1 mM final IPTG and harvested after 2 hours by spinning at 6,000 x g, 4ºC for 20 minutes. The pellets were flash frozen in liquid nitrogen and stored at -20ºC prior to purification.

The cell pellets were defrosted on ice and resuspended in lysis buffer (50 mM HEPES pH 7.0, 500 mM NaCl, 10 mM Imidazole, 3 mM β-mercaptoethanol, protease inhibitor tablets (cOmplete™ EDTA A-free, Roche) with 1 ml lysis buffer per 1 g of pellet. The resuspended

pellets were sonicated at 70% amplitude, 10 seconds on, 20 seconds off, for 7 - 10 minutes total and centrifuged at 20,000 x g, 4ºC for 1 hour. The His-tagged proteins were purified on Ni-NTA superflow resin (approx. 100 µl/1 g of pellet) (Qiagen). The resin was equilibrated with 20 CV lysis buffer, 20 CV high salt wash buffer (50 mM HEPES pH 7.0, 1 M NaCl, 50 mM Imidazole, 3 mM β-mercaptoethanol, protease inhibitor tablets) and final 20 CV lysis buffer prior to passing the soluble protein material twice through the column. The protein was eluted with an increasing imidazole gradient of 2 -3 CV of 150 mM, 300 mM and 500 mM concentrations, using dilutions of the elution buffer (50 mM HEPES pH 7.0, 500 mM NaCl, 500 mM Imidazole, 3 mM β-mercaptoethanol and protease inhibitor tablets). The protein elutions were analysed by SDS-PAGE and combined elutions were dialysed at 4ºC for 1 hour in 8 kDa MWCO MIDI GeBaFlex Tubes (Generon) in dialysis buffer (50 mM HEPES, 500 mM NaCl, 3 mM β-mercaptoethanol and protease inhibitor tablets). Where needed, the protein samples were concentrated in 10 K cut off 15 ml concentrators (Merck). The protein was flash frozen in liquid nitrogen in 30 µl aliquots and stored at -80ºC.

2.6 Live Cell Imagining

Fluorescent protein production in infected Sf21 cells was analysed by life cell imaging using widefield and confocal microscopy on either 6 well plates (WVR) or 8 chamber Nunc™ LabTek™ cover glass plates (Thermo Fisher). Widefield data was acquired using a Leica DM I6000 inverted epifluorescence microscope with the following filters: DAPI (350/50, BP 460/50) for mTagBFP, GFP (480/40, BP 527/30) for EYFP and Rhod (546/10, BP 580/40) for mCherry. Confocal data was acquired using a Leica SP8 AOBS confocal laser scanning microscope attached to a Leica DM I6000 inverted epifluorescence microscope. According to the fluorophore's excitation wavelengths (Table 2.4) and the emission spectrum, the following lasers were utilised for confocal imaging: 50 mW 205 nm diode laser, 65 mW Argon laser (458, 476, 488, 496 and 514 nm lines), 20 mW solit state yellow laser (561 nm) or a 2 mW Orange Hene (594 nm). The resulting fluorescence was detected by Lecia Hybrid detectors and the separation of excitation and emission was accomplished using an acusto-optical beam splitter (AOBS) in combination with notch filters to supress any reflected laser light. The utilised lenses included a 10X PL Fluotar and 63X HC PL APO CS2 with either 512 x 512 or 1024 x 1024 pixels. Images were processed using FIJI (ImageJ).

2.7 Flow Cytometry

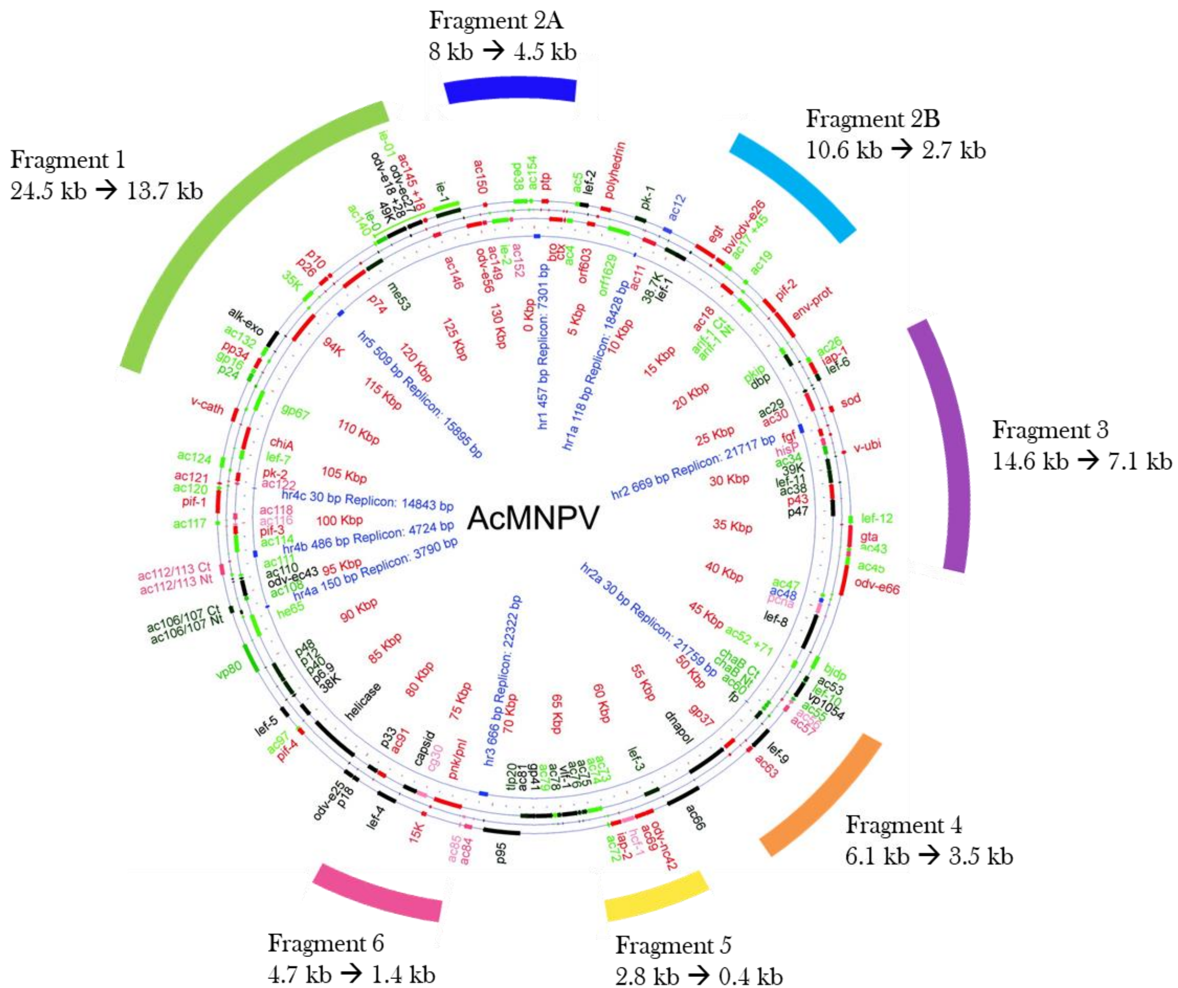
The presence of heterologous fluorescent proteins in Sf21 cells was analysed by flow cytometry after infection of recombinant baculoviruses. The samples were prepared as per Section 2.3.4, with the Draq7 staining for live and dead cell differentiation and data acquired on the same day with a Fortessa X 20 with Diva software (BD Biosciences). The data was analysed using FlowJo X (BD Biosciences). Lasers and band pass filters (BP) utilised for the fluorophores were as following: EYFP (488 nm laser, 530/30 BP), mTagBFP (405 nm, 450/50 BP) and mCherry (561 nm, 610/20 BP). The percentage of EYFP positive cells was used as a representation of the transfection efficiency of the different baculovirus genomes, with the number of mTagBFP and mCherry positive cells within the EYFP positive cells as a depiction of the stability of the baculovirus genomes.

Chapter 3: Minimal Baculoviral Genome: The Primary Iterative Approach

3.1 Introduction

3.1.1 SynBac Design

A streamlined, minimised baculoviral genome, SynBac, has been designed. The blueprint is based on extensive data mining of gene functions, homology alignments and phylogenetic analysis of all viruses in the nuclear polyhedrosis family, of which the precursor genome, MultiBac, belongs to. The proposed genome reduction of the AcMNPV genome was from approximately 133 kb to 90 kb. Among the DNA to be deleted from the genome are deleterious apoptotic factors, proteases, and non-essential genes for virus propagation in a laboratory setting. Examples of these include the genes involved in the infection of larvae or the survival of virion in the natural environment. The disruption of v-cath from the MultiBac genome and other commercial systems such as flashBACULTRA™ (Oxford Expression Technologies), demonstrated longer viability of the cell culture, thus increased protein expression yields. From the evaluations, seven broadly distributed regions were identified, and their original size and planned reduction is displayed in Figure 3.1. In total 59 genes were identified as potential candidates for deletion. Fragments 1-4 include the deletion of the native DNA and replacement with synthetic fragments created using sequence and ligation independent cloning (SLIC) methods. Fragments 5 and 6 comprise only deletions. The rewired fragments are to only contain the essential DNA elements and are to be successively introduced in the vector to create SynBac, a part wild-type and part-synthetic vector. After each fragment replacement, validation of viability, infectivity and protein expression is required.



Genome: 133,894 bp, 156 ORF, 9 Homology Regions, 59 Deletion Candidates, 97 Residual ORF

Count	Category
45	Type 1 and 2: Deletion was published as tolerable or likely harmless (loss in next neighbours or no mRNA)
9	Type 3: Deletion is perhaps harmless (loss in next neighbours)
5	Type 4: Deletion is perhaps harmless (variable gene lost in neighbours)
2	Variable in clade-I and NPV (no deletion category)
35	Unique in AcMNPV or conserved in clade-I and clade-Ib
10	Conserved in NPV
50	Conserved in lepidopteran baculoviruses or core genes

Figure 3.1– Assigned deletion fragments within the AcMNPV genome.

A schematic representation of the AcMNPV genome (top) indicating the fragments planned for engineering (Fragments 1, 2A, 2B, 3, 4, 5 and 6). The annotated genes are coloured based on the understanding of their essentiality using the scale (bottom), from essential (black and green) to non-essential and possible deletion candidates (red and pink). Image adapted from Vijayachandran et al. 2013.[159]

(Tables 8.2 to 8.9) in the appendix (section 8.4) lists all the genes planned for deletion and retention within each of the fragments as designed by Deepak Balaji, a former lab member. Likewise, (Table 8.10) lists the remaining genes found in the AcMNPV genome that were not included in the primary engineering design. (Tables 8.2 to 8.10) summarise the information about each gene's functions and thus its suspected essentiality.

3.1.2 Baculovirus Engineering

The approach chosen for the engineering of the baculovirus genome relies on homologous recombination (HR), specifically the Red/ET system (previously described in Section 1.4.1). The procedure (Figure 3.2) comprises two, successive HR reactions. Firstly, the native bacmid DNA segment is removed and replaced with GentR, allowing for selection of the modified genome. Subsequently, this resistance marker is replaced with the synthetic DNA containing the essential genes from within the fragment concerned. In both cases, the HR reactions are guided by the presence of homology arms (HA) comprising the viral DNA that flanks the region of interest. This first reaction to remove the native DNA is crucial to prevent undesirable off-target recombination reactions. Bypassing this initial reaction would create multiple undesirable homology regions between the viral genome and the essential genes within the synthetic fragment.

During the second HR reaction to re-insert the essential viral genes, additional DNA elements are incorporated into the synthetic fragment. These include an alternative selection marker, specifically ampicillin (AmpR), and the tools required for its subsequent removal. As the Red/ET HR occurs at relatively low efficiencies, the presence of a selection marker is vital. Attaining the correctly modified bacmid without this selection pressure would be a tedious task. However, the methods utilised for this reaction required thought, as one cannot perform the iterative modifications with identical antibiotic resistance markers. The use of multiple antibiotic resistances would seriously impact the freedom to use an appropriate resistance marker for the generation of recombinant bacmids during Tn7 transposition.

Among the planned methods for the removal of the selection marker are site-specific recombination (SSR) and the use of homing endonucleases (HE). Fragment 1 was designed to utilise the Cre/LoxP system, where the antibiotic marker is flanked by LoxP sites and excised upon the action of the Cre recombinase. Fragments 5 and 6 were designed to include homing endonuclease sites, PI-PspI and I-SceI, respectively, where subsequent restriction digestion removes the resistance marker. The loss of the resistance marker and lack of a positive selection pressure means the colonies attained need to be restreaked onto agar plates to confirm the loss of resistance, and the size of the modified region validated by PCR at each stage of the engineering.

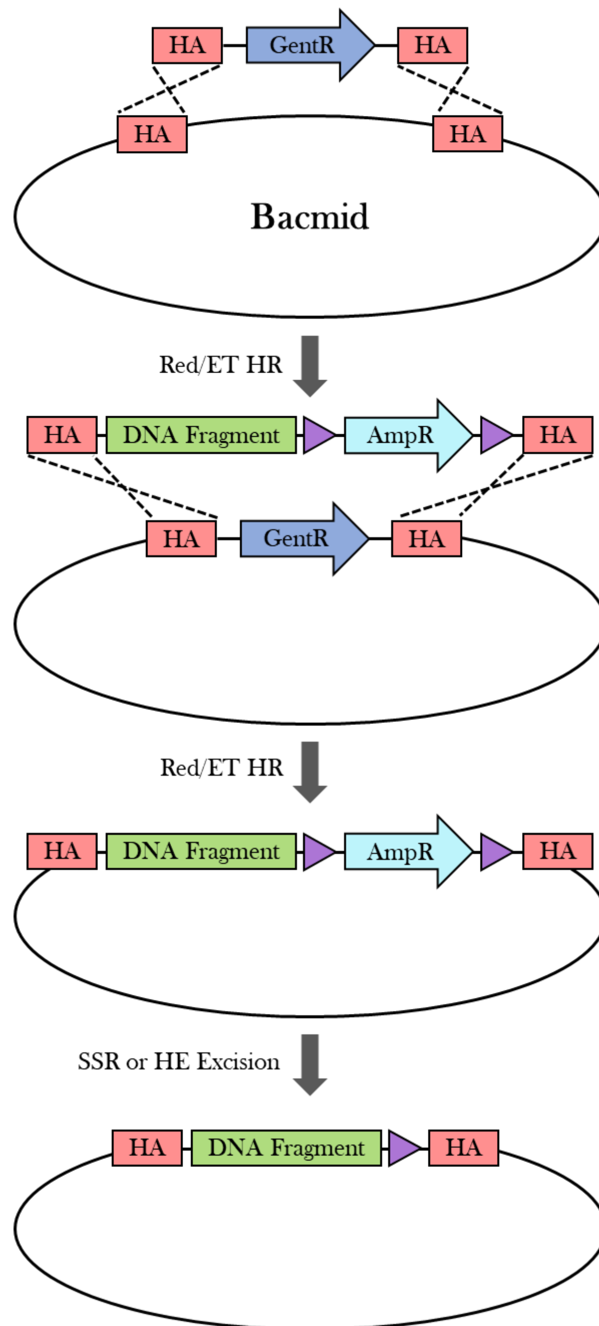


Figure 3.2 – Experimental procedure for the modification of the baculovirus genome.

A schematic representation of the three main steps undertaken during the fragment modifications in the generation of SynBac. Initially, the native fragment of DNA is replaced with a resistance marker, gentamicin (GentR), using Red/ET homologous recombination (Red/ET HR), with the homology arms (HA) acting as the junctions for the modification. A second Red/ET reaction replaces the GentR marker with a synthetic fragment of DNA that encompasses the essential DNA (green box), a different antibiotic marker, AmpR, and recognition sites for site-specific recombinases (SSR) or homing endonucleases (HE) (purple triangles). The SSR or HE sites are then utilised in the last step to extract the antibiotic resistance marker, ensuring the modification process can be performed again without accumulation of resistance markers.

3.1.3 First Genomic Modification

Prior to the start of this project, SynBac1 had already been generated through the replacement of Fragment 1 using the Red/ET recombination method and Cre/LoxP recombination. Fragment 1, the largest planned deletion, removed a total of 10,805 bp, thereby reducing the fragment from 24,479 bp to 13,674 bp. Combined with the genomic deletion of the non-essential genes within Fragment 1, an EYFP reporter gene was inserted. The incorporation of a reporter gene had previously proven to be highly beneficial in the real-time tracking of recombinant protein production by the routinely used EMBacY system. However, the EYFP reporter gene in EMBacY is driven by the late phase polyhedrin promoter. Conversely, the chosen promoter for the EYFP within Fragment 1of SynBac was the ie1 promoter. Ie1 is a gene that is activated earlier in the baculoviral life cycle, thus the EYFP would also be expressed earlier. Previous members of the Berger lab validated SynBac1 in terms of the virus amplification kinetics and protein expression capabilities. An SDS-PAGE gel (Figure 3.3) displays the expression of the test protein GDAP288 from the precursor genome, MultiBac, and the modified genome, SynBac1. A comparable yield of protein production can be observed for both viruses. The positive protein expression validated the modification of Fragment 1 and confirmed the predicted non-essentiality of the genes that were deleted within the fragment.

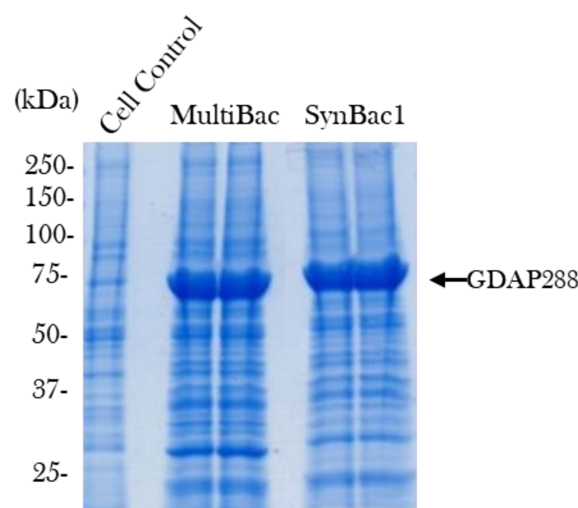


Figure 3.3 – Fragment 1 validation.

12% SDS-PAGE analysis of the protein expression capabilities of MultiBac and SynBac1 (Fragment 1 exchanged from starting genome, MultiBac) displaying the presence of a band around 75 kDa representing the test protein, GDAP288. Image kindly provided by Fred Garzoni.

3.1.4 Generation of Tools and Methods Required for Fragment Integration

While Fragment modifications 1, 5 and 6 utilise the Cre/LoxP system, PI-PspI and I-SceI homing endonucleases, respectively, to remove the antibiotic marker from the remaining Fragments (2A, 2B, 3 and 4), a different method for the removal of the resistance marker was required. The existence of other site-specific recombination systems similar to Cre/LoxP recombination have been discovered through homology alignments of P1-like phages.[273,274] These include Dre/RoxP, Vika/Vox, VCre/VloxP and SCre/SloxP (Table 3.1). They all share the quality of acting on sites with palindromic repeats and spacer regions. All apart from the Dre/RoxP system recognise 34 bp regions with 8 bp spacer, whereas Dre recognises a 32 bp site with a 4 bp spacer. Numerous studies have previously demonstrated the absence of cross-talk between the different recombinases and the recognition sites.[273-276] The high specificity of these systems suggest that they may be ideal for DNA engineering applications, where they can be used in parallel with no hesitation about cross-reactivity. The recombinases and their corresponding target sequences mentioned above, will be utilised in the generation of the synthetic baculoviral genome, SynBac. The use of these sites in the remaining fragments will enable a virtually scar-less genome engineering process, as the antibiotic resistance markers used for selection can be subsequently removed. Following each fragment modification, only one of the recognition sites will remain in the genome, which can be utilised for future insertions of any auxiliary DNA sequences with desirable functions.

Recombination System	Recognition sequence *5'-3'*
Cre/LoxP	ataacttcgtata**GCATACAT**tatacgaagttat
VCre/VLoxP	tcaatttctgaga**ACTGTCAT**tctcggaaattga
SCre/SLoxP	ctcgtgtccgata**ACTGTAAT**tatcggacatgat
Vika/Vox	aataggtctgaga**ACGCCCAT**tctcagacgtatt
Dre/RoxP	taactttaaataat**GCCA**attatttaaagtta

Table 3.1 Site specific recombinases and their recognition sequences.

The spacer region is in **bold**.

Once the minimal synthetic baculoviral genome, SynBac, has been produced and the different recognition sites for the site-specific recombination have been added, the use of these sites will require validation. In order to achieve this, new acceptor and donor plasmids that contain the four different recognition sites are required. Likewise, their corresponding recombinase enzymes need to be expressed and purified. The acceptor plasmids will be based upon pACEBac1, however containing the different recognition sites in the place of LoxP. The synthesis of these plasmids will be carried out by GenScript Biotech. The donor plasmids all require different antibiotic resistance markers if they are to be tested collectively in the SynBac genome. Resistances that are not already present in the bacmid or the helper plasmid have been chosen. The donor plasmids are to be generated based upon pUCDM, a donor from the MultiBac system. The LoxP site will be exchanged with one of the four recognition sites and the existing antibiotic resistance marker (chloramphenicol) exchanged. To confirm the functionality of the recognition sites, a measurable readout needs to be incorporated. Therefore, fluorescent protein markers are to be inserted into the donor plasmids so that their expression can be monitored through fluorescence microscopy and fluorimetry. Refer to (Table 3.2) for the planned donor plasmids and their corresponding antibiotic marker, site-specific recombination system and fluorescent marker.

Test colour plasmid	Antibiotic Marker	Site-specific recombination system	Fluorescent marker
pUADM_RoxP_AmCyan	Ampicillin	Dre/RoxP	AmCyan
pUBDM_SLoxP_mTagBFP	Blasticidin	SCre/SLoxP	mTagBFP
pUPDM_VLoxP_EGFP	Puromycin	VCre/VLoxP	EGFP
pUZDM_Vox_mCherry	Zeocin	Vika/Vox	mCherry

Table 3.2 Colour donor plasmids and their assigned antibiotic marker, site-specific recombination system and fluorescent marker.

This chapter encompasses a collaborative project, thus multiple co-workers participated in its progression. Where applicable, the work undertaken by others as part of this project is appropriately indicated. The fragment modifications are not described in numerical order but detailed in chronological order. This progression was chosen based on the ease of fragment modification. The mutant genome nomenclature consists of SynBacX, SynBacΔX, SynBacXΔY and SynBacXAmp, where X represent the fragment(s) modified since original MultiBac genome, Δ indicates an intermediate stage where the native DNA has been replaced with a fragment containing a resistance marker, XΔY signifies fragments X and Y have been modified but the resistance marker remains within fragment Y, and XAmp indicates the presence of the ampicillin marker within the X fragment, which has not yet been excised from the baculoviral genome.

The results presented within this thesis utilise the use of standard error or mean (SEM) as error bars within all the fluorescence quantification data. SEM of a statistic is the standard deviation of its sampling distribution, or an estimate of that standard deviation. SEM was chosen in this thesis due to the smaller error bars displayed in the bar graphs, as a way to maximise area of the bar graph displayed. In hindsight, the use of true standard deviation (SD) should have been considered as SEM underestimates the standard deviation, particularly when the sampling population is low. Additionally, where a statement is made suggesting a particular SynBac version is better/comparable/displays similar levels of fluorescence to the control virus within that experiment, the veracity of the statement is subject to performing a statistical analysis.

3.2 Chapter Aims

The aim of this chapter is to generate a minimal synthetic baculoviral genome utilising the primarily proposed iterative method. This will be explored through the following approaches:

1. Removal of genes considered non-essential for laboratory cell culture in the form of fragments containing multiple genes.
2. Generation of tools and methods required for the fragment engineering and validation of final SynBac genome.

3.3 Results

3.3.1 Iterative Fragment Removal

3.3.1.1Confirmation of Starting Material

The first SynBac prototype was previously generated in the Berger Group by replacing wild-type DNA sequences with synthetic Fragment 1, generating SynBac1. SynBac1 was then used to generate SynBac1Δ5 (Fragments 1 and 5 modified) and SynBac1Δ6 (Fragments 1 and 6 modified), with both genomes still containing the gentamicin marker utilised in the removal of the fragments. While the validation of the SynBac1 genome has been confirmed through protein expression screening (Figure 3.3), the data for SynBac1Δ5 and SynBac1Δ6 had not yet been collected. The two mutant genomes were subject to examination by protein expression of a test protein, HexaHis_GDAP288 (72.6 kDa). The testing of SynBac1Δ5 was carried out by myself in the context of this thesis, whereas SynBac1Δ6 testing was undertaken by a colleague and will not be detailed here.

Using the standard methods, pFB-GDAP288-pDS was used to generate recombinant bacmids, which were prepared and transfected in Sf21 cells. The EYFP protein expression levels were measured after the V0 amplification on the 6 well plates and the best clones were chosen for V1 amplification. The protein expression capability of SynBac1 (A) and SynBac1Δ5 (B) was compared using SDS-PAGE and WB analyses (Figure 3.4). Here, samples were collected from a V1 amplification at DPA+24h and comparable amounts of GDAP288 protein were detected for both baculoviruses. However, it should be noted that the time taken to reach DPA differed between the two baculovirus genomes: 72hpi for SynBac1 and 120hpi for SynBac1Δ5.

Amongst the genes deleted in Fragment 1 was the strongly expressed baculoviral protein, P10 (10kDa). This was used to unambiguously distinguish between the original genomes such as MultiBac and EMBacY, which express P10, and the recently created SynBac1 and SynBac1Δ5 mutants, which do not. Samples from the four baculoviral genomes were collected at the V1 amplification stage and resolved on an 18% SDS-PAGE (Figure 3.4 (C)). The bands present at approximately 10 kDa in the MultiBac and EMBacY samples, in comparison to the SynBac1 and SynBac1Δ5, confirm lack of the P10 protein and therefore verify the correct genomes were provided.

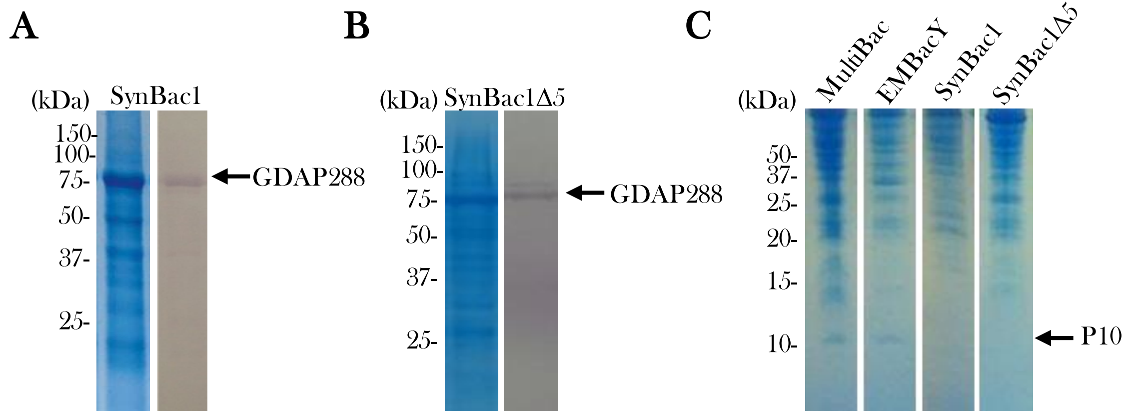


Figure 3.4 – Inspection of starting baculovirus genomes.

(A-B) 12% SDS-PAGE (left) and WB (right) of SynBac1 **(A)** and SynBac1Δ5 **(B)** demonstrating the presence of a GDAP288 test protein band situated approximately at 75 kDa. Protein band detected in WB analysis using Anti-polyHistidine-Alkaline Phosphatase antibody. **(C)** 18% SDS-PAGE of MultiBac, EMBacY, SynBac1 and SynBac1Δ5 validating the removal of fragment 1 through the absence of the P10 protein in the SynBac1 and SynBac1Δ5 samples. (A-C) 1 x 10^6 cell samples were collected at DPA+24 during V1 amplification.

3.3.1.2 SynBac16

Since the quality control of the starting baculoviral genomes was confirmed, the first modifications could be performed. As Fragment 1 utilised the Cre/LoxP system for the removal of the antibiotic marker that leaves behind a Cre site, different methods for removal of the subsequent fragments were required. Homing endonucleases were chosen for Fragments 5 and 6, with Fragment 6 already containing the PI-PspI sites. Due to the presence of the PI-PspI sites, the genome of SynBac1Δ6 was taken and the GentR removed to generate SynBac16.

3.3.1.2.1 Helper Plasmid and SynBac1Δ6 Isolation

Prior to the removal of the GentR from SynBac1Δ6, a protocol for the isolation of the bacmid from the pHelper was required. The removal of pHelper was crucial to facilitate subsequent modifications without the presence of the plasmid, which could influence the efficiency of the engineering due to the high copy nature of the plasmid. Bacmid preparations of SynBac1Δ6 were performed according to protocol detailed in Methods section, desalted and transformed into electrocompetent DH10β cells. The recovered reaction was split and plated on two sets of plates (tetracycline, IPTG, BluOGal and kanamycin, IPTG, BluOGal). A schematic representation for the isolation of the pHelper and bacmid can be seen in Figure 3.5 (A). The colonies representing pHelper should be white as they are not able to process the BluOGal

present in the agar plates. The bacmid carries the LacZ gene that is activated by IPTG to express β-galactosidase, which processes the BluOGal, thus blue colonies are expected. The colonies from both sets of plates need to be restreaked to confirm only the presence of pHelper (tetracycline+, kanamycin-) and bacmid (kanamycin+, tetracycline-) alone. This was particularly important for the isolation of the bacmid as the chances of the pHelper transforming alongside the bacmid were higher due to the smaller plasmid size. The colour of the colonies representing the bacmid was additionally crucial, as the presence of white colonies would indicate that damage to the baculoviral vector has occurred within the LacZ gene.

After carrying out the procedure, plentiful colonies were found on both sets of plates, with restreaked colonies only growing on antibiotics corresponding to the resistance found on the pHelper (tetracycline) and bacmid (kanamycin). A prep of the pHelper was performed and restriction mapping performed to confirm that the correct plasmid was present. The pHelper would be used in subsequent experiments for transformation into newly generated SynBac variants. Bacmid preparations from both sets of plates were performed to visualise only pHelper and bacmid alone were present in the cells. Figure 3.5 (B) shows successful isolation of the SynBac1Δ6 bacmid (blue arrow) and no pHelper present in the preparation. The pHelper sample displays only the corresponding bands to pHelper (red arrow).

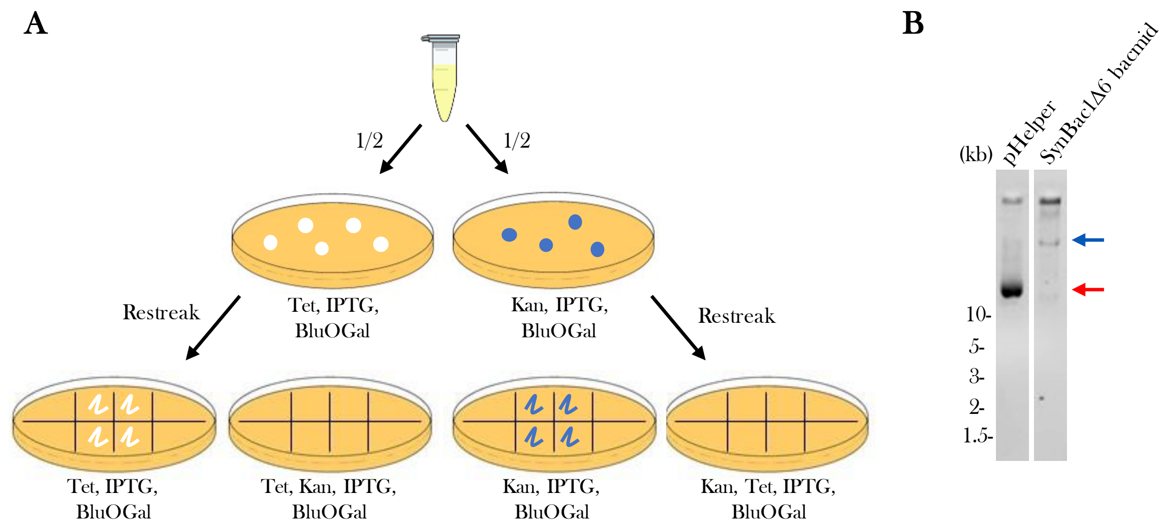


Figure 3.5 - SynBac1Δ6 and pHelper isolation.

(A) Schematic representation of SynBac1.1Δ6 and pHelper isolation post electroporation. Half of the recovered cells were plated on tetracycline, IPTG, BluOGal plates and half on kanamycin, IPTG, BluOGal plates. The colonies were restreaked to confirm only pHelper and SynBac1.1Δ6 bacmid presence. **(B)** DH10β isolated SynBac1Δ6 and pHelper bacmid preparations. Blue arrow- SynBac1.1Δ6 bacmid, red arrow- pHelper (≈13 kb).

3.3.1.2.2 SynBac16 Generation and Validation

The gentamicin marker from the isolated SynBac1Δ6 bacmid could then be removed to generate SynBac16. Homing endonuclease restriction enzyme sites (PI-PspI) were previously cloned on either side of the GentR and inside of the homology arms (Figure 3.6 (A)). This fragment was used to generate SynBac1Δ6, the starting genome for the following modification, by replacing the native piece of DNA with the gentamicin containing fragment.

An increased volume bacmid preparation of the SynBac1Δ6 was performed. 2 x 10 ml cultures were resuspended in 900 µl of buffer P1, split into 3 eppendorf tubes (6 total) and subsequently treated as usual. All 6 tubes were resuspended collectively in 20 µl water. A 30 µl reaction with 20 µl bacmid, 1.5 µl PI-PspI, 10X PI-PspI buffer, 10X BSA) and water was set up and incubated at 65 °C overnight. The restriction reaction was dialysed, ligated using T4 DNA ligase, electroporated into DH10β cells and plated on Kan, IPTG, BluOGal plates. The following day, numerous colonies were present and restreaked on Kan, IPTG, BluOGal and Gent, IPTG, BluOGal plates. All but one of the 40 restreaked colonies did not grow on Gent plates, confirming the successful removal of the gentamicin marker and re-ligation procedures. Preceding pHelper transformation and competent cell preparation, several clones were picked, prepped and transfected into Sf21 cells to check for the presence of EYFP. The number of clones displaying positive EYFP absorbance was low. However, clones showing EYFP were present and the best one was picked for pHelper transformation. Multiple clones were additionally checked at this stage before competent cell preparation and transformation of the test protein, GDAP288. An SDS-PAGE gel, (Figure 3.6 (B)) was used to detect the expression of the GDAP288 test protein for SynBac1 and SynBac16 samples collected at DPA+48h. The EYFP data, (Figure 3.6 (C)) for the best clones from V1 amplification of SynBac16 Precursor, SynBac16, SynBac16 GDAP288 and SynBac1 illustrate that the SynBac16 variants were functional, however the EYFP fluorescence was slightly lower at DPA and lower still at DPA+24h when compared to SynBac1. In addition, the day to reach proliferation arrest had increased to 120 hpi, in contrast to 48-72 hpi for SynBac1.

The SynBac16 genome was then used to apply the next modification, Fragment 5, which utilised a different homing endonuclease site, I-SceI. The generation of SynBac165 was performed by a colleague and will not be detailed. Furthermore, the optimisation of the transfection protocol was performed by a colleague, where the volume of overnight culture and the amount of transfection reagent were explored in the hopes of improving the virus amplification. Even though the SynBac16 and SynBac165 variants previously displayed a slower time to reach

proliferation arrest, the application of the subsequent fragment went ahead due to the moderate improvement of the transfection protocol.

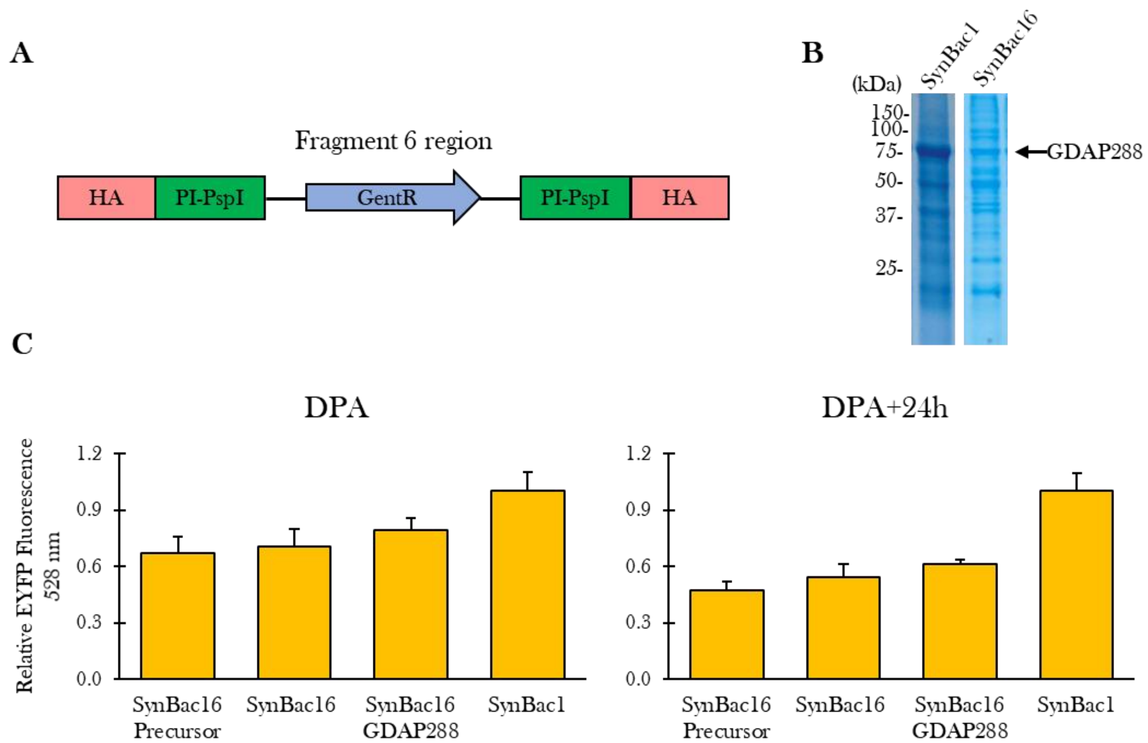


Figure 3.6 - SynBac16 generation and validation.

(A) Baculoviral region representing the replaced Fragment 6 with a gentamicin marker. Homing endonuclease (PI-PspI) sites flanking both sides of the gentamicin gene (GentR) inside the homology arms (HA). **(B)** 12% SDS-PAGE of SynBac1 and SynBac16 demonstrating the presence of GDAP288 test protein band situated approximately at 75 kDa. Samples collected at DPA+48h. **(C)** Average EYFP production detected from SynBac16 Precursor (no pHelper), SynBac16, SynBac16 GDAP288 and SynBac1 from 1,000,000 cells collected at DPA and DPA+24h. Values represent averages normalized to SynBac1 = 1; error bars indicate SEM (n=3).

3.3.1.3 Generation of Tools and Methods Required for Further Fragment Integration

While the previous Fragment modifications (1, 6 and 5) utilised the Cre/LoxP system, PI-PspI and I-SceI homing endonuclease, respectively, to remove the resistance marker, a different method for the removal of the antibiotic marker from the remaining Fragments (2A, 2B, 3 and 4) was required. The addition of the different approaches for each fragment would also give rise to a baculoviral genome that contained numerous DNA insertion sites for potential future engineering. As previously described (Section 3.1.4), in addition to the Cre/LoxP system, the

existence of other site-specific recombination systems has been discovered. These include Dre/RoxP, Vika/Vox, VCre/VloxP and SCre/SLoxP.[273,274] The lack of cross-reactivity between these SSRs has been previously demonstrated. Therefore, these sites could appropriately be utilised for the remaining fragment modifications.

3.3.1.3.1 Recombinase Purification and Validation

The alternative recombinase enzymes are not commercially available and thus, prior to their use in the engineering of the bacmid genome, the expression and purification of each recombinase protein was required. Constructs were designed to contain the genes encoding for the recombinases in a pet28+ bacterial expression vector, with expression inducible by IPTG, an N-terminal DecaHis tag and a TEV cleavage site for ease of purification. The establishment of the purification protocol was carried out by a colleague (H. Crocker) and the subsequent purifications of the enzymes was a collective effort including other lab members. The purified proteins (ranging from 0.5-1.3 mg/µl) can be seen on the SDS-PAGE gel, (Figure 3.7 (A)).

The recombinases to be used were first tested with modified plasmids (pDS-Frag4-RoxAmpRox, pDS-Frag3-VoxAmpVox, pDS-2A-VloxPAmpVloxP and pIBH2-2B-SLoxPAmpSLoxP) that contained ampicillin flanked by two of the site-specific recombination sites (RoxP, Vox, VloxP and SLoxP, respectively). 400-600 ng of the different plasmids were mixed with 2-6 µl of their corresponding recombinases (Dre, Vika, VCre and SCre, respectively) and 3 µl NEB 3.1 buffer in a 30 µl reaction. A negative control with no recombinase was also set up. First, the activity of Dre recombinase at 37 ° C was assessed with a time course experiment. Here, 10 µl samples were taken after 30 minutes, 1 hour and 2 hours and heat inactivated immediately for 20 minutes at 70 ° C. The samples were subsequently digested with XhoI and analysed by agarose gel electrophoresis. The results of this time course (Figure 3.7 (B)) show the presence of a 1143 bp band corresponding to the ampicillin fragment already at 30 minutes, with the 1 hour sample possibly showing a stronger band corresponding to the ampicillin fragment. From the time course of Dre activity, it was decided that all future reactions including the four different recombinases would be incubated for 1 hour. The samples for all the recombinases were set up as previously described with an hour incubation, digested (XhoI for Dre/RoxP, Vika/Vox, SCre/SLoxP and Bsu36I for VCre/VloxP) and the samples run on an agarose gel to see the excised ampicillin marker (approx. 1.1 kb), (Figure 3.7 (C)). The samples with no added recombinases did not show the band corresponding to the size of the ampicillin marker.

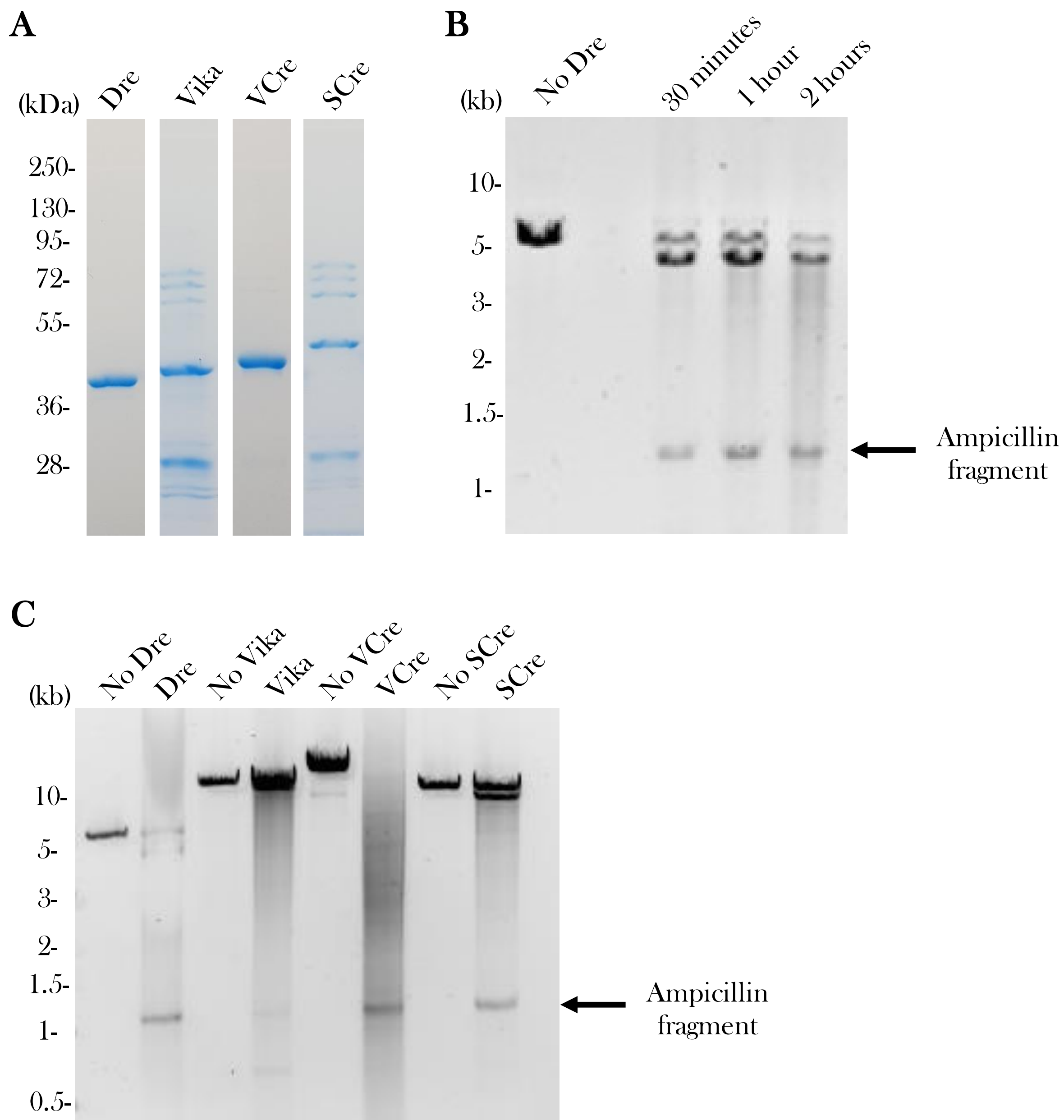


Figure 3.7 - Purified recombinase enzymes and their activity.

(A) 12% SDS-PAGE of the purified recombinase enzymes Dre (42 kDa), Vika (44 kDa), VCre (45.4 kDa) and SCre (47 kDa). Approximately 2-3 mg of protein was loaded. **(B)** 1% agarose gel showing the time course results for Dre recombinase activity to remove the ampicillin (1143 bp) cassette found between two RoxP sites. **(C)** 1% agarose gel demonstrating the recombinase activity of all 4 enzymes. Plasmids containing an ampicillin cassette flanked by two site specific recombination sites were incubated with their corresponding recombinases resulting in the excision of the ampicillin fragment visualised around 1 kb (arrow). The following restriction pattern was expected; Dre – 5568, 4425, 1143 bp; Vika- 10182, 9104, 1078 bp; VCre- 7649, 6571, 1078 bp; SCre- 8205, 7127, 1078 bp.

3.3.1.3.2 Generation of Fluorescent Donor and Acceptor Plasmids

Preceding the use of the different site-specific recombination systems on the bacmid, the activity and efficiency of the method needed to be tested. To be able to carry this out, we had the acceptor plasmid, pACEBac1 synthesised to contain the RoxP, SLoxP, VLoxP and Vox sites. The donor plasmids were modified in-house based on pUCDM (collective effort between other lab members, see (Table 8.1) in appendix for which constructs were generated in this study), to contain an alternative resistance marker and each of the different recombinase recognition sites. The four different donor plasmids generated can be visualised as a schematic in Figure 3.8 (A).

Firstly, pUCDM would be modified by exchanging the resistance marker, generating pUADM, pUBDM, pUPDM and pUZDM (ampicillin, blasticidin, puromycin and zeocin, respectively). The pUDM backbone and the resistance marker was amplified using primers that contained overhangs complementary to the pUCDM backbone (Tale 8.2 in Appendix). The fragments were ligated together through a Gibson reaction and plated on corresponding antibiotic resistance plates. Clones were checked for the correct ligation product through restriction digestion of the original pUCDM and modified plasmids.

The next step included modifying the LoxP site for the 4 different recombinase recognition sites (RoxP, SLoxP, VloxP and Vika). Phosphorylated primers were designed to amplify the backbone, however omitting the LoxP site, and the different recombinase recognition sites were inserted as an overhang on one of the primers. To be able to distinguish between the original plasmid containing the LoxP site and the different sites, a unique restriction site, EcoRV, was inserted as the other primer's overhang. The PCR reactions were analysed by an agarose gel electrophoresis, the PCR amplicon extracted, and the ligation reaction initiated with 200-300 ng of the gel extracted DNA, 1 µl T4 Ligase and 1 µl T4 ligase buffer. The reaction was incubated at 22 ˚C for 1 hour, transformed and plated on corresponding agar plates. Clones were checked through restriction digestion using the inserted EcoRV site.

Subsequently, the fluorescent proteins (AmCyan, mTagBFP, EGFP and mCherry) were PCR amplified. The previously modified plasmids: pUADM_RoxP, pUBDM_SLoxP and pUZDM_Vox were cut with StuI and pUPDM_VLoxP was cut with SmaI. The backbones were gel extracted and Gibson reactions set up with their corresponding fluorescent protein. Clones were miniprepped and checked by restriction digestion for correct size after fluorescent marker addition. The fluorescent protein encoding plasmids were additionally sent for sequencing to confirm correct plasmid generation.

Lastly, the donor plasmids were fused with their corresponding acceptors, i.e., pUADM_RoxP_AmCyan with pACEBac1_RoxP and so on for SLoxP, VLoxP and Vox containing pairs. The recovered cells were plated on agar plates containing their corresponding antibiotics and gentamicin (pACEBac1). Clones were miniprepped and checked by restriction digestion for the correct fusion product. Fusion maps were generated using the Cre ACEMBLER software by exchanging the specific sites for LoxP. To verify that there was no cross-reactivity between the different sites and their corresponding recombinases, controls using Cre recombinase were set up with the different site-specific plasmid pairs. A LoxP containing donor and acceptor were also tested with 4 of the different recombinases. None of the control reactions yielded colonies confirming no cross-talk between the different sites and recombinases. Each of the donor (~3.3-3.8 kb), acceptor (~2.9 kb) and fusion (~6.3-6.8 kb) plasmids were cut once to illustrate the size of the linear products, (Figure 3.8 (B)).

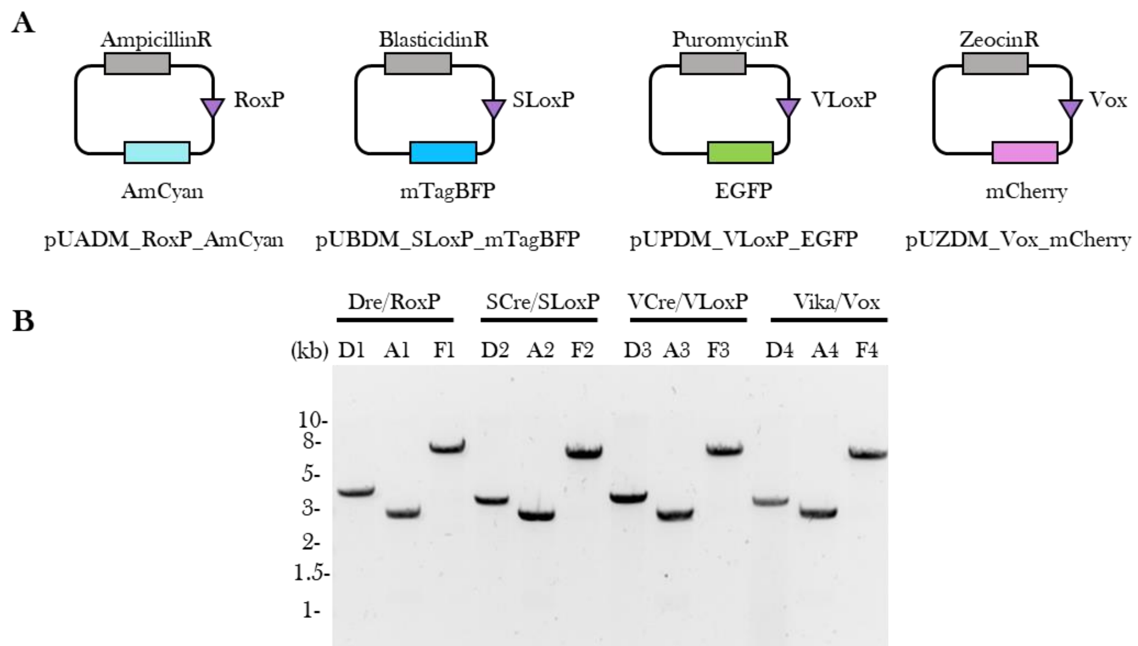


Figure 3.8 - Fluorescent donor plasmids and their fusions.

(A) Schematic diagram showing the generated fluorescent donor plasmids, containing a different resistance marker (ampicillin, blasticidin, puromycin and zeocin), site-specific recombination site (RoxP, SLoxP, VloxP and Vox) and fluorescent protein (AmCyan, mTagBFP, EGFP and mCherry). **(B)** 1% agarose gel of the fusions of donor and acceptor plasmids *in vitro* via different site-specific recombinases and their recognition sites. The donor (D) and fusion (F) plasmids were analysed using NheI restriction, and acceptor (A) plasmids using BamHI restriction.

3.3.1.3.3 Validation of Donor Plasmids

Following the production of the fusion plasmids, their ability to express the fluorescent proteins was tested using EMBacY. The plasmids were transformed into EMBacY containing cells by Tn7 transposition and triplicates of four of the different fluorescent expressing fusions were transfected into Sf21 cells. Samples were taken daily during V1 amplification to track the EYFP (EMBacY bacmid) and fluorescent protein (donor plasmids) expression. The fluorescence values of the V1 samples taken 3 days post infection were plotted to show their relative fluorescence units. Images of the Sf21 cells were taken on the DMI6000 Leica confocal microscope 4 days post transfection of the colour fusion plasmids to illustrate the expression of the fluorescent markers, (Figure 3.9). All the tested colour fusion plasmids demonstrated successful fluorescent marker expression in the EMBacY system.

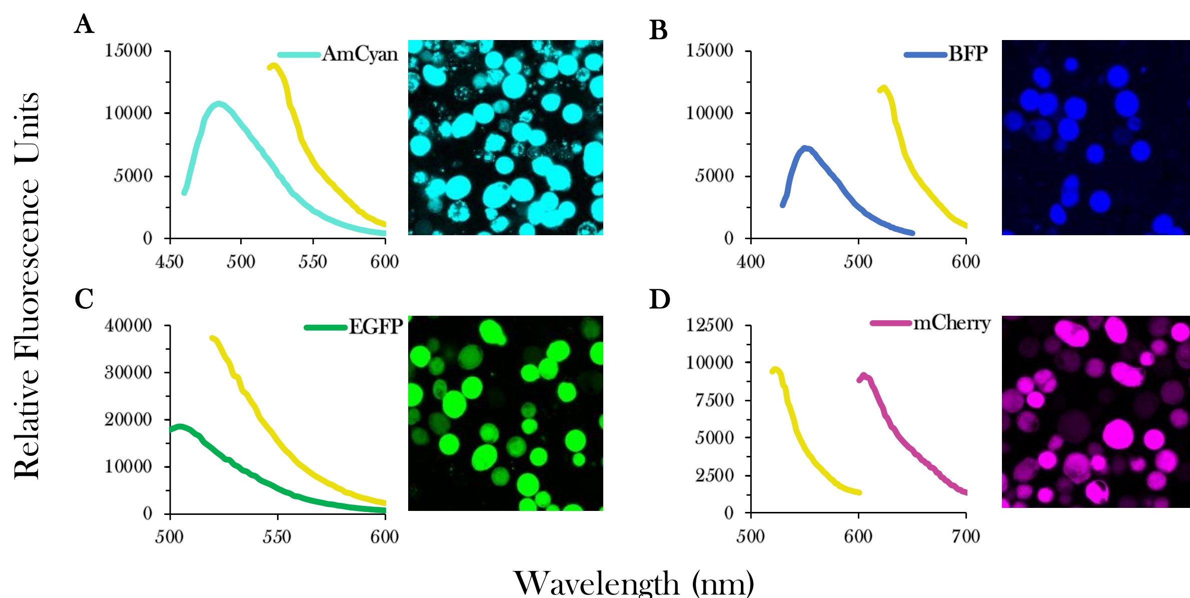


Figure 3.9 - Fluorescent protein expression in EMBacY transfected Sf21 cells.

EMBacY containing cells were transformed with four of the different colour plasmid fusions, transfected and V1 amplified. EYFP (yellow plot), AmCyan, mTagBFP, EGFP and mCherry values were measured from 1 million cell samples collected 3 days post infection of V1, according to their excitation and emission spectra, (Table 2.4, Materials and Methods). Images of the Sf21 cells were taken on the DMI6000 Leica confocal microscope 4 days post transfection. Briefly the cells were resuspended from the 6 well transfection plate, diluted in PBS and approximately 70,000 cells in 300 µl total volume transferred to 8 chamber Nunc™ Lab-Tek™ II chamber slides (Thermo Fisher). 63X magnification, oil immersion objective used. **A-** pUADM_RoxP_AmCyan*pACEBac1_RoxP, **B-** pUBDM_SLoxP_mTagBFP*pACEBac1_SLoxP, **C-** pUPDM_VLoxP_EGFP*pACEBac1_VLoxP, **D-** pUZDM_Vox_mCherry*pACEBac1_Vox

3.3.1.4 SynBac1654

3.3.1.4.1 Fragment 4 Modification

Prior to applying the next alteration to the SynBac165 genome, the fragment under scrutiny, Fragment 4, needed to be modified. Our strategy for exchanging the fragments has been updated by using differing types of recombinases and their corresponding recognition sites. The RoxP site and Dre recombinase were assigned to Fragment 4.

pDS-Frag4, the plasmid containing the minimised Fragment 4 was modified by adding an ampicillin resistance marker, for which one can select after the Red/ET recombination, and two flanking RoxP sites that allow the subsequent removal of the ampicillin marker by Dre recombinase. The cloning strategy was designed to amplify the AmpR gene with RoxP sites as overhangs on the primers, which would be complementary to the overhangs on the pDS-Frag4 amplified fragment. Figure 3.10 (A) shows the 1142 bp ampicillin fragment (RoxPAmpRoxP) amplified from pUC57 and the 4386 bp backbone from pDS-Frag4. The primers have been designed to include unique restriction sites, NheI and NotI. The enzyme restriction sites were used to generate the new modified pDS-Frag4RoxPAmpRoxP through restriction-ligation of the two PCR products. The PCR products were subject to restriction digestion and a 2:1 ratio of insert to vector with 400 ng of vector was combined with 2 µl T4 Ligase and 3 µl T4 Ligase buffer in a 30 µl reaction. The ligations were incubated at room temperature for 2 hours and transformed. Clones were miniprepped and checked for the correct ligation product through restriction digestion of the original pDS-Frag4 and pDS-Frag4RoxPAmpRoxP for which a larger band would be observed due to the presence of the additional ampicillin containing fragment. Figure 3.10 (B) displays the pattern seen after the restriction digestion, with pDS-Frag4 cut once with BamHI (4386 bp) and cut twice with NdeI (2485+1901 bp) and pDS-Frag4RoxPAmpRoxP cut once with BamHI (5568 bp) and cut twice with NdeI (3667+1901 bp).

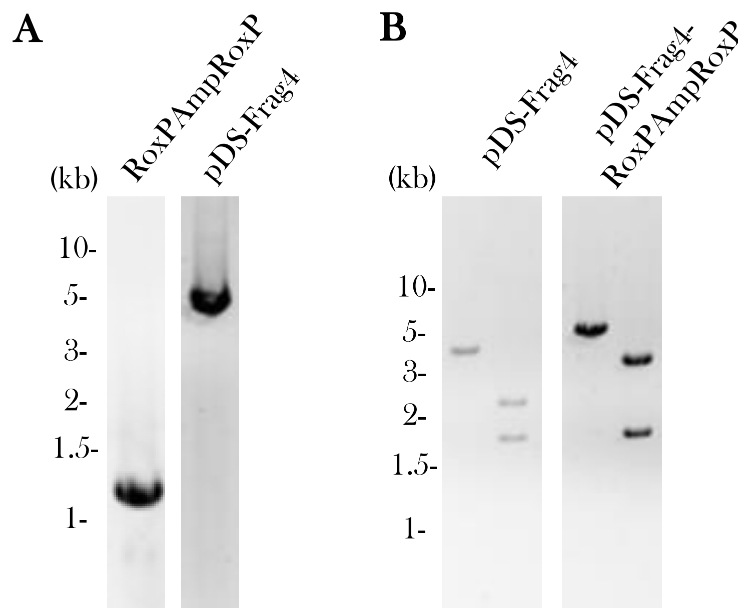


Figure 3.10 - Fragment 4 modification.

(A) 1% agarose gel of the PCR amplified products of RoxPAmpRoxP (1142 bp) and pDS-Frag4 backbone (4386 bp). **(B)** 1% agarose gel of the restriction pattern of pDS-Frag4 and pDS-Frag4-RoxPAmpRoxP after linearization with BamHI (4386 bp and 5568 bp, respectively) and double cut with NdeI (2485+1901 bp and 3667+1901 bp, respectively). The band shift can be visualised for the modified pDS-Frag4-RoxPAmpRoxP compared to the original pDS-Frag4.

3.3.1.4.2 SynBac1654 Generation and Validation

SynBac1654 was generated using the Red/ET recombination method with SynBac165 as the precursor genome and PCR amplified fragments, HA-GentR-HA and HA-Frag4-HA (Figure 3.11). Firstly, the original Fragment 4 in the baculoviral genome was removed and replaced by a fragment harbouring the gentamicin resistance, HA-GentR-HA. The procedure was performed with SynBac165 precursor cells transformed with pRedET. The recovered cultures were plated on Kan, Gent, IPTG, BluOGal plates to select for successful recombination, thus removal of the endogenous Fragment 4 and insertion of the gentamicin containing fragment. The generated SynBac165Δ4 was then taken through the Red/ET recombination procedure again with the modified Fragment 4 containing an ampicillin marker, HA-Frag4-HA. The recovered cultures were plated on Kan, Amp, IPTG, BluOGal plates. The colonies were restreaked on gentamicin containing plates to check for complete removal of the gentamicin containing fragment. A bacmid prep from the generated SynBac1654Amp clones was prepared and 20 μl was incubated with 3 μl 3.1 NEB buffer, 3 μl Dre recombinase and 4 μl water, at 37°C for 1 hour. The reaction was subsequently dialysed for one hour and 10 μl electroporated into DH10β cells. The

recovered cells were plated on Kan, IPTG, BluOGal plates. The colonies were restreaked on Kan only and Kan, Amp plates to confirm ampicillin resistance loss.

To confirm that the correct baculoviral genome was obtained, PCR reactions that amplify the region of modification were performed at each step, SynBac165, SynBac165Δ4, SynBac1654Amp and the final SynBac1654 genome. The amplified fragments can be visualised at approximately 5.9kb, 1.1kb, 3.3 kb and 2.2 kb, respectively (Figure 3.11(C)). Several clones of the SynBac1654Amp and SynBac1654 genome were chosen, prepped and transfected into Sf21 cells, alongside the control genome SynBac1, for measurement of relative EYFP fluorescence. To better understand the virus amplification kinetics compared to the previous SynBac16 validation where only the DPA and DPA+24h samples were illustrated, samples were taken during V1 amplification at 24 hour intervals until 168 hpi for SynBac1654Amp and SynBac1654, and 72 hpi for SynBac1. The reason for including the SynBac1654Amp variant in the testing was to determine if the Dre recombinase treatment to remove the ampicillin marker resulted in any damage to the bacmid elsewhere. The average EYFP fluorescence for the three SynBac mutants shown in Figure 3.11 (D), illustrates that the SynBac1654 mutants reach similar levels of EYFP fluorescence. However, the time taken for the cultures to reach proliferation arrest and start expressing the EYFP is extended when compared to the SynBac1 control. On average the SynBac1 V1 cultures reach proliferation arrest at 48 hpi, whereas the SynBac1654 mutants required 144-168 hpi.

Due to the slow virus amplification observed in the case of SynBac1654, the remaining fragments were not modified and the approach to the project had to be re-evaluated. In parallel to the SynBac1654 validation, virus kinetics experiments were performed by a colleague (H. Crocker) on the genomes of SynBac16 and SynBac165. The viruses similarly displayed an extended timescale to reach DPA and express protein. SynBac165 was found to be slower than SynBac16, suggesting that the additional fragment removal negatively affected the virus. Interestingly, the SynBac165 genome required 144-168 hpi to reach comparable protein amounts to the control, as did the SynBac1654 genome validated in this thesis. Thus, the virus kinetics appear comparable between SynBac165 and SynBac1654, suggesting that removal of Fragment 4 alone, may not negatively affect the baculoviral genome. The project tactic was subsequently changed to revert to the SynBac1 genome as the starting genome and modify the remaining fragments separately. Fragments 3 and 4 were assigned to myself, whereas Fragments 2A and 2B were implemented by a colleague and are detailed elsewhere (Dr Hannah Crocker, PhD Thesis).

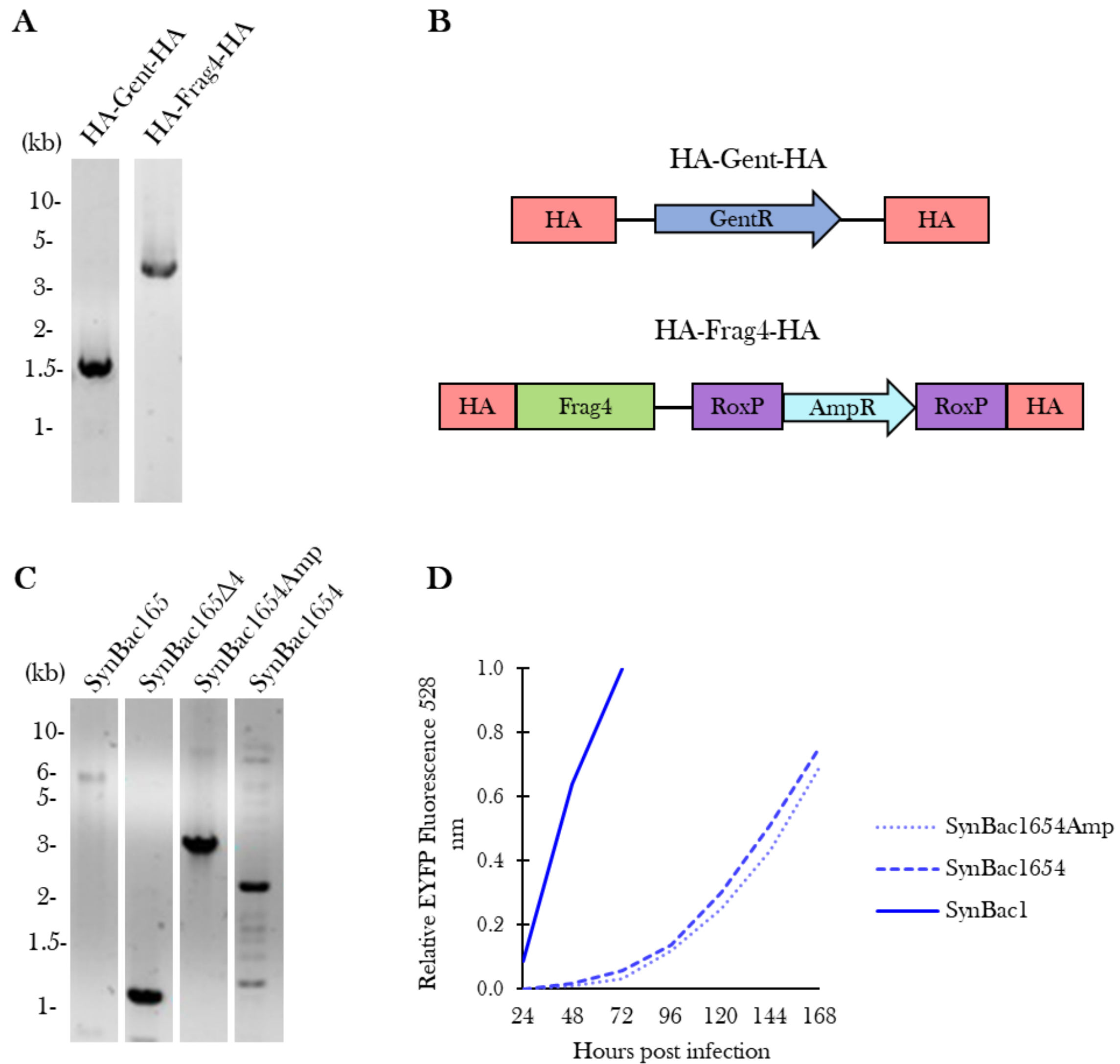


Figure 3.11 - SynBac1654 generation and validation.

(A) 1% agarose gel of PCR amplified HA-Gent-HA fragment (1480 bp) from pDS-Frag4A and HA-Frag4-HA (3789 bp) from pDS-Frag4-RoxPAmpRoxP. **(B)** Schematic representation of the PCR amplified fragments, HA-Gent-HA, used in the first Red/ET reaction and HA-Frag4-HA, used in the second Red/ET reaction. The HA-Frag4-HA fragment contains the minimised selection of genes within Fragment 4 and an ampicillin marker flanked by two RoxP sites. **(C)** 1% agarose gel showing PCR checks of the Fragment 4 baculoviral genome. The PCR products show the expected fragment sizes at each step of the engineering, beginning with the starting genome, SynBac165 (5903 bp), the intermediate genomes, SynBac165Δ4 (1115 bp) and SynBac1654Amp (3339 bp), and the final genome, SynBac1654 (2228 bp). **(D)** Average relative EYFP fluorescence at 528 nm detected for the control SynBac1 and the newly generated SynBac1654Amp and SynBac1654 baculoviruses. 1 million cells taken every 24 hours until 168 hpi during V1 amplification. SynBac1 samples taken until 72 hpi due to visualised cell mass death at 96 hpi.

3.3.2 Individual Fragment Removal

3.3.2.1 SynBac14

SynBac14 was generated as previously described by using the Red/ET recombination method with SynBac1 as the precursor genome, and PCR amplified fragments, HA-GentR-HA and HA-Frag4-HA (Figure 3.11). Firstly, the original Fragment 4 in the baculoviral genome was removed by a fragment harbouring the gene encoding for gentamicin resistance, HA-GentR-HA. The procedure was performed with SynBac1 precursor cells transformed with pRedET. The generated SynBac1Δ4 was then taken through the Red/ET recombination procedure again with the modified Fragment 4 containing an ampicillin marker, HA-Frag4-HA. The colonies were restreaked on gentamicin containing plates to check for complete removal of the gentamicin containing fragment. A bacmid prep from the generated SynBac14Amp clones was prepared and 20 µl was incubated with 3 µl 3.1 NEB buffer, 3 µl Dre recombinase and 4 µl water, at 37°C for 1 hour. The reaction was subsequently dialysed for one hour and 10 µl electroporated into DH10β cells. The colonies were restreaked on Kan only and Kan, Amp plates to confirm ampicillin resistance loss, thus generation of SynBac14.

The correct baculoviral genome obtained at each stage of the modification was checked through PCR reactions that target the region of modification. Figure 3.12 (A) shows the PCR amplified bands from SynBac1, SynBac1Δ4, SynBac14Amp and the final SynBac14 genome. The amplified fragments can be observed at approximately 5.9kb, 1.1kb, 3.3 kb and 2.2 kb, respectively. Alongside measuring the EYFP fluorescence, an AmCyan expressing plasmid (pCPSS2185) was transformed through Tn7 transposition into the SynBac variants. The EYFP fluorescence is illustrative of the virus amplification as it is found in the virus genome and generally detected faster, whereas the AmCyan fluorescence is demonstrative of the recombinant protein expression levels from the Tn7 site. Analysing the two different fluorescence measurements aimed to discriminate between viruses that may have had normal virus amplification but altered recombinant protein expression. Triplicate clones of the SynBac14Amp and SynBac14 genome were picked, prepped and transfected into Sf21 cells, alongside the control genome SynBac1. Samples were taken during V1 amplification at 24 hour intervals until 120 hpi for SynBac14Amp and SynBac14, and 96 hpi for SynBac1. The average EYFP and AmCyan fluorescence for the three SynBac mutants is shown in Figure 3.12 (B and C). This experiment illustrates that the SynBac14 mutants reach comparable levels of EYFP

fluorescence (B) to the SynBac1 control. However, the levels of AmCyan fluorescence (C) for the SynBac14 mutants were much lower compared to the SynBac1 control.

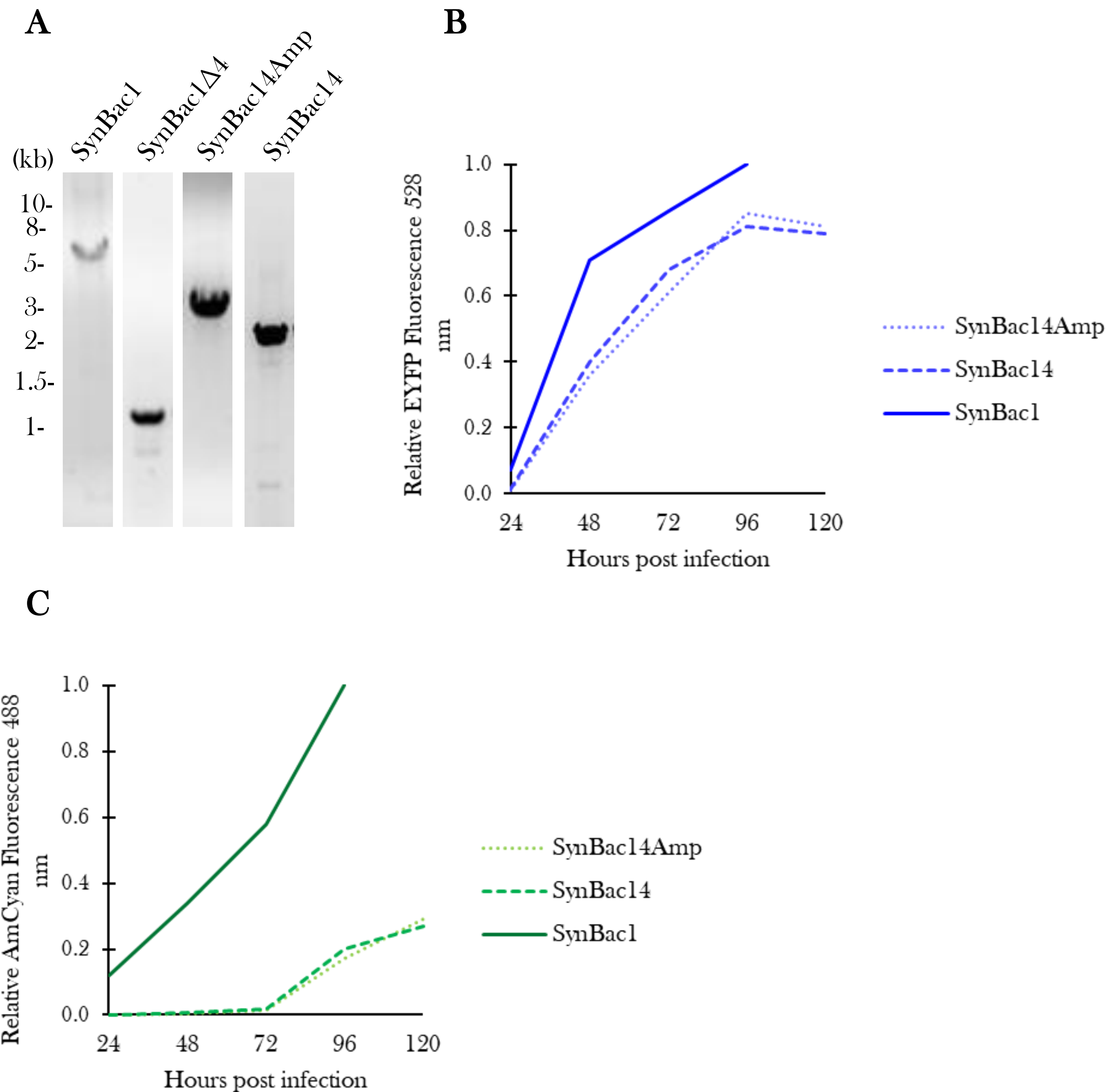


Figure 3.12 - SynBac14 generation and validation.

(A) 1% agarose gel showing PCR checks of the Fragment 4 baculoviral genome. The PCR products show the expected fragment sizes at each step of the engineering, beginning with the starting genome, SynBac1 (5903 bp), the intermediate genomes, SynBac1Δ4 (1115 bp) and SynBac14Amp (3339 bp), and the final genome, SynBac14 (2228 bp). **(B)** Average relative EYFP fluorescence at 528 nm detected for the control SynBac1 and the newly generated SynBac14Amp and SynBac14 baculoviruses. **(C)** Average relative AmCyan fluorescence at 488 nm detected for the control SynBac1 and the newly generated SynBac14Amp and SynBac14 baculoviruses. EYFP and AmCyan measured from 1 million cells taken every 24 hours until 120 hpi during V1 amplification. SynBac1 samples taken until 96 hpi.

3.3.2.2 SynBac13

The Vox site and Vika recombinase have been allotted to Fragment 3. However, as Fragment 3 is much larger than Fragment 4, the modification to incorporate the ampicillin marker and two flanking Vox sites was outsourced to GenScript. The fragments required for the two Red/ET reactions, HA-Gent-HA (2081 bp) and HA-Frag3-HA (8253 bp) were PCR amplified from pDS-Frag3-HomAGentHomB and pDS-Frag3-VoxAmpVox (GenScript), respectively, Figure 3.13 (A). The schematic representation of the PCR amplified fragments is shown in Figure 3.13 (B). The generation of SynBac13 was performed as previously described for SynBac14 (Section 3.3.2.1), starting with the SynBac1 precursor genome but utilising the Fragment 3 PCR amplicons.

To confirm the correct baculoviral genome was obtained, PCR reactions amplifying the region of modification were performed. Figure 3.13 (C) shows the PCR amplified bands from SynBac1, SynBac1Δ3, SynBac13Amp and the final SynBac13 genome. The amplified fragments can be observed at approximately 12.6kb, 1.3kb, 7.5 kb and 6.4 kb, respectively. Similarly to validating the genome of SynBac14, SynBac13 was tested by measuring EYFP and AmCyan fluorescence. Triplicate clones of the SynBac13Amp and SynBac13 genome were picked, prepped and transfected into Sf21 cells, alongside the control genome SynBac1. Samples were taken during V1 amplification at 24-hour intervals until 120 hpi for SynBac13Amp and SynBac13, and 96 hpi for SynBac1. The average EYFP and AmCyan fluorescence for the three SynBac mutants shown in Figure 3.13 (B and C), illustrates that the SynBac13 mutants do not reach comparable levels of EYFP fluorescence (B) compared to the SynBac1 control. Correspondingly, the levels of AmCyan fluorescence (C) for the SynBac13 mutants were much lower compared to the SynBac1 control.

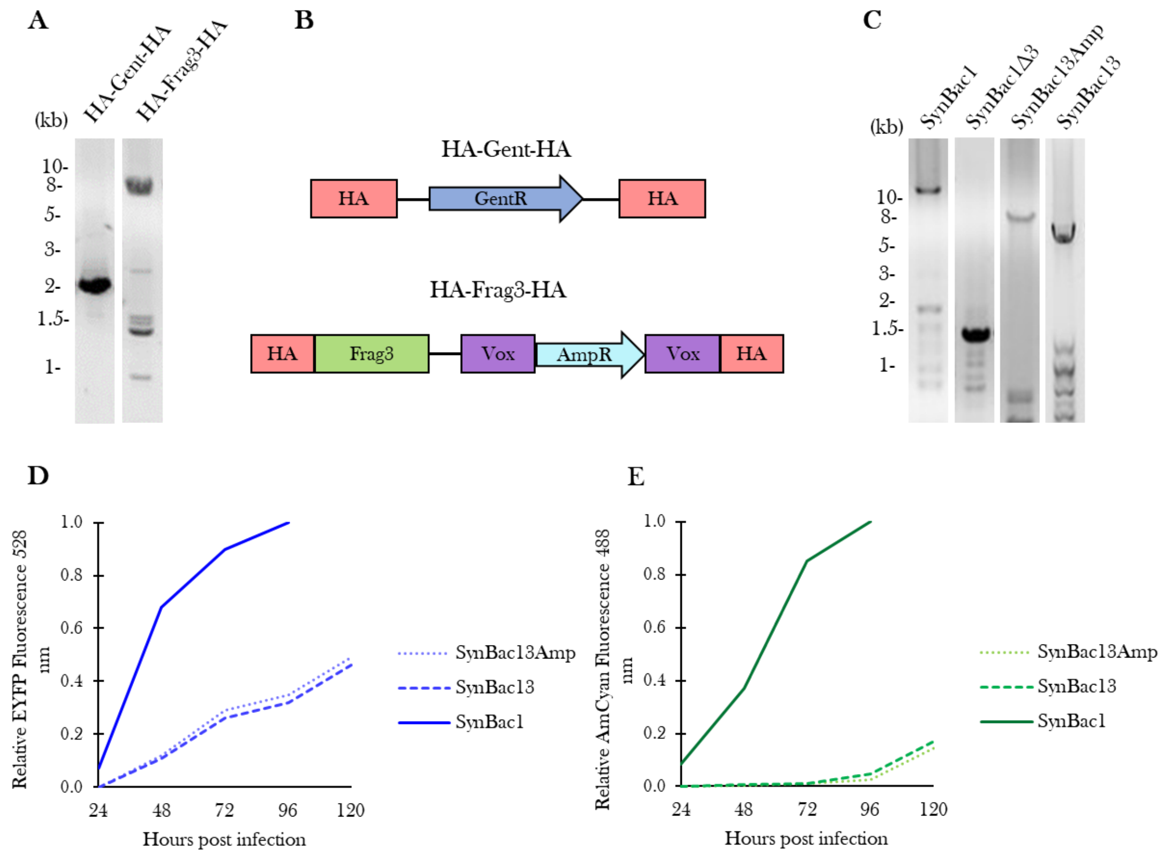


Figure 3.13 - SynBac13 generation and validation.

(A) 1% agarose gel of PCR amplified HA-Gent-HA fragment (2081 bp) from pDS-Frag3-HomAGentHomB and HA-Frag3-HA (3789 bp) from pDS-Frag3-VoxAmpVox. **(B)** Schematic representation of the PCR amplified fragments, HA-Gent-HA, used in the first Red/ET reaction and HA-Frag3-HA, used in the second Red/ET reaction. The HA-Frag3-HA fragment contains the minimised selection of genes within Fragment 3 and an ampicillin marker flanked by two Vox sites. **(C)** 1% agarose gel showing PCR checks of the Fragment 3 baculoviral genome. The PCR products show the expected fragment sizes at each step of the engineering, beginning with the starting genome, SynBac1 (12609 bp), the intermediate genomes, SynBac1Δ3 (1306 bp) and SynBac13Amp (7450 bp), and the final genome, SynBac13 (6375 bp). **(D)** Average relative EYFP fluorescence at 528 nm detected for the control SynBac1 and the newly generated SynBac13Amp and SynBac13 baculoviruses. **(E)** Average relative AmCyan fluorescence at 488 nm detected for the control SynBac1 and the newly generated SynBac13Amp and SynBac13 baculoviruses. EYFP and AmCyan measured from 1 million cells taken every 24 hours until 120 hpi during V1 amplification. SynBac1 samples taken until 96 hpi.

3.4 Discussion

In this chapter, the genome minimisation project, previously initiated in the Berger group, was continued and expanded. The project focused on deleting genes within predetermined fragments that were considered non-essential for laboratory practice by replacing wild-type sequences in the genome with minimal synthetic DNA segments devoid of the 'unnecessary' DNA elements. The aim of the genome minimisation was to investigate if the stability of the generated virus would increase and whether there would be any change in the protein expression levels. Prior to the start of this thesis, a first prototype virus, SynBac1, was prepared with seemingly unchanged virus kinetics and unaffected recombinant protein expression capacity, as compared to the progenitor genome, MultiBac. However, the subsequent modification of the remaining fragments, both iteratively and individually, did not follow this initial trend but instead, resulted in viruses with delayed virus amplification and reduced protein expression capabilities. To streamline the genome engineering, new strategies for the removal of each individual fragment were explored. Specifically, unique site-specific recombinases and their corresponding recognition sites, like the commonly utilised Cre/LoxP system, were investigated and implemented in the engineering of SynBac.

3.4.1 Iterative Fragment Removal

3.4.1.1Confirmation of Starting Material

The SynBac1Δ5 virus genome was shown to express the test protein, GDAP288, thus validating the deletion of Fragment 5 was not deleterious to the baculoviral vector. However, the protein expression as observed by SDS-PAGE was reduced as compared to SynBac1, which may be due to the fragment deletion or, alternatively improper handling during the experiment. Typically, the MultiBac or EMBacY baculoviruses would reach DPA around 48-72 hpi, compared to 96-120 hpi in the case of SynBac1Δ5. The effect on DPA could be due to the deletion of Fragment 5 or through the insufficient preparation and handling of the bacmid, thus leading to the production of weak viruses. Weaker viruses take longer to replicate and spread within the cell culture, leading to a longer period before cell division stops. To rule out any technical uncertainty, the expression test should have been scrutinised again and the DPA closely examined alongside MultiBac or EMBacY.

The validation of the SynBac1Δ6 performed by a co-worker, similarly, showed GDAP288 expression thus validating Fragment 6 deletion was not harmful to the baculoviral vector.

Presence of the test protein in the SynBac1Δ5 and SynBac1Δ6 expression experiments meant further modifications to remove the GentR marker and combine the two deletions were undertaken.

3.4.1.2 SynBac16

Prior to removing the GentR marker from SynBac1Δ6, the baculoviral genome and the pHelper needed to be isolated. The procedure to isolate the pHelper and bacmid through electroporation and antibiotic selection was proven to be effective. The method was easy to follow and the separation of the bacmid from the pHelper was successful. However, at the beginning only one colony was present carrying just the bacmid compared to numerous colonies corresponding to the pHelper. The size difference between the baculoviral vector and the pHelper is large, therefore the chances of the bacmid crossing through the cell membrane during electroporation is reduced. Additionally, the protocol to isolate the bacmid where a restriction enzyme targeting the pHelper could have been used. The enzyme Bsu36I, which is specific for the pHelper and not the bacmid was applied overnight after the bacmid preparation. The purpose of this was to increase the number of colonies containing just the bacmid by digesting pHelper, preventing its replication. Although the pHelper digestion did not result in a major improvement, the possibility of using the restriction enzyme in future experiments exists to increase the probabilities of bacmid-only transformation and isolation. The high copy number of pHelper leads to high amounts of this contaminating plasmid DNA present in the bacmid preparations, thus increasing the concentration of Bsu36I may aid in the complete digestion of the pHelper. Incidentally, at a later stage in the project, the isolation of the bacmid from the pHelper did not prove to be difficult, with reactions yielding plentiful colonies. One possible explanation may be that this was due to the improved handling during bacmid isolation leading to cleaner bacmid preparations, hence higher chances of the baculoviral vector crossing solely through the cell membrane during electroporation.

The *in vitro* manipulation of the isolated SynBac1Δ6 to remove the GentR marker by the action of homing endonucleases was successful. The process yielded many colonies as an increased volume of bacmid was used for the digestion/ligation reaction than previously for the bacmid isolation experiment. Out of 40 restreaked colonies, all but one failed to grow on gentamicin plates, indicating a high efficiency of digestion by PI-PspI to excise the GentR marker. However, out of the first 16 clones tested for EYFP absorbance, only 2 clones displayed EYFP presence. The *in vitro* modification to the baculoviral vector had potentially adverse effects on the bacmid during the PI-PspI digestion where possibly, other sites may have been targeted, thus damaging

the vector. Decreasing the time of the digestion from overnight to few hours could decrease the occurrence of damaging cuts and increase the ratio of EYFP containing clones. The EYFP results of the clones showed how critical subsequent testing is to ascertain that the correct clone was used to proceed with the next competent cell preparation. Although, similar amounts of EYFP were reached for the SynBac16 variants compared to the SynBac1 control, the time taken to reach these levels was significantly longer. A further experiment collecting samples each day of the virus amplification needed to be carried out to determine the virus kinetics and eliminate ambiguity about virus performance.

3.4.1.3 Generation of Tools and Methods Required for Further Fragment Integration

A purification protocol for the recombinases was tested by a co-worker based on the HEPES system, as previously seen was better at solubilising the Dre recombinase than PBS. Using the protocol outlined in the Methods (Section 2.5.3) all four of the recombinases have been purified. However, SCre and Vika required additional concentration steps to reach similar quantities (Figure 3.7 (A)). In addition, SCre and Vika samples also displayed higher levels of contaminants. During the expression trials of the recombinases, high levels of some of the proteins were observed in the insoluble fractions, depending on the harvest time of the cultures. High levels of protein were also present in the flow through samples of the purification. These therefore required re-binding of the protein to the purification column. For an increased final protein concentration in future purifications, the flow through samples need to be monitored and the sample rebound if required. An additional size exclusion purification step could also be applied to remove the remaining contaminants and thus improve the activity of the enzymes.

The expressed and purified recombinases were tested prior to using the enzyme on the bacmid and for the donor/acceptor fusion reactions. Before the experiment, the technical parameters for using the recombinases were investigated. The buffer composition for the Dre recombination reaction contained: 50 mM Tris, 100 µg/ml BSA, 10 mM $MgCl_2$, 50 mM NaCl, and 1 mM DTT as per a recent publication regarding Dre recombination strategies.[273] The reaction was decided to be carried out in the NEB 3.1 buffer (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$ and 100 µg/ml BSA) as the buffer is widely available and the composition is similar. The successful removal of the ampicillin marker by all the recombinases was observed. The time course samples at 30 minutes, 1 and 2 hours revealed that 30 minutes is sufficiently enough for the removal reaction, however the 1-hour sample showed a possibly stronger band for the ampicillin marker.

From this experiment the incubation time and recombination buffer were established for future reactions.

A number of variations of the donor plasmids were successfully generated (Figure 3.8 (A)). Specifically, the antibiotic resistance marker of pUCDM was changed to ampicillin, blasticidin, puromycin or zeocin, and the LoxP site was exchanged for RoxP, SLoxP, VLoxP and Vox. In addition, fluorescent protein markers including AmCyan, mTagBFP, EGFP and mCherry were inserted for visual confirmation of the systems functionality. All the described steps showed no complexity, except in the use of blasticidin where the optimal concentration required adjustment for successful transformations and growth in liquid culture.

Through the adaption of the Cre recombination protocol with the parameters tested in the previous recombinase experiment, successful fusions of all donors and acceptors were generated. The clones were screened by restriction mapping and the correct pattern was acquired. Performing this procedure in parallel with Cre recombinase and the LoxP site containing donor and acceptor, confirmed that there was no cross-talk between the LoxP and the other four recognition sites as well as the different recombinases due to lack of colonies present after recombination event. The recombinase tests confirmed that the activity and specificity of the systems was appropriate, and thus enabled the use of the enzymes on the bacmid in subsequent experiments. The donor, acceptor and fusion plasmid restrictions, (Figure 3.8 (B)), confirmed the expected size of the plasmids, and furthermore sequencing results of the donors verified that the correct modifications were undertaken.

The expression of the fluorescent protein markers from the fusion plasmids was validated in the EMBacY system. The proliferation arrest of the cell culture during the virus amplification step was as expected (2-3 days post infection), with highest levels of the fluorescence measured at 3 days. All the fusions displayed similar values for the EYFP fluorescence, apart from pUPDM_VLoxP_EGFP*pACEBac1_VLoxP, where a higher value was observed, (Figure 3.9 (C)). The higher fluorescence could be a result of the EGFP and EYFP excitation and emission maximums being the most alike to each other, see (Table 2.4) in Methods and Materials, thus potentially the EGFP presence increased the readout of the EYFP emission.

3.4.1.4 SynBac1654

Fragment 4 was successfully modified to include the ampicillin marker and flanking RoxP sites, which facilitates the genome modification by providing an easy method of subsequent removal of the resistance marker by Dre recombination. Before the successful cloning, as illustrated in

the results section, the first approach of modifying pDSFrag4 was based upon a SLIC reaction, with the RoxP site present as an overhang on the PCR amplified fragments. However, this method was not effective after several attempts and it was decided that the repetitive nature of the RoxP site, 5' TAACTTTAAATAATGCCAATTATTTAAAGTTA 3', may be able to form secondary structures between overhangs in the same fragment. The secondary structures may inhibit the SLIC reaction by not providing the required overhangs for the ligation of the two fragments. To overcome this, primers were designed to include new overhangs that included unique restriction sites. The restriction ligation cloning approach worked on the first attempt, yielding the correctly modified pDS-Frag4-RoxPAmpRoxP. PCR fragments for the removal of the native Fragment 4, HA-Gent-HA from pDS-Frag4A and the new Fragment 4 from pDS-Frag4-RoxPAmpRoxP were amplified for later use in the Red/ET recombination reaction.

The generation of SynBac1654 from SynBac165 followed as planned, where the Red/ET recombination procedure was very efficient in exchanging the fragments and yielding plentiful colonies. The Dre recombination reaction to remove the ampicillin marker resulted in numerous restreaked colonies not containing the ampicillin marker. Furthermore, blue/white screening indicated that minimal damage to the bacmid occurred as observed by the small percentage of restreaked colonies that displayed a white colour. Only blue colonies were chosen for further testing by PCR amplification of the modified region. The PCR testing proved to encounter the most complications, where several temperatures and addition of DMSO to some reactions were required. For the amplification of the 5.9 kb fragment from SynBac165, the PCR temperature needed to be increased to 70 °C from the recommended melting temperature of ~57 °C. The amplification of the 3.3 kb fragment from SynBac1654Amp on the other hand was not improved by an increase in temperature, but with the addition of DMSO. Prior to the final PCR test of SynBac1654, new primers of longer length were ordered with the hope of achieving improved binding specificity. The PCR reaction did yield the desired 2.2 kb fragment, however numerous unspecific bands were also present. For future PCR tests of new SynBac variants, the use of a different polymerase should be considered. Although, the Phusion High-Fidelity DNA polymerase used in these tests is an excellent polymerase for cloning with minimal mutation rates, it is prone to falling off the elongating strand, and thus generating unspecific products. As the tests are purely for quantifying the correct fragment size amplified, the use of an alternative enzyme such as Taq DNA polymerase may be used. Taq polymerase is not high fidelity but is more stable when compared to Phusion polymerase, where it can aid in the amplification of longer targets. Another option is the Herculase II Fusion DNA Polymerase that shows superior

yields and can amplify genomic DNA targets of up to 23 kb. The Herculase enzyme also displays high sensitivity for amplification of low starting material, which would be an excellent choice due to small occurrence of target site compared to the large size of the bacmid. [277,278]

The EYFP results from the V1 amplification of the SynBac1654 variants and the SynBac1 control virus show the SynBac1654 virus was majorly delayed in terms of the EYFP fluorescence measured. The virus amplification was evidently affected, and the strategy for the genome minimisation project was reconsidered and consequently changed. As the results for the virus kinetics were similar in SynBac1654 and SynBac165 (performed by a colleague), it was suggested that the removal of Fragment 4 alone may not hinder the virus amplification. The remaining fragments were separately removed from the starting genome, SynBac1, thus circumventing the negative impact that Fragments 5 and 6 had on the virus amplification according to the data presented.

3.4.2 Individual Fragment Removal

3.4.2.1 SynBac14

The generation of SynBac14 followed in the same way as SynBac1654, with the PCR fragments already in place and no problems with the Red/ET reactions or finding clones that were ampicillin negative after restreaking. As longer primers for checking the Fragment 4 region were acquired during the checking of SynBac1654, the PCR reactions proved to be less difficult, with the correct bands obtained without too many unspecific bands.

The V1 amplification trails of the generated SynBac14 variants provided interesting results in terms of the differences seen between the EYFP and AmCyan expression. The SynBac14 mutant showed comparable levels of EYFP fluorescence, suggesting that the virus amplification was not negatively influenced. However, the amounts of AmCyan measured were significantly lower when compared to the SynBac1 control samples. The reduced recombinant protein expression could represent that one or more of the genes deleted within Fragment 4 were involved in the successful expression of protein from the inserted DNA sequence in the Tn7 site.

3.4.2.2 SynBac13

The amplification of Fragment 3 required for the second Red/ET reaction was expected to be difficult due to the long sequence (8253 bp). Although unspecific bands were seen on the agarose gel, a sufficient amount of Fragment 3 was extracted, and sequencing results validated that no mutations were present within the sequence. For future experiments where a long piece of DNA

would be required for engineering the baculovirus, the inclusion of restriction sites on the ends of the fragment during the cloning procedure/ordering of plasmids is advised. The restriction sites enable the fragment to be extracted from the plasmid via digestion, and thus the need for PCR amplification is avoided.

SynBac13 was generated and the region of modification checked by PCR. The primers designed for the inspection of Fragment 3 were carefully selected based on the previous problems experienced with Fragment 4. The temperature of the primers was increased to 60-61°C and their possible complementarity checked against the rest of the baculovirus genome. While unspecific bands were detected in the PCR reactions, presence of the expected bands was clearly observed. The results generated from the V1 amplification of the SynBac13 mutants and the control SynBac1 virus revealed that the SynBac13 variants were seriously hindered. The EYFP fluorescence, although, slowly accumulating, did not reach comparable levels to the control, and the AmCyan expression yet again was slower and lower.

3.4.3 Concluding Remarks and Future Directions

The reagents needed to validate the future engineered SynBac genome have been developed and implemented. Purification protocols for the site-specific recombinases, Dre, SCre, VCre and Vika have been presented. The enzymes were purified, and their activity confirmed. Acceptor and donor plasmids containing the different recognition sites for the site-specific recombination, and a selection of resistance and fluorescent markers were generated. The expression of the fluorescent marker from the four donor plasmids was validated in the EMBacY system.

A series of SynBac variants were generated and tested (SynBac1Δ5, SynBac16, SynBac1654, SynBac14 and SynBac13), with the remaining fragments (2A and 2B) checked by a co-worker in parallel. The rewiring of the additional fragments did not lead to a viable baculovirus capable of high-level protein expression. However, the deletion of Fragment 4 indicated that the virus amplification seemingly remained unaffected.

The data acquired from the deletion of the fragments offered valuable insights into the essential nature of the DNA elements of the baculoviral genome. The data was not always fully consistent with the available data the original SynBac blueprint was based upon. Although the genes selected for removal in the genome minimisation project had previous evidence to suggest their dispensability for laboratory practice, the results of the fragment deletions proved otherwise in the context of this project. The results highlight our still sizeable lack of understanding of the baculovirus genome, which is more complex than previously thought. While the deletion of a

single gene may be shown as tolerable, the additive effect of multiple gene deletions may critically alter the outcome. Regardless of the results that followed, the minimisation of Fragment 1 to form SynBac1, generated a virus genome with intact virus kinetics and a protein production capacity comparable to the precursor genome, MultiBac. In terms of the available data in the literature that was used to design the original minimal SynBac blueprint, it should be noted that the paraments analysed were frequently related to the life cycle of the baculovirus in nature. These parameters may be indicative of the essential or non-essential nature of a particular gene but are evidently not the same as the parameters this project prioritises, namely genome stability, virus amplification kinetics and target protein production yield. In this respect, the original strategy to rewire and minimise the baculoviral genome, appears to have been too optimistic, exacerbated by the success with creating the first specimen, SynBac1.

All the original fragments for engineering the SynBac genome have been checked, and no additional fragments on top of the already removed Fragment 1 proved to generate a virus with the desired properties, sufficient virus amplification and protein expression levels. Subsequently the best specimen, SynBac1, was used as the starting point for further engineering by single gene deletion analysis, as described in Chapter 4. All of the genes within the previously designed fragments, as well as additional genes of interest outside of the fragments will be investigated. The single gene deletions were achieved through recombineering of the bacmid by exchanging the gene with a Blasticidin marker, similarly to the first step of the Red/ET reaction in the fragment removal. The exchange is a result of homologous recombination between the homology arms flanking the blasticidin cassette that are complementary to the regions outside of the gene to be deleted. The single gene deletion versions of the baculoviral genome underwent standard testing to determine viable virus amplification, EYFP, and protein expression, AmCyan. The data gained from these experiments enabled us to decide which genes can be deleted for the ultimate, optimal SynBac genome.

Chapter 4: Generation and Testing of Single Gene Deletions in the Baculoviral Genome

4.1 Introduction

Prior to project start, the prototype SynBac1 has been generated by replacing Fragment 1 using the Red/ET recombination method and Cre/LoxP recombination. The vector was validated by showing unaltered protein expression yields. During the first year of the project, SynBac15, SynBac156 and SynBac1654 have been generated. It turned out that the time to reach proliferation arrest was significantly delayed for all these versions of SynBac, substantially impacting on virus performance. As a consequence, the project strategy was adjusted by reverting to SynBac1 and scrutinising each fragment individually (SynBac14 and SynBac13 in this thesis) to determine which of the fragments can be replaced additionally in the SynBac1 framework to obtain a final SynBac version. The removal of the individual fragments however did not result in a virus with the same viability and protein expression levels as SynBac1.

4.1.1 Single Gene Knockout Virus Generation and Testing

Removing more fragments consisting of multiple genes did not result in SynBac genomes satisfying the target characteristics, notably comparable or even enhanced heterologous protein expression levels. Therefore, it was decided to scrutinise all individual genes within the previous fragments separately to determine the outcome of single gene deletions, and then reintegrate into a new genomic blueprint. All the genes within the fragments and additional genes outside of the fragments were to be included in the investigation, totalling 67 genes for which general information can be found in the Tables in the appendix, Chapter 8 (Table 8.3-8.10). The single gene knockout baculoviral genomes are generated through recombination-mediated genetic engineering, known as recombineering.

The method has become very popular in the engineering of large constructs as it precludes the need for suitable restriction sites as in conventional cloning.[225] The technique relies on linear DNA substrates that can be single or double-stranded and contain approximately 50 bases of homology on both sides corresponding to the target site, denoted (HA) for homology arms (Figure 4.1). The linear DNA is commonly made using high-fidelity PCR, where the primer design determines the exact junction of the recombination event.[229] Using double-stranded DNA one can produce deletions, insertions, gene replacements and inversions. In the case of modifying the baculoviral genome to delete the gene, the DNA region will be deleted by recombination with a PCR fragment. The fragment will contain exact homology arms (HA) to the target region and an antibiotic resistance gene (blasticidin) for selection of recombinants, as illustrated in Figure 4.1. The cassette will also contain two I-SceI restriction sites to facilitate the

removal of the cassette in a second recombineering event, and a rpsL gene that confers streptomycin sensitivity to normally streptomycin resistant DH10β cells.[279] Selection on streptomycin plates should enhance the second recombineering step to remove the cassette and achieve a scar-less deletion.

For the purpose of analysing single gene knockout viruses, only the first recombineering step was performed and the virus genomes were subject to testing with the blasticidin cassette in place of the deleted gene. To enable testing, a fluorescent protein, AmCyan, was inserted in the Tn7 transposition site prior to generating the single gene deletions and the pHelper plasmid removed.[164] The pHelper plasmid was removed as previously described in section 3.3.1.2.1 by transforming bacmid preparations into empty DH10β cells and selecting on kanamycin only (baculoviral backbone). The colonies were restreaked on kanamycin and tetracycline plates to confirm loss of the pHelper. This procedure was performed first to save time transforming pHelper and subsequently AmCyan into each of the 67 single gene knockout genomes that were generated. The single gene deletion viruses were analysed and compared to the starting genome by measuring the levels of fluorescence from AmCyan and EYFP already present in the baculoviral backbone during V1 amplification of the virus.

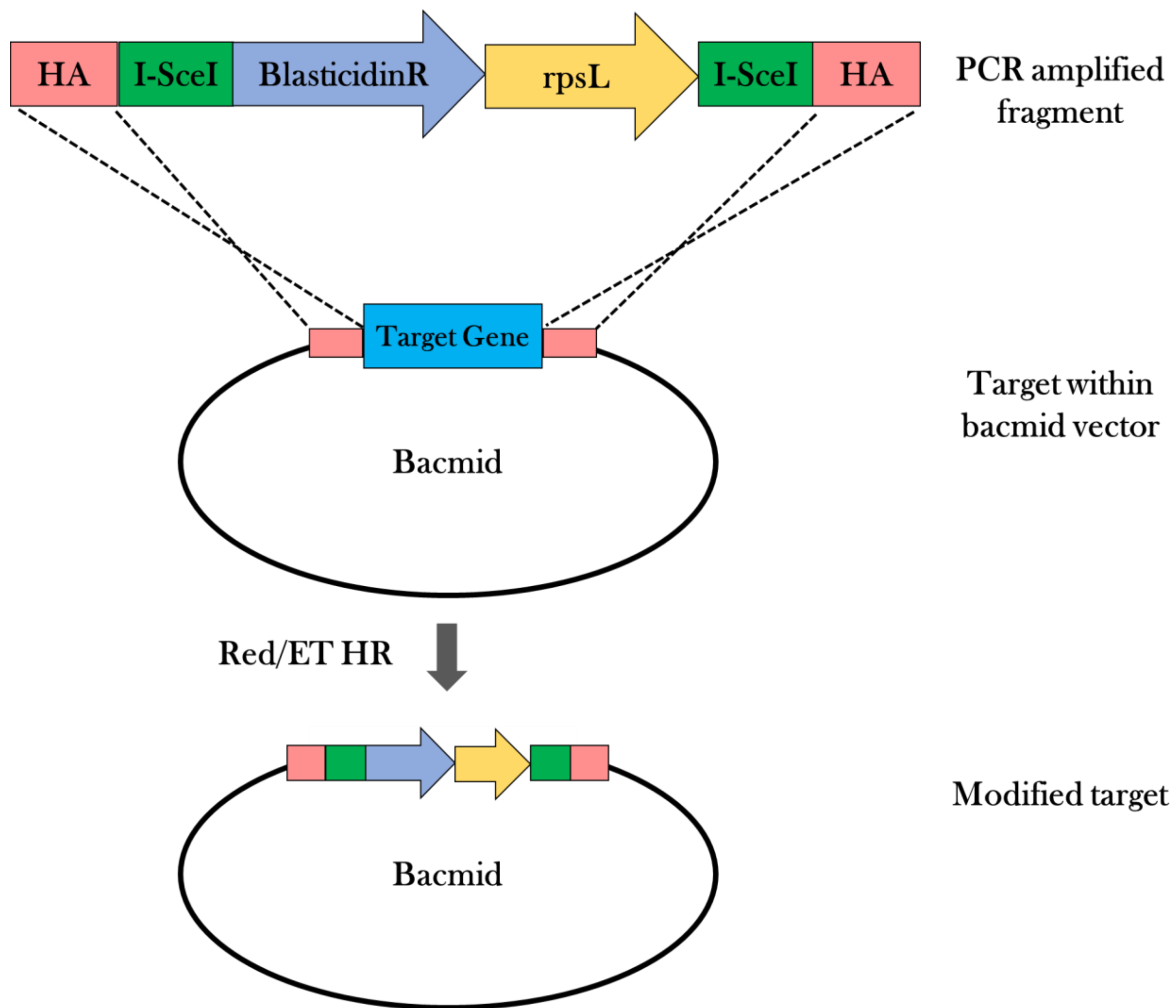


Figure 4.1 - Generation of single gene deletion baculovirus genomes by Red/ET homologous recombination.

Gene deletion achieved by Red/ET mediated homologous recombination (HR) with a linear, single or double stranded DNA substrate made using high-fidelity PCR. A blasticidin cassette with I-SceI restriction sites and rpsL counterselection for enhanced removal in second recombineering step was used. The PCR amplified fragment contains approximately 50 bases of homology arm (HA) on both sides correspoding to target site. Primer design determines the exact junction of the recombination event.

4.2 Chapter Aims

The aim of this chapter is to provide information about the essential nature of the individual genes within the proposed fragments in the original design of the project and other genes outside of the fragment's boundaries. This will be explored through the following approaches:

1. Generation of 67 single gene deletions by homologous recombineering.
2. Investigation of the knockout viruses by measuring the amount of fluorescence from two reporter proteins: EYFP located in the bacmid backbone (to report on viral replication) and AmCyan located in the Tn7 transposition site (to report on recombinant target protein production).

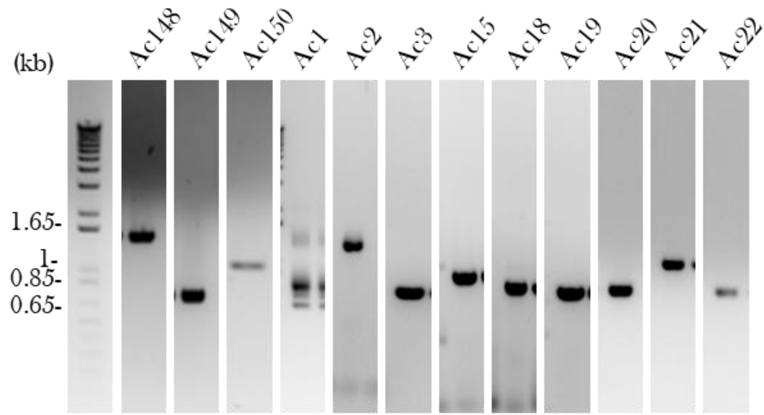
4.3 Results

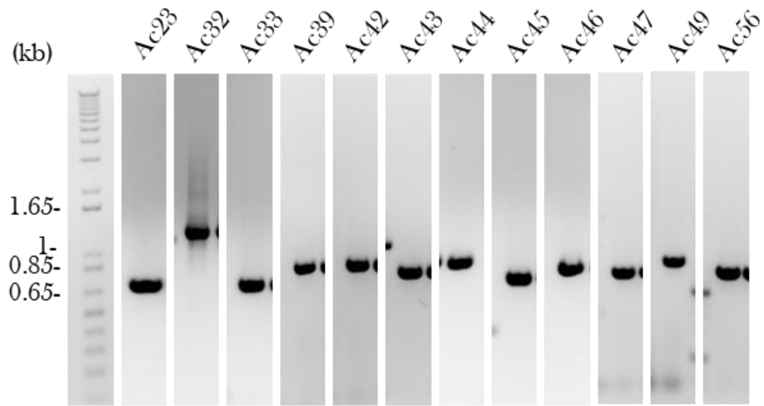
4.3.1 Blasticidin Cassette Amplification and Recombineering

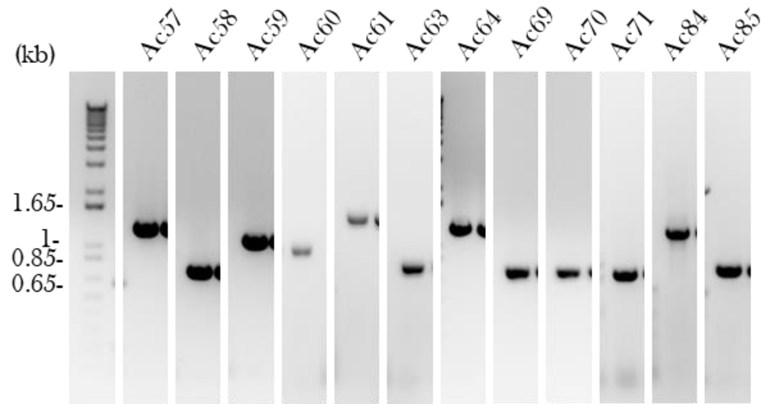
The blasticidin (BSD) cassette was amplified from a plasmid available at AstraZeneca, using primers that contain I-SceI overhangs (ISceIBSDFor and ISceIBSDRev). The 958 bp I-SceI_BSDcassette_I-SceI PCR product was gel extracted, Dpn1 digested and cleaned up. The I-SceI_BSDcassette_I-SceI PCR product was used as the template for PCR reactions using #HAFor and #HARev primers (where # is the number assigned to the different genes from 1-67 and HA is the homology arm). The primers contain overhangs of 35-50 nucleotides, which function as the homology arms for the recombineering reaction. The HA_I-SceI_BSDcassette_I-SceI-HA PCR product of 1043-1058 bp depending on the length of the homology arms was gel extracted. The HA_I-SceI_BSDcassette_I-SceI-HA PCR products for gene 1-67 were then electroporated into cells containing the bacmid and the recombineering pKD46 plasmid. The recombineering procedure occurred as explained in the introduction (Figure 4.1).

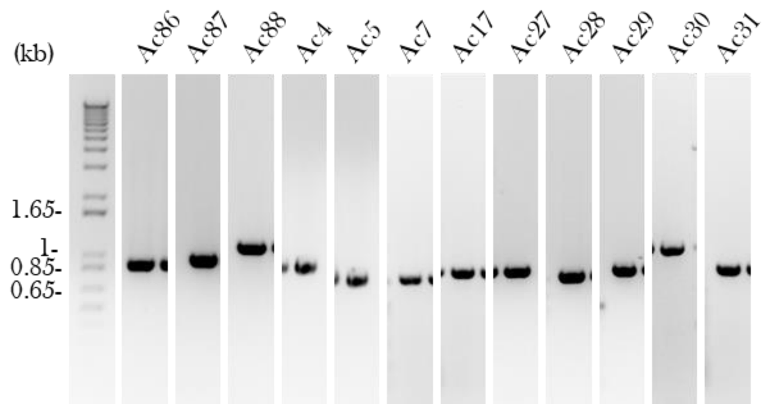
4.3.2 Blasticidin Integration Validation

Colony PCR was used to check for the presence of the BSD cassette in the generated single gene knockout bacmid genomes. All PCR validations used BSDCheckRev as reverse primer and #CheckFor (where # is a different number corresponding to the 67 genes) as forward primer. The original backbone (SynBac1) was used as control for no BSD cassette integration. 3 clones were checked for each gene. All but one (Ac124) of the single gene knockouts were confirmed by PCR, (Figure 4.2) and (Table 8.11) in the appendix for expected size of PCR product. No PCR products were detected for all the control runs with SynBac1. A positive clone for each of the genes was saved and proceeded to testing stage.


(kb)
Ac148
Ac149
Ac150
Ac1
Ac2
Ac3
Ac15
Ac18
Ac19
Ac20
Ac21
Ac22
1.65-
1-
0.85-
0.65-


(kb)
Ac23
Ac32
Ac33
Ac39
Ac42
Ac43
Ac44
Ac45
Ac46
Ac47
Ac49
Ac56
1.65-
1-
0.85-
0.65-


(kb)
Ac57
Ac58
Ac59
Ac60
Ac61
Ac63
Ac64
Ac69
Ac70
Ac71
Ac84
Ac85
1.65-
1-
0.85-
0.65-


(kb)
Ac86
Ac87
Ac88
Ac4
Ac5
Ac7
Ac17
Ac27
Ac28
Ac29
Ac30
Ac31
1.65-
1-
0.85-
0.65-

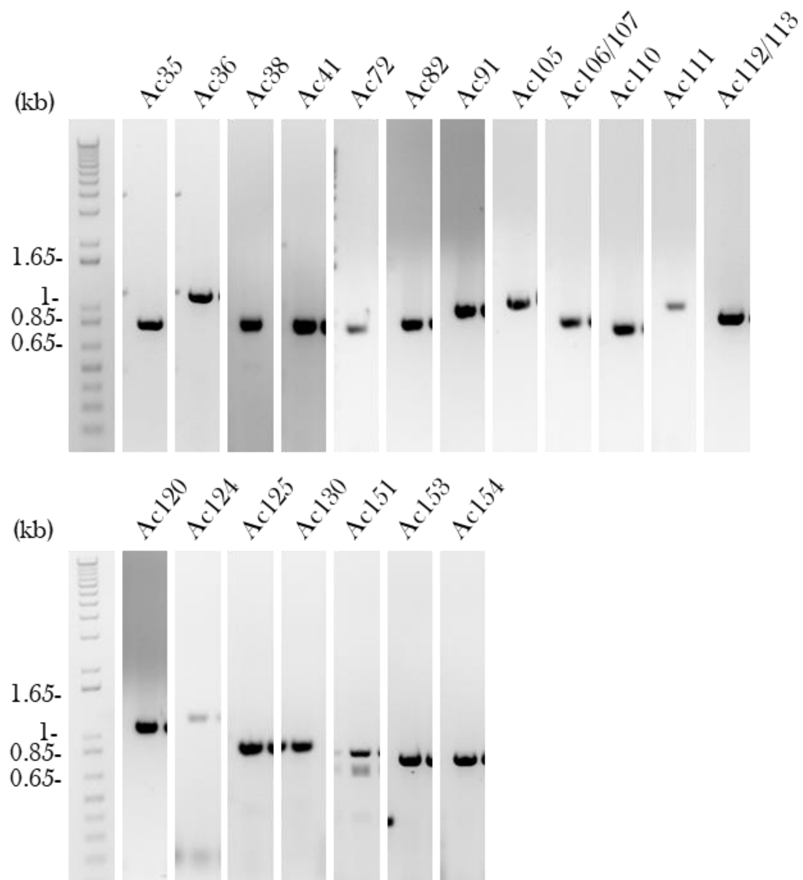


Figure 4.2 - Blasticidin integration validation by PCR on SynBac1 single gene deletion variants.

1% agarose gels showing PCR amplification of the 5' junction. For primer information and expected PCR products refer to Table 8.11 in the appendix.

4.3.3 Single Gene Knockout Virus Testing

The single gene knockout viruses were now subjected to testing by following through the standard procedure of bacmid preparation, transfection and virus amplification as per Methods and Materials section. Prior to generating the single gene deletions in the genome of the baculovirus, the AmCyan expression cassette was put into the Tn7 site to be able to test the modified genomes quicker. The EYFP (bacmid backbone) and AmCyan expression was monitored by measuring the fluorescence of the two proteins from samples taken at 24 h intervals for 4 days. The EYFP fluorescence would be the indicator for successful virus amplification and the AmCyan for recombinant protein expression. Due to the presence of the blasticidin cassette the overnight cultures for bacmid preparation were grown in TB media instead of LB, as low salt LB media proved to be less reliable. The SynBac1 genome was used as the control virus for comparison of the single gene knockout viruses.

4.3.3.1 First Round - All Genes

All 67 single gene deletion viruses were tested subsequently in batches. The following results are displayed by grouping the single gene knockout viruses into the fragments that they would have been deleted in the original design of the project. This has been done to be able to determine which of the genes in the fragments can be deleted and which should be kept in the design. Other genes not included in the original design have been grouped together. Although all the results have been normalized to the control SynBac1 = 1 in each of the experiments, due to the nature of biological systems, variations between experiments were expected and should be considered when comparing the data to each other. Single gene knockouts with the potential for deletion would be grouped into ranks with highest and medium performance. The highest rank contains single gene deletions displaying EYFP and AmCyan of at least 80% of the control, and the medium rank represents single gene deletions with fluorescence values below 80% but above 30%. The medium and highest performing group was to be re-examined again to confirm if any of the gene knockouts in the medium rank belong in the highest rank instead. Likewise confirm if the highest performing single gene deletions still demonstrate comparable values of EYFP and AmCyan to the SynBac1 control genome.

4.3.3.1.1 Fragment 2A

Among the single gene deletions from Fragment 2A (Figure 4.3), gene ac149 EYFP and AmCyan values follow closely to the control, approximately 80% of control for EYFP at both time points and AmCyan at 72 hpi, but the AmCyan value at 96 hpi displays half the value of SynBac1. Ac150 and ac2 range in between 30-60% of the control, ac1 shows little EYFP fluorescence and ac148 and ac3 displayed no real measurable EYFP and AmCyan fluorescence. Ac149 single gene deletion was put in the highest rank, with ac150 and ac2 in the medium rank that need retesting.

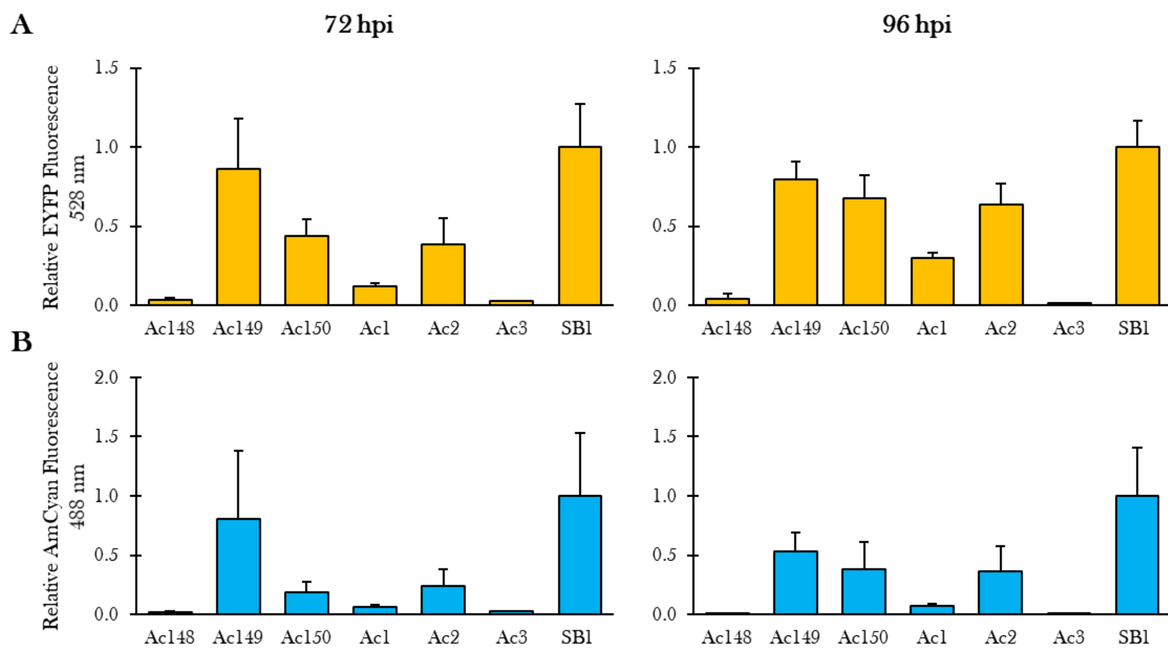


Figure 4.3 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 2A.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.2 Fragment 2B

As seen in Figure 4.4, ac20 single gene deletion shows high EYFP and AmCyan values of 90-100% of the control. Ac22 displays values of 50-70% of the control for EYFP fluorescence but is lower for AmCyan. Lower EYFP (below 50%) and AmCyan expression (below 20%) was seen further in ac15, ac21 and ac23. There was no detectable fluorescence for genes ac18 and ac19. The ac20 single gene knockout was ranked in the highest performing group and ac22 put forward for reassessing in the medium rank.

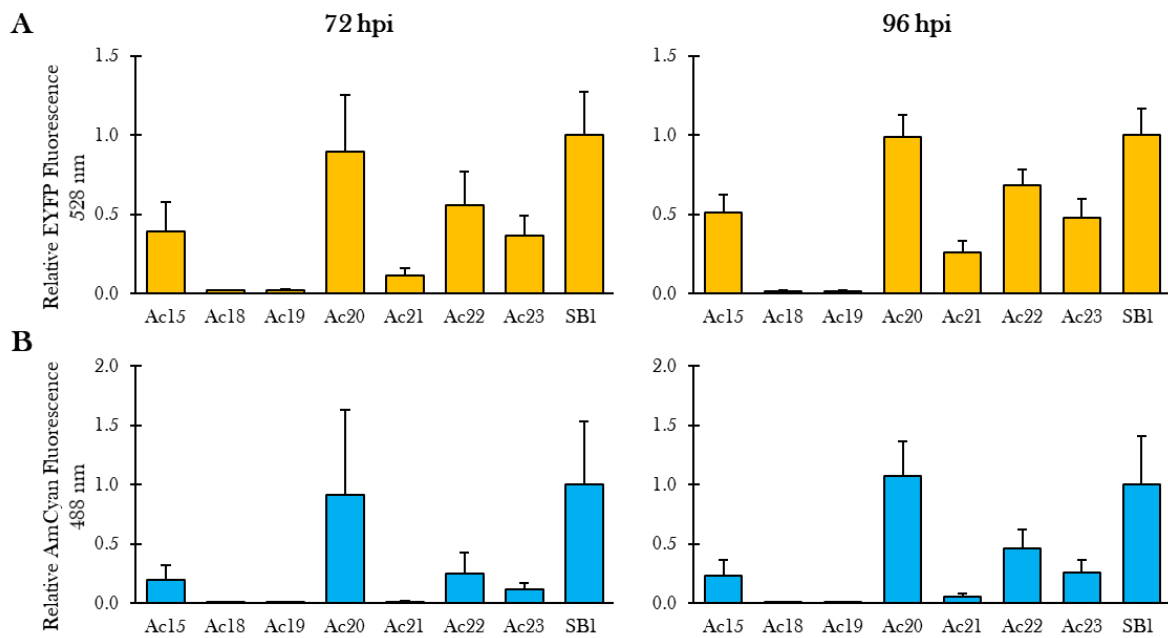


Figure 4.4 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 2B.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.3 Fragment 3

The single gene deletion data from Fragment 3 (Figure 4.5), show a higher number of genes than previously within Fragment 2A and 2B, with similar values of EYFP and AmCyan to the control. Among the single gene deletions that performed well in terms of EYFP and AmCyan fluorescence are ac32, ac42, ac44, ac45 and ac46. The ac47 and ac49 genes showed comparable data for EYFP values, however the AmCyan values were much lower than the control. Genes ac33, ac39 and ac43 displayed very low EYFP and AmCyan fluorescence. From this data, genes ac32, ac42, ac44, ac45 and ac46 were taken further for verification of their performance. Although, ac47 and ac49 display high EYFP fluorescence, these single gene deletions were not put forwards to the second round of testing due to the low AmCyan values detected.

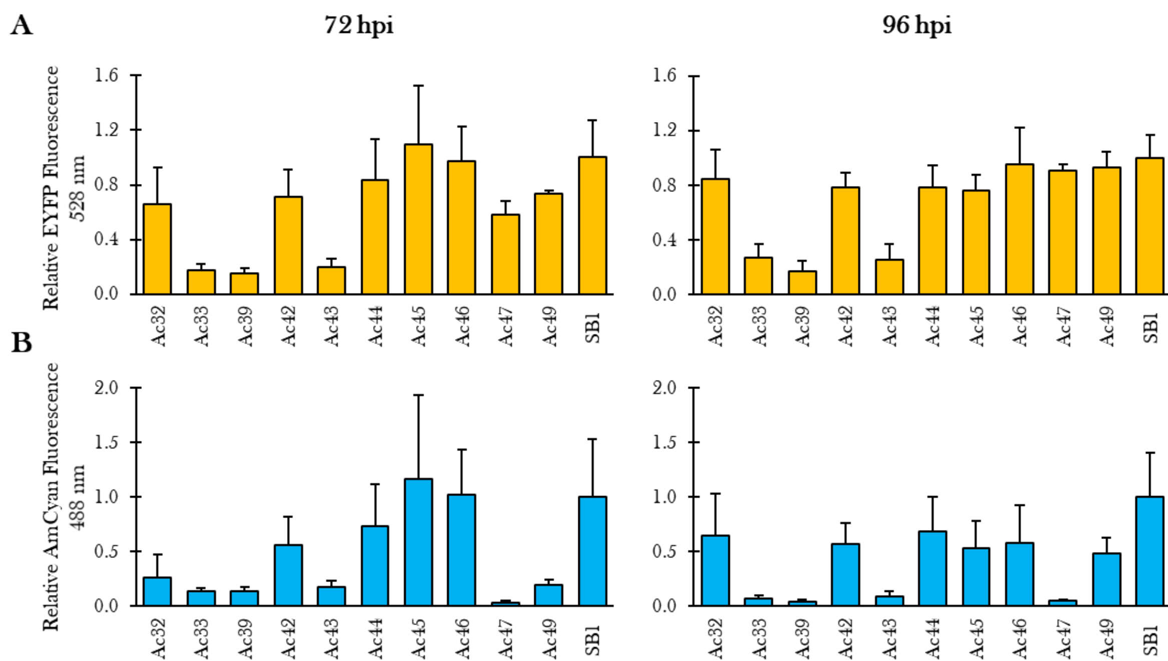


Figure 4.5 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 3.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.4 Fragment 4

The EYFP and AmCyan data for genes found within Fragment 4 (Figure 4.6), reveals ac56 single gene deletion values were very similar to the control and therefore put into the highest performing group for second testing. Gene ac57 EYFP data shows lower EYFP value at 72 hpi, but the value reaches the same level at 96 hpi and displays considerable (almost 70% of control) AmCyan fluorescence at 96 hpi. The ac57 single gene deletion has hence been put into the medium performing group for retesting. The remainder of the genes only showed little EYFP levels and even less AmCyan when compared to the SynBac1 control, thus were not put forward for reassessment.

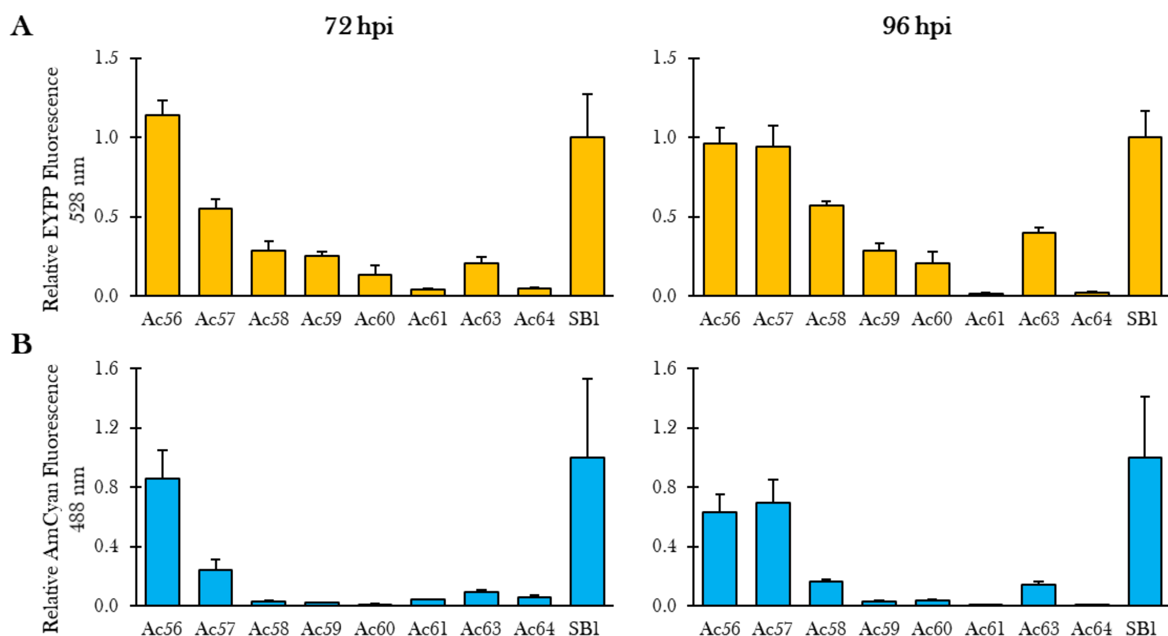


Figure 4.6 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 4.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.5 Fragment 5

The single gene deletions subjects within Fragment 5 show no gene that can be deleted based on the low levels of EYFP and AmCyan fluorescence detected (Figure 4.7). However, gene ac70 has been placed in the medium performing group due to its 50-70% average EYFP and ~50% AmCyan values compared to the SynBac1 control.

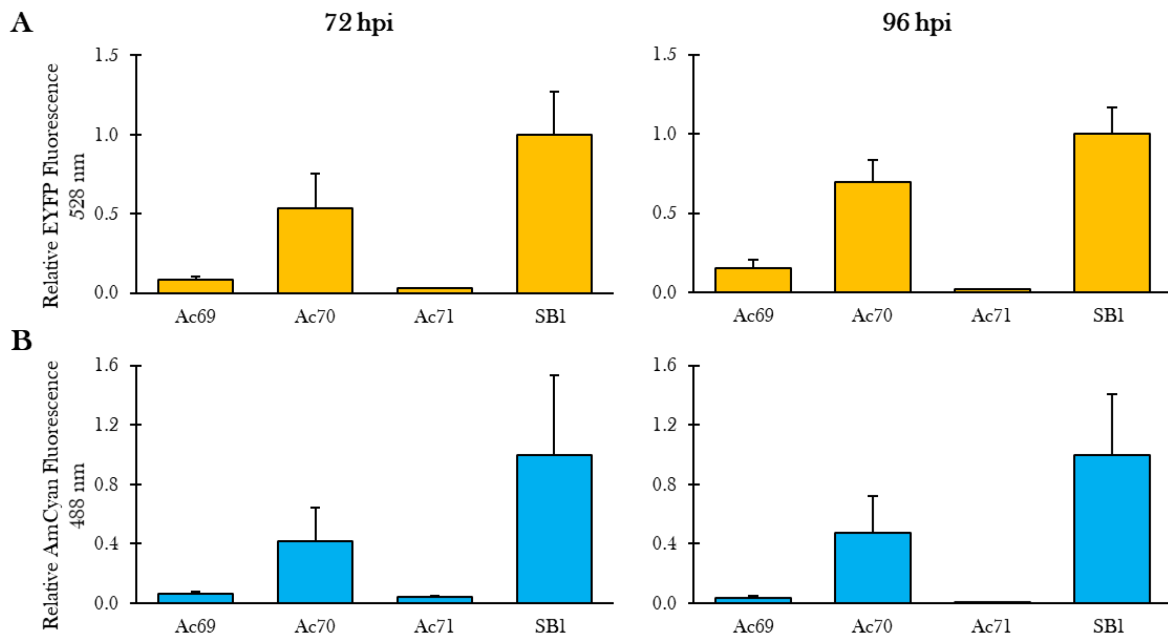


Figure 4.7 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 5.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.6 Fragment 6

Among the single gene deletions from Fragment 6, the fluorescence data (Figure 4.8) reveals gene ac87 shows comparable EYFP and AmCyan levels of 80-90% of the control at 96 hpi, and hence was placed in the highest performing gene group. Although the EYFP values for genes ac84 and ac86 are similar to the control, the AmCyan values are not as high and were therefore placed in the medium performing group for re-examination. Except the partially decent EYFP levels at 96 hpi for gene ac85, the EYFP and AmCyan fluorescence for genes ac85 and ac88 was considerably lower than the control, thus disregarded from further testing.

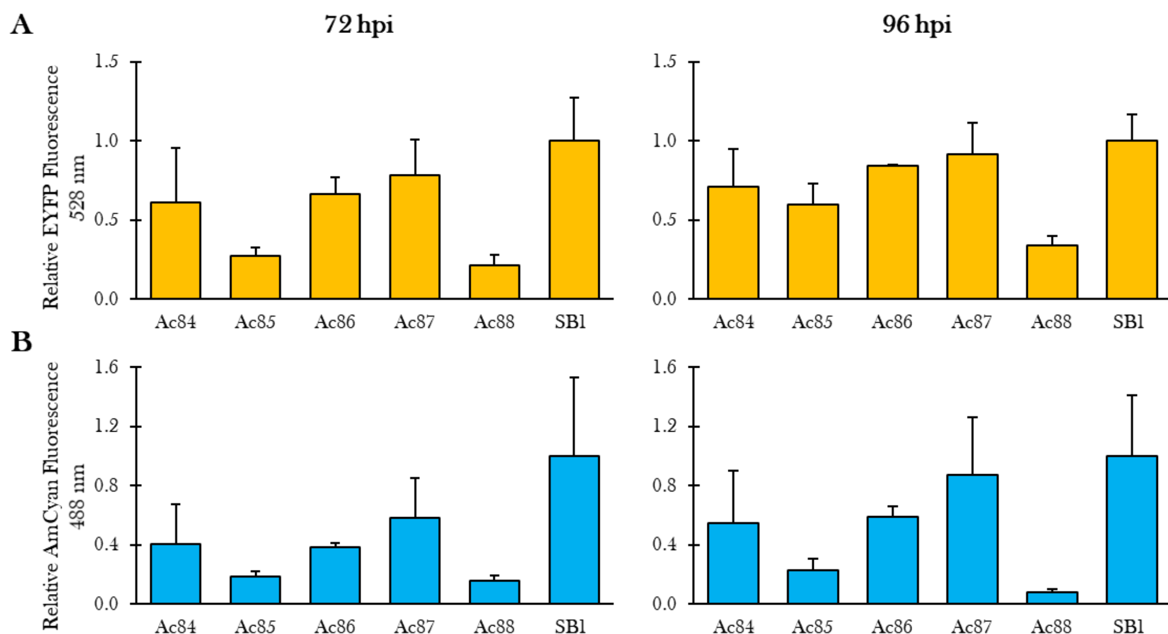


Figure 4.8 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located within Fragment 6.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.1.7 Other Genes

As observed in Figure 4.9, the EYFP and AmCyan data shows two gene deletions, ac91 and ac112/113, which display higher expression levels than the control. Gene ac91 deletion shows a steady level for both values and time points that are just above the control and ac112/113 presents much higher EYFP expression and even higher AmCyan expression compared to the control. Another two genes that are on par with EYFP and AmCyan levels of the control samples are gene deletions ac7 and ac110. These four gene deletions, ac7, ac91, ac110 and ac112/113, have therefore been taken into the highest performing group for second examination. Gene ac111 displays slightly lower values for EYFP and AmCyan fluorescence compared to the control, thus was put into the medium performing group. Gene deletions ac72 and ac124 display some EYFP expression, however the AmCyan levels are considerably lower than the EYFP and the control values. The rest of the gene deletions were noticeably worse compared to the SynBac1 control, with minimal to none EYFP and AmCyan fluorescence detected.

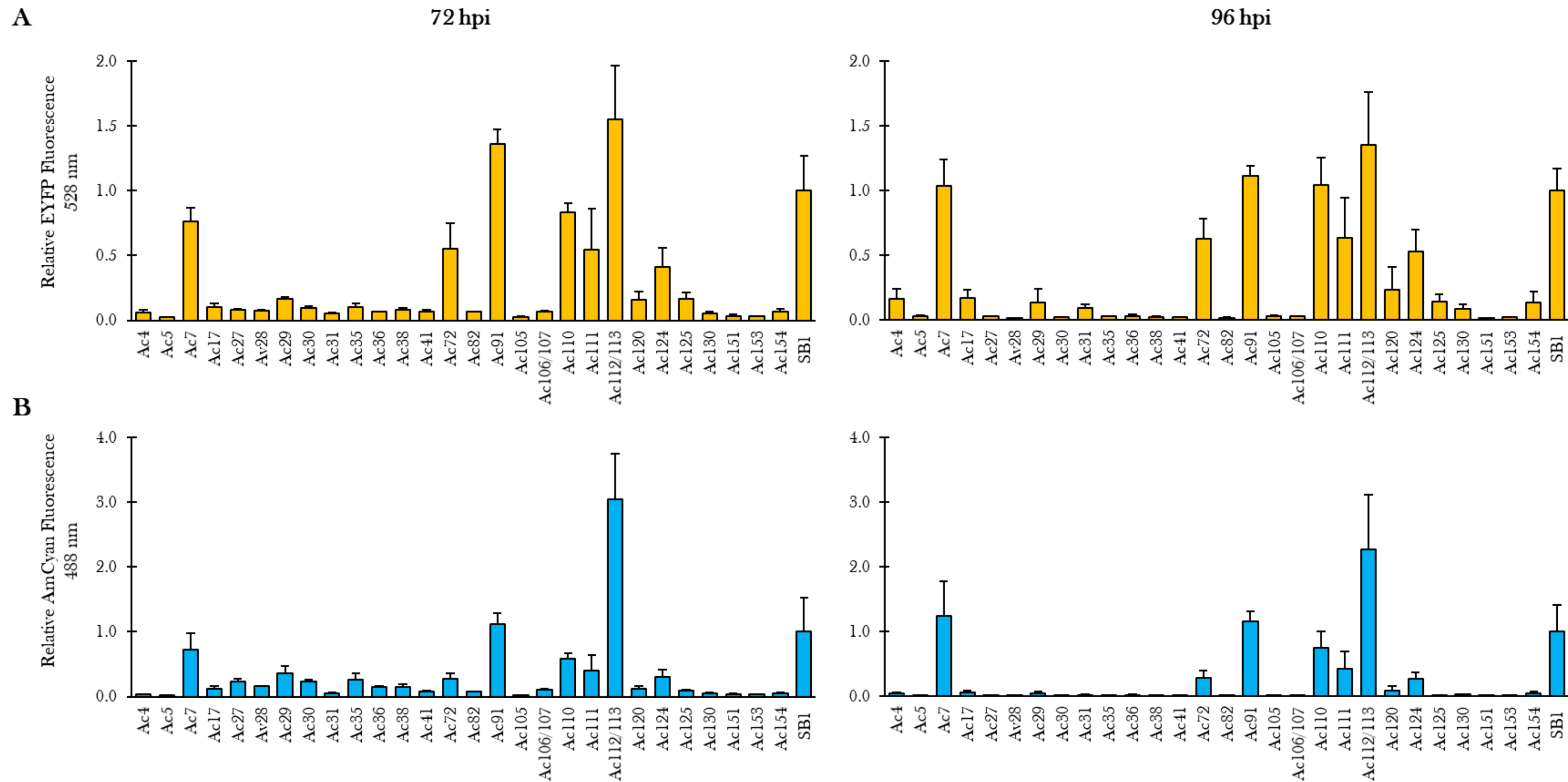


Figure 4.9 – EYFP and AmCyan production at 72 and 96 hpi from SynBac1 single gene deletions located outside of original fragment design.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.2 Second Round - Verification of Single Gene Knockout Candidates

After the initial testing of all the 67 genes, 12 and 9 single gene deletions were grouped in the highest and medium performing ranks, respectively. The two groups of genes were tested together for ease of comparison to the control samples. The medium performing group was tested to check that none of the deletions perform better and if not, they can be fully ruled out as potential targets for elimination. The highest performing group was tested to confirm the single gene deletion viruses perform as well as the control SynBac1. Additionally, the retesting highlights if any of the single gene deletions in fact minimally reduce the levels of EYFP and AmCyan expression, for further exclusion from potential targets for final deletion.

4.3.3.2.1 Single Gene Deletions - Medium Performance

The second round of testing the medium performing single gene deletion viruses confirmed what was seen beforehand in the first round of testing. All the single gene deletion viruses still experienced lower levels of EYFP and AmCyan expression compared to the control SynBac1 (Figure 4.10), where in some cases the levels are yet lower than in the first round of testing. Therefore, these genes were not suitable for deletion from the baculovirus genome.

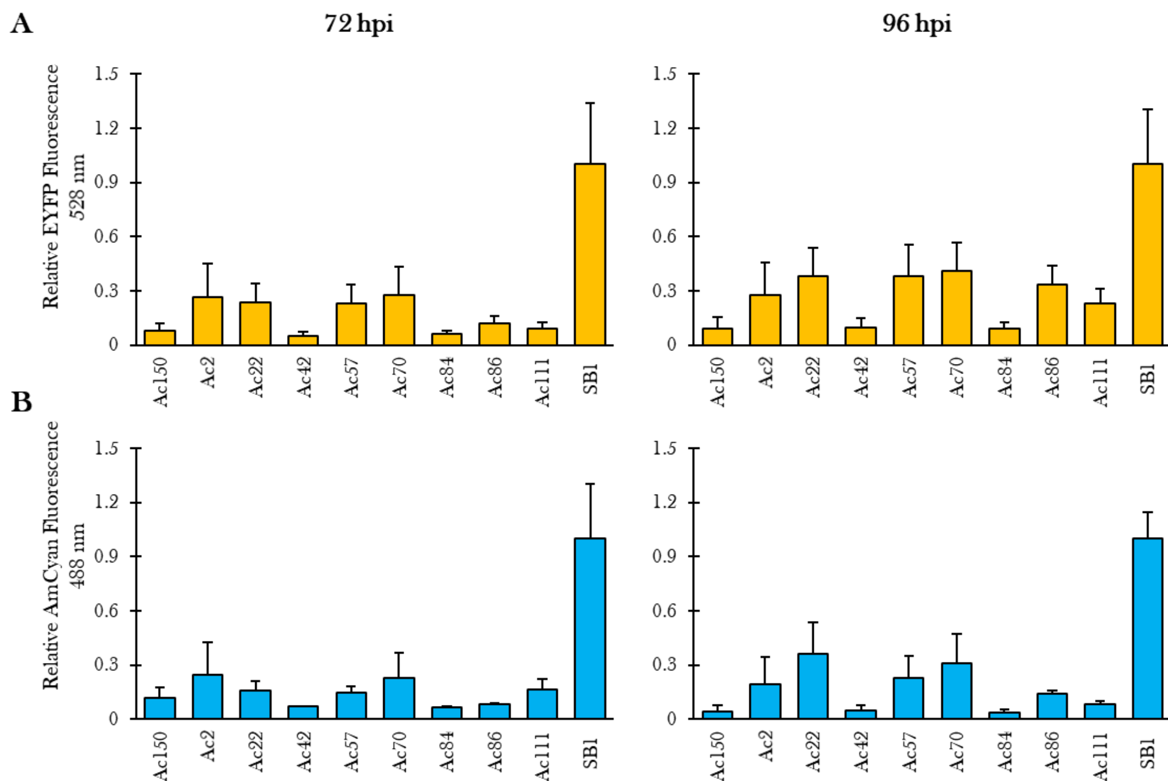


Figure 4.10 – EYFP and AmCyan production at 72 and 96 hpi from the middle rank of SynBac1 single gene deletion candidates.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.2.2 Single Gene Deletions - Highest Performance

The EYFP and AmCyan fluorescence data from the second round of testing the single gene deletions from the highest performing group (Figure 4.11), exposed several differences between the expression levels acquired in the first round of testing. Unlike in the first screening, where ac32 was similar if not slightly lower than the control, the results in the second screen showed ac32 had the highest levels of EYFP and AmCyan, with 5 and 3 times more AmCyan fluorescence at 72 and 96 hpi, respectively, compared to the control. Single gene deletions ac87 and ac110 previously showed very comparable amounts to the control for both proteins, with even higher levels of AmCyan expression detected in the second round, 1.5 and 3.5 times more at 72 hpi for ac87 and ac110, respectively. Although ac112/113 gene knockout still displayed higher AmCyan expression at 72 hpi and the levels of EYFP and AmCyan at 96 hpi are the same compared to the control, the overall amount obtained was smaller than in the first round of testing. Single gene deletions ac20 and ac45 showed similar levels of EYFP and AmCyan (80-

100%) compared to the control, which was also previously seen in the first round. While ac91 displayed a slightly higher level of EYFP and AmCyan fluorescence compared to the control beforehand, the levels were somewhat lower this time with 70-80% AmCyan expression correlated to the control. In the first round of testing, single gene deletions ac46, ac56 and ac7 all showed comparable amounts of EYFP and AmCyan fluorescence measured against the control, however they all exhibited a decrease in the levels of the expressed proteins compared to the control in the second screening. Previously, single gene deletions ac149 and ac44 were the lowest performing genes in this rank with approximately 70-80% fluorescence values compared to the control, and once again displayed lesser amounts of fluorescence, however the amounts were drastically lower in the second examination.

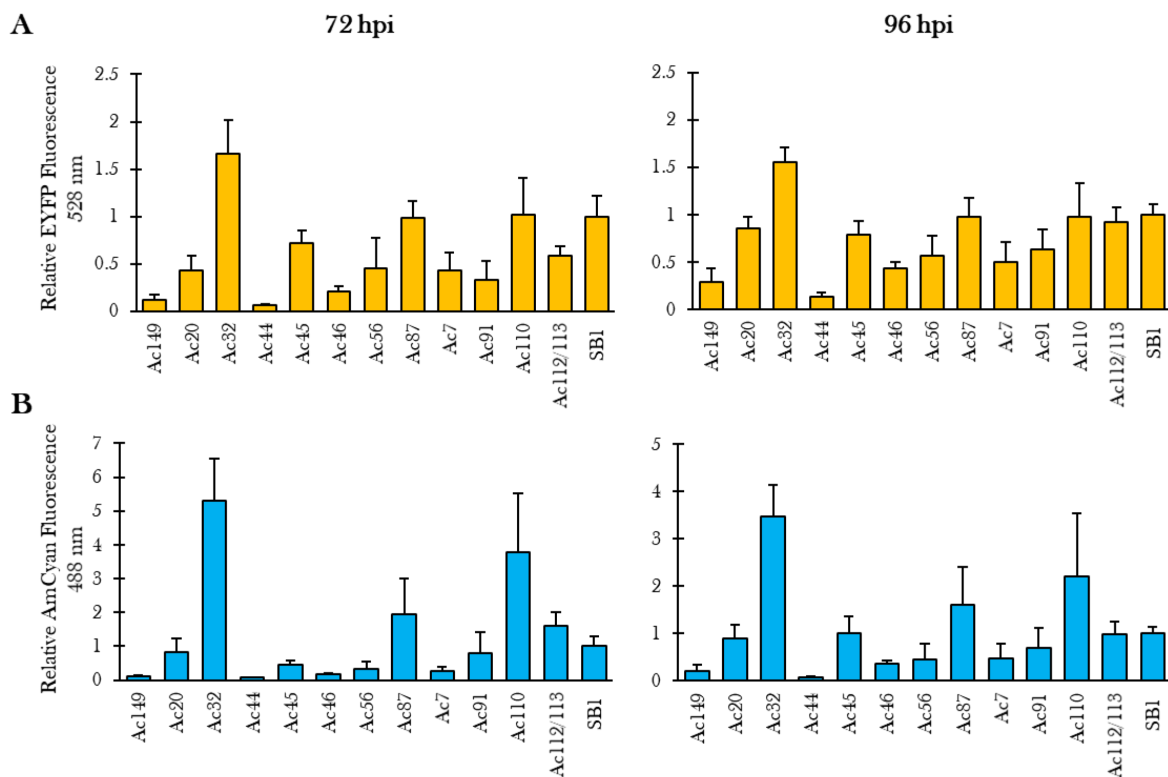


Figure 4.11 – EYFP and AmCyan production at 72 and 96 hpi from the highest rank of SynBac1 single gene deletion candidates.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.3.3.3 Third Round - Retesting Highest Performing Single Gene Deletions

A third round of testing was undertaken (Figure 4.12) to assess any of the discrepancies previously seen between the results obtained in the first and second round of screening the single gene deletion viruses. The three highest expressing gene knockout viruses, ac32, ac87 and ac110 from the second screen, still showed higher levels of fluorescence compared to the control virus, with a particularly more substantial increase in levels for ac87 gene deletion. The expression levels of both proteins for gene deletion ac112/113 were similar to the control, though slightly lower than previously seen in the first and second round of testing, however the levels can be possibly correlated to the quality of the bacmid preparation as the overnight cultures failed to grow to a suitable density. Gene deletion ac20 displayed similar levels of fluorescence compared to the control in the second screen, whereas in the third round the data showed the fluorescence rise to 2 and 2.5 times more than the control for EYFP and AmCyan at 72 hpi, respectively. The results obtained for gene deletions ac45 and ac91 were comparable to the second round of testing, where both showed 70-95% fluorescence compared to the control. The results obtained for single gene deletions ac56 and ac7 are closer to the initial data seen in the first round of testing, rather than the weaker fluorescence seen in the second screen. Ac56 exhibited levels of 85-100% fluorescence of the control for both EYFP and AmCyan, and ac7 displayed an increase of up to 1.5 times more fluorescence compared to the control. Gene deletion ac46 similarly performed better in this round, as it did in the first round of testing, with EYFP levels of 90-110% but lower AmCyan levels of 60% compared to the control. The single gene knockout viruses ac149 and ac44 performed slightly better than in the second screen, as in the first screen, however similar to the first screen the fluorescence levels for both were lower compared to the control, especially for AmCyan.

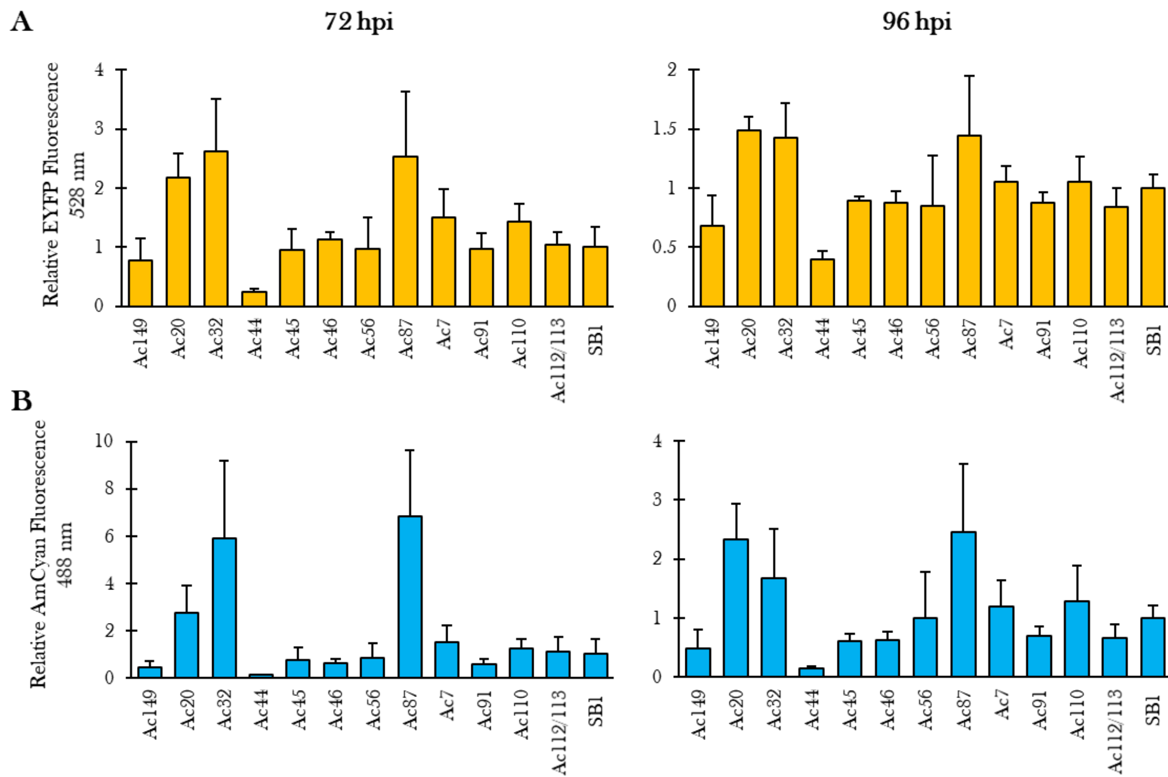


Figure 4.12 – EYFP and AmCyan production at 72 and 96 hpi from the retest of highest ranked SynBac1 single gene deletion candidates.

500,000 cells taken at 72 and 96 hpi during V1 amplification. **(A)** EYFP (yellow) value taken at 528 nm. **(B)** AmCyan (blue) value taken at 488 nm. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

4.4 Discussion

In this chapter, the generation and testing of 67 single gene knockout baculoviruses within the SynBac1 genome framework has been detailed. The gene deletions were achieved using recombineering methods, where a blasticidin cassette was inserted in place of the gene. The single gene knockout viruses were tested for EYFP and AmCyan expression and the fluorescence intensity measured and compared to the original SynBac1 virus.

4.4.1 Generation of Single Gene Knockout Viruses

Using homologous recombineering methods to generate the single gene knockout viruses proved to be both successful and efficient. Numerous engineering examples have been previously described using systems based on homologous recombination and the use of the Red/ET plasmid or similar approaches.[85,229,280,281] One of the most crucial steps to effective homologous recombination is the design of the homology arms, where most commonly 50 nucleotide bases are used, but can go as low as 20 bases. Due to the repetitive nature of some regions in the baculovirus, notably the hr regions, in some cases the design of the homology arms was adjusted by skipping the hr regions and deleting less of the target gene. Each of the designed homology arms were checked for potential complementarity to other regions in the genome. In some instances, the homology arm selected exhibited complementarity to another region with only 4 and 3 mismatches for gene ac148 and ac150, respectively. In these occasions, the homology arms were shortened to 40 nucleotides as opposed to making the homology arms longer as the homologous recombination mismatching is stricter and more precise with a shorter sequence. Despite having occasionally another probable location for a homology arm, all recombineering reactions were successful. The reactions yielded many colonies on plates, with only a handful of the reactions resulting in a small number of colonies. The numbers of colonies can most likely be linked to the design of the homology arm and the quality of the prepared PCR fragment. All but one of the single gene knockouts were confirmed by PCR for correct blasticidin cassette integration. Gene deletion ac124 lacked a correct PCR fragment that may have been caused by another possible primer annealing site present. The testing of the single gene knockout viruses proceeded to include ac124 nevertheless the lack of PCR confirmation. Further sequencing of the genome could have been carried out if the ac124 virus proved to be functional.

4.4.2 Initial Testing of Single Gene Knock-out Viruses

The first screening of the 67 single gene knockout viruses revealed less genes that could be deleted compared to the information gathered in the available literature. Overall, the results pointed to 12 potential targets grouped in the highest performing rank compared to the progenitor genome and 9 in the medium performing group. Included among the genes were two negative controls, ac36 and ac106/107, which were foreseen to generate non-functional viruses, refer to (Table 8.10) in appendix. Deletion of gene ac36, expressing protein 39K, was previously stated to reduce late gene transcription with a 100-fold reduction in infectious BVs.[282-284] A BmNPV homolog mutant for gene ac106/107 displayed a virus phenotype that was unable to spread infection.[285,286] The data gained from the first round of testing agreed with the notion that deletion of gene ac36 and ac106/107 would result in non-functional viruses.

Based on the literature, there was also a substantial number of genes (15) that were predicted not to work or experience reduced protein expression levels. Deletion of gene ac46, encoding a part of the occlusion derived virus envelope, odv-e66, was previously reported to generate functional spread of infection in cell culture, as it is not involved in the production of budding viruses. The EYFP fluorescence data collected showed the virus was viable with 90-100% EYFP of the control confirming the deletion of the gene did not impair the spread of virus between cells, however the recombinant protein expression for AmCyan was reduced by about 30%, which was also reported for a BmNPV deletion mutant for ac46 (bm37).[285] Gene deletions ac61 and ac71 encoding fp and iap2, respectively, were previously reported not to be lethal but resulted in a 'few polyhedra phenotype' for ac61 and decreased protein expression for ac71 in a BmNPV homolog (bm58) deletion mutant.[26,27,285,287] Both of the single gene deletion viruses however displayed no detectable fluorescence at all for EYFP and AmCyan in the initial testing. The protein, cg30, encoded by ac88 was shown to result in a 10 -100 fold reduction in virus titers[288],and re insertion of this gene into the Tn7 site of a Fragment 6 null virus mostly rescued the infection kinetics (data collected by other member of lab). The gene was therefore essential, which was confirmed by the data collected for the single gene deletions within Fragment 6, where deletion of ac88 showed very low levels of EYFP and AmCyan compared to the control and other deleted genes within Fragment 6. There were conflicting reports about bm23, a AcMNPV homolog of ac31. In one study, bm23 was classified as essential for virus replication, but in another study deletion of bm23 apparently did not negatively impact the virus.[285,289] No data exists about deletion of ac31 from AcMNPV however the results acquired in the initial testing lean in favour of the gene being essential due to lack of detectable EYFP and AmCyan in the experiments. Deletion of gene ac17

was reported not to interfere with virus DNA replication, however it was needed for rapid expression of viral genes that led to reduced viral titers.[54] The data for deletion of gene ac17 showed very low levels of EYFP and AmCyan supporting the earlier claim. Although the importance of late expression factors (lefs) in the onset of late transcription has been described previously, there are examples of mutants with infectious phenotypes and therefore three of the lef genes (ac28, ac41 and ac125) were included in the single gene deletion screen for confirmation of results.[285,288,290–293] However, all three of the single gene deletions failed to display measurable EYFP and AmCyan levels, which highlights their importance in late gene expression and production of infectious viruses. Gene ac41 further included an overlooked deletion of the AT dinucleotide within ATG start codon of the neighbouring essential gene, p47, which would have been recovered after the blasticidin cassette removal and could potentially be retested, nevertheless the information gathered predicts the deletion is detrimental. A deletion mutant for the ac35 homolog in BmNPV (bm26) was shown to be viable but with less protein expression detected and a 5-10 fold reduction of BV in AcMNPV.[285,294] The lack of EYFP and AmCyan measured for gene ac35 supports the previously observed virus titer reduction in AcMNPV. Although a BmNPV deletion mutant (bm29) for the ac38 homolog showed the virus was viable with only reduced levels of protein expression, the deletion mutant was severely compromised in AcMNPV, which was also seen in the initial testing.[285,295] Deletion of genes ac82 and ac153 has been previously shown to significantly reduce BV production, which was illustrated in the initial testing with the absence of EYFP and AmCyan fluorescence.[296–300] Deletion of gene ac151 was reported to not affect the infection in Tn-5B1-4 cells however in Sf21 cells the spread of infection was hindered.[301,302] The lack of detectable EYFP and AmCyan in the first round of testing confirmed that infection of the ac151 null virus was affected in a major way in Sf21 cells.

Closer scrutiny of the data revealed that some single gene knockout viruses displayed lower levels of AmCyan compared to EYFP, which may indicate that these genes are responsible for late gene expression and therefore expression of a recombinant protein. Two gene deletions, ac47 and ac49, illustrated this phenomenon particularly well. Gene ac47 showed 90% EYFP at 96 hpi compared to control, but no measurable AmCyan. Apart from a study where the BmNPV mutant homolog (bm38) to ac47 exhibited a viable virus, no additional information can be found about its function in the AcMNPV.[285,303] Gene ac49 also displayed 90% EYFP at 96 hpi compared to the control but with only half the amount of AmCyan. It has been previously predicted to be an early gene needed for correct and timely expression of late genes, which explains why the

EYFP levels were normal (virus replication unaffected), but the AmCyan expression was reduced.[304,305]

4.4.3 Retesting the Single Gene Knockout Viruses

The second round of testing the single gene deletion viruses grouped in the medium performing rank confirmed the viruses still exhibited lower levels of EYFP and AmCyan expression and therefore would not be ideal candidates for deletion. Overall, the measured EYFP and AmCyan fluorescence was lower compared to the control than in the first round of testing, however the control virus exhibited higher levels of EYFP and AmCyan (raw data) than previously. The amounts of the expressed proteins can correlate to the strength of the virus, which could be influenced by the quality of the bacmid preparation. Analysing bacmid preparations by agarose gel electrophoresis provides a visual estimation on the amounts of bacmid present and helps predict if transfections are less likely to generate good quality viruses. The presence of genomic DNA in the preparation also impacts the quality of the transfections by sequestering transfection reagent and thus the technical procedure of bacmid preparation should be performed with precision.

The second and third round of testing the single gene deletion viruses in the highest performing group displayed some differences among the data acquired on the three separate occasions of testing. This illustrates the importance of repeating experiments for maximising confidence in the results, as opposed to only including replicates within experiments. It is particularly crucial to be able to generate reproducible results, at the same time it is expected that working with a biological system would display slight differences between separate experiments. In addition to the quality of the bacmid preparation, other factors like cell culture health and age can influence virus quality. The cell culture should be maintained within appropriate densities, in the logarithmic growth phase, especially on the planned day of transfections or V1 virus amplification. If the cell culture overgrows, is seeded too low or has a high passage count, the cells are likely to experience slowed down replication cycles and affect the virus infection. Although the cell culture was maintained within suitable ranges, cell culture health between experiments can fluctuate and in turn effect the virus infection and protein expression levels.

Although some differences were displayed between the three rounds of testing, overall a pattern of best performing single gene deletion viruses clearly emerged. Single gene knockout viruses ac20, ac32, ac87, ac110 and ac112/113 constantly displayed higher or similar levels of EYFP and AmCyan compared to the SynBac1 control virus (subjective to a statistical test) and will therefore

be included in the plan for deletion from the SynBac1 genome. Genes ac32, ac87 and ac110 were previously shown as non-essential as deletion in AcMNPV or BmNPV displayed no differences between wildtype virus.[46,285,306–311] Ac32 displays homology to fibroblast growth factor (fgf), and the AcMNPV fgf is secreted and stimulates insect cell motility. The deletion of ac32 resulted in no change from *wt* in cultured cells, however time of larvae death was delayed. Thus, ac32 is likely required for stimulating cell migration in the insect larvae, but the function is redundant in cultured cell, hence the deletion was tolerable.[307] Ac110 encodes one of the Per os infectivity factors (pif-7) involved in the initiation of infection in larvae by ODVs.[43] The spread of infection in cultured cells does not rely on ODVs, rather is mediated by BVs, thus the deletion of ac110 did not impact the virus in a cell culture setting. The BmNPV homolog (Bm70) of ac87 was suggested to encode a capsid protein called p15. There is a lack of functional classification of ac87 in the AcMNPV virus. Gene ac20 is the C-terminal region of the actin rearrangement-inducing factor (Arif-1) protein, expressed together with gene ac21. It is expressed as an early gene and was shown to induce actin rearrangement. The expression of Arif-1 is greater in midgut cells rather than cultured cells and deletion mutants showed delayed infection in several tissues of the insect larvae, suggesting it plays a role in establishing infection in the host, but it is not necessary in cultured cells. It was previously shown that deletion of the C-terminal half of the Arif-1 gene or insertion of the LacZ gene in the middle produced viable viruses, which agrees with the gained results.[312] There is a lack of gene deletion data for ac112/113 in AcMNPV and no homolog is present in BmNPV. The absence of known homologs indicates the gene was not well conserved in baculoviruses, likely due to a non-essential function. A related orf in LdMNPV (ld109) encodes a suppressor of apoptosis, however ac112/113 showed no apoptotic activity and is likely non-essential due to the results acquired in the present study. Single gene deletions ac45, ac56 and ac7 showed very comparable levels of EYFP and AmCyan compared to the control across the three rounds of testing and thus will be included in the deletion plan as well. Homologs for ac45 are only found in four other closely related viruses. Structural predictions of ac45 show the protein is similar to a S-phase kinase associated/F-box protein, which normally mediate protein-protein interactions. Although homologs of ac56 are found in most of the Alphabaculoviruses, there is no information about the gene products function. It was previously shown homologs of gene ac45 and ac56 deletions in the BmNPV (bm36 and bm45, respectively) resulted in viable viruses with no noticeable differences between the wild type virus, thus the deletions were also tolerable in AcMNPV.[285] Gene ac7, is only found in 3 other baculoviruses in addition to AcMNPV and has no known function. Ac7 deletion displayed no effect on virus replication in Sf21 cells, although only 170 nucleotides were deleted, the deletion of the whole

gene was also tolerated.[313] Single gene deletions ac46 and ac91 displayed equal EYFP fluorescence compared to the control, however the AmCyan levels were slightly lower in all three rounds of testing and hence will not be included in the deletion plan. Both of the homolog genes for ac91 and ac46 were deleted in BmNPV (bm74 and bm37, respectively) resulting in viable viruses however a slight decrease in protein expression was observed for ac46, which is also true based on the results gained in this study.[285] Gene deletions ac149 and ac44 only displayed similar levels of protein expression compared to the control in the initial testing, with lower amounts of the proteins in the second and third round of testing, thus will be disregarded for the deletion plan. Though the deletions of the homolog genes for ac149 and ac44 in BmNPV (bm125 and bm35, respectively) resulted in viruses that appeared normal, the AcMNPV single gene deletion viruses were affected.[285]

Finally, it should be noted that all the single gene deletion viruses were tested with their overnight cultures for bacmid preparations grown in TB media instead of the normally used LB media, due to the presence of the blasticidin cassette. Although all the samples including the controls were treated in the same way, the growth of culture in TB media influenced the overall quality of the bacmid preparations hence the quality of the virus. In general, the viruses that originated from overnight cultures grown in TB media were approximately half a day to a day slower than ones grown in LB media (data not shown). The virus infection could have been slowed down due to the presence of more genomic DNA present in bacmid preparations during transfection as a result of the high-density *E. coli* growth of the overnight culture. It should be taken into consideration that the virus infection and protein expression is likely to differ once the BSD cassette will be removed for the confirmed genes. To avoid variations in the quality of the bacmid preparation, the use of a different antibiotic selection cassette that can grow in LB media, should be considered for future experiments.

4.4.4 Concluding Remarks and Future Directions

The generation and testing of 67 single gene knockout viruses of the SynBac1 baculoviral genome has been presented. The single gene deletions were achieved using homologous recombineering and the presence of the blasticidin cassette was confirmed by colony PCR of the baculoviral genome. The viruses were subjected to testing and the fluorescence levels of the expressed reporter proteins, EYFP and AmCyan, were measured and compared to the control, SynBac1, with the aim to scrutinise virus replication (EYFP) and late phase heterologous target protein production (AmCyan). Overall, the results obtained revealed that the original fragment design for the genome minimisation contained a considerable number of genes that could not be

deleted, and the fragments would not be reengineered for deletion. The results gathered in this study pointed to eight suitable candidates for gene deletion from the genome, which includes genes ac20, ac32, ac45, ac56, ac87, ac7, ac110 and ac112/113.

From the results presented in this chapter, five sites from the eight confirmed gene deletions, were earmarked to be tested for the relocation of the Tn7 transposition site to improve the stability of baculoviral genome. The Tn7 site is known to be a deletion hotspot in bacmid-based baculoviruses, presumably because it is embedded in a large section of prokaryotic DNA encoding for the replicon and associated modules.[314] In this approach, the Tn7 site will replace the gene using homologous recombineering, with an ampicillin resistance marker flanked by two RoxP sites. The antibiotic marker will be removed by the site-specific recombination between the RoxP sites and acting enzyme, Dre, leaving behind one RoxP site and the Tn7 site.[273,274] Additionally, a VloxP site will be inserted in place of one of the genes tested in this report as dispensable.

The confirmed single gene deletions that yielded a positive or neutral effect on the virus replication and target protein expression can then be combined to generate the minimised baculoviral genome. The genes should be deleted sequentially, and viruses tested at each of the stages to check for any negative effects when gene deletions are combined. The blasticidin cassette would be removed by a second homologous recombination event with a fragment containing just the homology arms. The reaction would be aided by the presence of the I-SceI restriction sites located on either side of the blasticidin cassette. The cassette also contains a rpsL gene that makes the normally streptomycin resistant DH10β *E. coli* cells sensitive to streptomycin, thus encouraging the removal of the cassette when streptomycin is present.[279]

Baculovirus can be used for gene therapy purposes, where it is used as a DNA cargo for numerous editing techniques, like CRISPR/Cas9.[315] The single gene deletions that yielded viruses with good EYFP levels but decreased AmCyan levels can be explored in the development of SynBacMam – a mammalian transduction virus. The viruses used for mammalian transduction require good virus replication, however the late phase protein expression is not needed as the genes of interest in the cargo are specifically mammalian expressed.

In the future, a second refined round of testing can be planned, which omits the already positively deleted genes and the predicted deleterious deletions. In this study, the single gene deletions were taken from the start of the open reading frame to the end, whereas in numerous studies beforehand, the classification of gene essentiality has been mostly determined by inactivating the

gene by inserted mutations or partial deletions. In those cases, it does not necessarily mean the genes are not essential for virus infection, as parts of the genes are still present but that the proteins expressed from these genes are not needed. For the genome minimization goal of this project, whole genes were taken to determine the maximum DNA sequences in the genome that can be deleted. However due to the close location and overlapping genes on the plus and minus strand of the baculoviral genome, the deletion of whole gene sequences proved to generate negative effects on the virus. The refined single 'gene' deletion test should include deleting smaller fragments than the original genes, where 30-50 bases are omitted on either side of the gene to avoid any effects on neighbouring genes and their promoter regions.

Chapter 5: Tn7 Site Relocation and VloxP/RoxP and EYFP Insertion

5.1 Introduction

The original, bold plan to minimise the baculoviral genome through rewriting large segments of wild-type code with synthetic, minimal code encountered the obstacles described in the previous chapters, necessitating a change in approach. This new approach utilised a progenitor genome, SynBac1, in which one DNA segment had been successfully rewired following the original approach, yielding a virus with wild-type virus kinetics and heterologous target protein expression capability. In the new approach, 67 single gene deletion baculoviruses were implemented using recombineering techniques and analysed. Of note, entire coding regions were deleted, earmarking overlapping genes or regulatory elements for a subsequent, future round of optimisation. One of the key objectives in our genome minimisation endeavour is to generate baculoviruses with improved stability during amplification and passaging to benefit industrial application where large virus volumes are required for bioreactors. One source of instability in baculoviral genomes relying on Tn7 transposition for heterologous gene insertion, including SynBac1, is the location of the Tn7 attachment site that is embedded in a swath of prokaryotic DNA and constitutes a hotspot of elimination.[97] Evidence exists that relocating the Tn7 attachment site into a safer site on the genome could be beneficial.[316] It was therefore reasoned in the present thesis that the gene loci identified as dispensable in the previous Chapter could represent such safer sites and thus could be exploited to harbour the Tn7 attachment site, adding stability and utility to the condensed baculoviral genome.

5.1.1 Tn7 Transposition Site Relocation

As previously described in the introduction of this thesis, Section 1.3, the instability of the baculoviral genome in serial passaging may stem from the location of the Tn7 transposition site (attTn7) within a large region of *E. coli* DNA in the baculoviral genome.[222] An approach taken to explore the stability of the Tn7 site was to relocate the attTn7 region, containing the lacZ cassette, but leave behind the resistance marker and bacterial origin of replication, the loss of which during amplification and upscaling would, probably, not adversely affect the baculovirions produced. In other words, loss of prokaryotic DNA in this region would be likely tolerated with no detrimental effects, as long as the DNA encoding the heterologous target protein, inserted by Tn7 transposition would be located in a safer site elsewhere. In summary, this could be instrumental in increasing genome stability and maintaining recombinant protein expression that would be likely impacted by deletions occurring in the prokaryotic DNA.

To test if the stability of the genome increases with the Tn7 site in a different genomic location, the current Tn7 site was removed, and four sites based on the single gene deletion screen were chosen for the Tn7 relocation. The eight highest performing single gene deletions from Chapter 4 were mapped onto the genome of SynBac1, (Figure 5.1 blue circles). The hr regions (grey circles) were also mapped, as one criterion for choosing the potential location for the Tn7 relocation is the proximity of the site to a hr region. The importance of the hr regions has been described in Section 1.1.4.1 in more details, but briefly the sites act as the virus's origins of replication, therefore are not exposed to as higher levels of spontaneous deletion as other regions in the genome.[73,80,84,85] The placement of the Tn7 site in the vicinity of a hr region may therefore reduce its likelihood of elimination during virus passaging. Based on the location of the eight genes and the hr regions, the four genes chosen for deletion and relocation of the Tn7 site were ac32, ac87, ac110 and ac112/113, with ac32 and ac87 close to hr2 and hr3, respectively and both ac110 and ac112/113 located between hr4a and hr4b. The ac45 location for the Tn7 site was also explored by another member of the lab, although the DPA and virus kinetics were similar to a virus containing the original Tn7 site, the virus with the Tn7 site at ac45 displayed lower protein expression (data shown in results section). Ac7 was not chosen as one of the potential sites for Tn7 relocation as the original Tn7 site (green circle) is located nearby and the repositioning of the site would not be significant. Ac56 was also not chosen for the relocation of the Tn7 site as it is not in close proximity to any of the hr regions and the other four locations performed better in terms of protein expression in the single gene deletion screen but may be considered in the future if further sites need to be examined. The remaining site, ac20, is similarly not positioned close to hr regions and was instead chosen for the insertion of a site-specific recognition sequence (details in next section).

Prior to the commencement of this work, the current Tn7 site needed to be removed and the work performed was undertaken by a fellow lab member (H. Crocker). In brief, a plasmid containing homology arms corresponding to the region surrounding the current lacZ_attTn7 cassette, with an AmpR flanked by I-CueI sites was utilised for the removal of the original lacZ_attTn7 region. The removal of the existing Tn7 site was performed using the Red/ET homologous recombination, as previously utilised in earlier baculoviral modifications in this thesis. The AmpR marker was subsequently removed by restriction digestion and re-ligation at the I-CeuI sites. The mutant lacking the Tn7 site was confirmed by the absence of blue colonies on IPTG and BluOGal plates as a result of the removal of the lacZ thus inability to process the BluOGal substrate into a visible blue colour. The mutant was further confirmed by colony PCR

of the region of modification and sequencing of the PCR fragment. As the Tn7 transposition site was removed, no experiments about the heterologous protein expression capabilities were performed. The mutant lacking the original Tn7 site was then utilised for the re-insertion of the lacZ_attTn7 cassette at the four sites chosen earlier.

Due to the original fragment modifications not taking place, the site-specific recombination systems previously explored for the removal of the selection markers were available for use. One of the site-specific recombination systems, Dre/RoxP, was utilised as a means to successfully integrate the new Tn7 site through antibiotic selection, with the resistance marker removed by the presence of the flanking RoxP sites and Dre treatment.

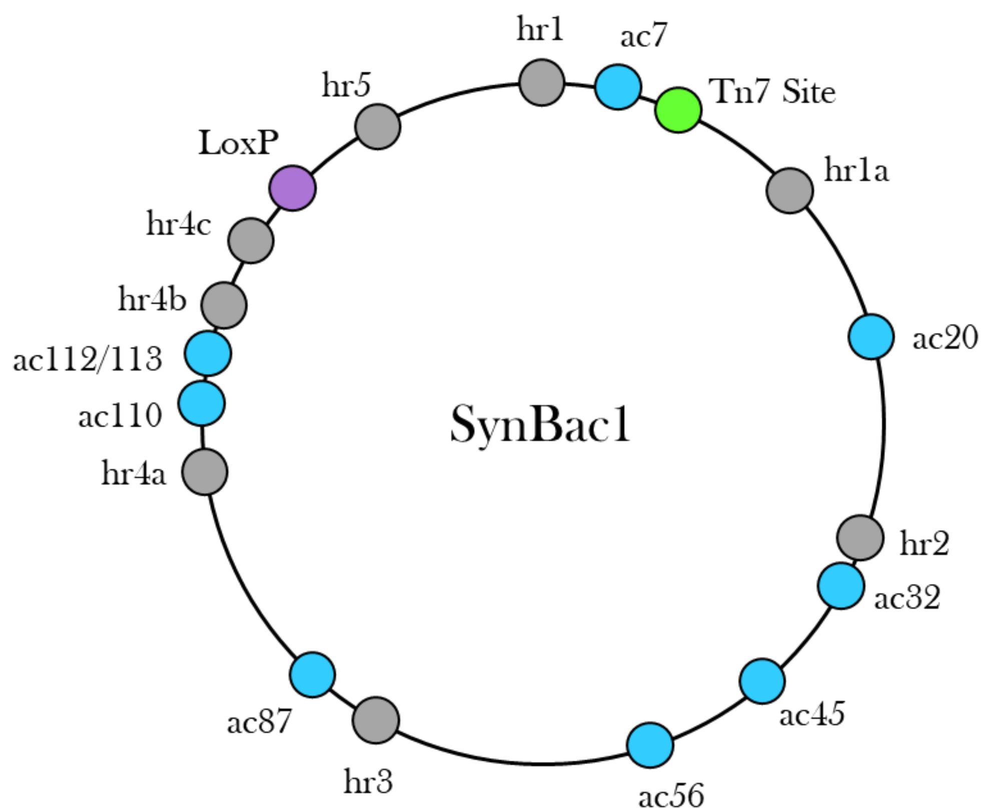


Figure 5.1 – Spatial distribution of hr regions and potential genes for deletion and insertion of the Tn7, VloxP and RoxP sites.

Schematic representation of the SynBac1 genome displaying the location of the hr regions (grey circles), potential genes (blue circles) for the insertion of the Tn7 site and VloxP and RoxP sites, as well as the current location of the Tn7 site (green circle) and the LoxP site (purple circle).

5.1.2 VloxP and RoxP Insertion

One of the confirmed single gene deletions, ac20, would be utilised for the insertion of an additional site-specific recognition site. The VloxP/VCre system was chosen as alongside the RoxP/Dre system, they were the best performing in terms of excision of antibiotic marker and donor-acceptor fusions that were most likely due to the cleanest prep of the purified recombinase enzymes (Figure 3.7 (A)). In case none of the new relocated Tn7 site viruses perform better in terms of protein expression and stability compared to the original Tn7 site, the RoxP site would be inserted at one of the locations, ac87. Originally the RoxP site would be inserted alongside the Tn7 modification, however if the results show the original Tn7 site is better, the addition of the RoxP site would already be in place and no further cloning would need to be designed and implemented later. We have decided to only implement two out of the four site-specific recombination systems previously explored in Chapter 3, Section 3.3.1.3. Having two additional sites for the insertion of supplementary modalities should be sufficient for a working baculoviral expression system, and various laboratories rarely require the need to add more than two extra sequences on top of their protein of interest expressed through the Tn7 site.

5.1.3 EYFP Insertion

An additional modification implemented in this chapter is the re-insertion of the EYFP reporter gene. The EYFP present in the established EMBacY genome is under control of the polh promoter,[317,318] however the EYFP inserted within the rewired Fragment 1 modification in SynBac1 is under an early promoter, ie1. To be able to compare the SynBac1 genomes with the different Tn7 locations to EMBacY, the expression of the fluorescent marker needs to be under the control of the same promoter. The different promoters are active in different stages of the virus replication, with ie1 being an earlier promoter and polh being a very late promoter.[319,320] The strength levels of the open reading frame translation and expression also differs with the two promoters, as the polh promoter is vastly more active and results in a higher readout of the EYFP fluorescence. Therefore, the promoter controlling the EYFP expression needs to be changed in the SynBac1 genomes to be able to properly compare with EMBacY as the benchmark in the forthcoming stability experiments.

The EYFP was previously removed from the SynBac1 genome with no Tn7 present by another member of the lab. Similarly to the removal of the original Tn7 site, a Red/ET reaction was performed to remove the EYFP cassette, and the ampicillin selection marker excised by the presence of flanking LoxP sites and Cre treatment. The reaction yielded a virus with one

remaining LoxP site (purple circle in Figure 5.1) that would be utilised for the re-insertion of the EYFP under the control of the polh promoter. The mutant lacking the Tn7 site and the EYFP cassette was used for the work in this chapter. Therefore, in the testing stage of the mutant clones prior to EYFP re-insertion, no EYFP fluorescence was measured but instead only AmCyan fluorescence.

5.2 Chapter Aims

The aim of this chapter is to apply the information acquired from the deletion of the individual genes in Chapter 4 and utilise the genes which can be deleted as potential sites for genome engineering. Additionally, an EYFP reporter gene is to be inserted into the backbone of the baculoviral genome. This will be explored through the following approaches:

1. Relocate the Tn7 transposition site to four potential locations.
2. Insert VloxP and RoxP sequences as additional sites for future engineering.
3. Insert an EYFP reporter gene at the LoxP site, which contains the same promoter as in EMBacY.

5.3 Results

5.3.1 VloxP, RoxP and Tn7 Fragment Preparation

Firstly, the plasmids utilised for the amplification of the VloxP, RoxP and Tn7 cassettes for the Red/ET reaction needed to be designed. Due to the presence of the RoxP and VloxP sites twice in the design and their repetitive nature, primer design for cloning of the cassettes in house would have caused difficulties, therefore the synthesis of the plasmids was outsourced to GenScript. The cassettes for the insertion of the VloxP and RoxP sites were designed to contain the homology arms corresponding to the ac20 and ac87 genes, respectively, with an AmpR flanked by the VloxP or RoxP sites, (Figure 5.2 (A)). The cassettes were cloned through EcoRV digestion into the pDS plasmid. The cassettes for the relocation of the Tn7 site were designed to contain the homology arms corresponding to the ac32, ac87, ac110 and ac112/113 genes, the lacZ_attTn7 cassette, with an AmpR flanked by the RoxP sites, (Figure 5.2 (B)). The cassettes were cloned through PmeI digestion into the pDS plasmid. For sequences information, refer to Section 8.2.3 in the Appendix.

To prepare the fragments used for the Red/ET reactions, primers were designed to amplify the cassettes by binding just outside of the homology arms. The agarose gel (Figure 5.2 (C)) displays the PCR amplified fragments for insertion of the VloxP and RoxP sites, ac20VloxP (1237 bp) and ac87RoxP (1261 bp), as well as the four fragments for relocation of the Tn7 site, ac32Tn7, ac87Tn7, ac110Tn7 and ac112/113Tn7 (2199 bp).

The fragments were then used in the Red/ET reaction as previously stated in the methods and prior chapters. In brief, the fragments were electroporated into cells induced with L-arabinose and containing the pRed/ET plasmid and appropriate baculoviral genome, SynBac1 with original Tn7 for VloxP and RoxP insertion and SynBac1 lacking the original Tn7 for Tn7 relocation. The AmpR was subsequently removed by incubation of the baculoviral DNA with the Dre or VCre recombinase and checked for successful removal of the AmpR.

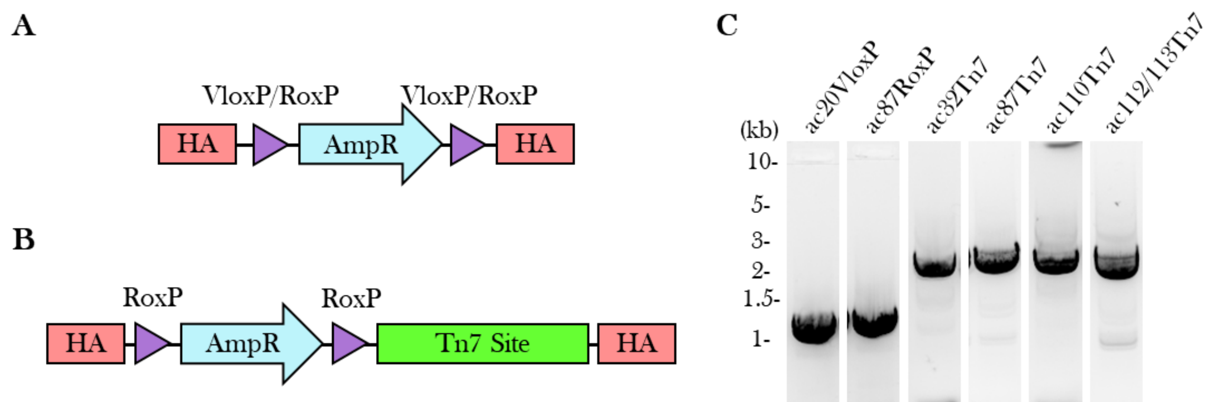


Figure 5.2 VloxP/RoxP and Tn7 site cassette design and preparation.

(A-B) Schematic representation of the designed cassettes for the insertion of the **(A)** VloxP (ac20) and RoxP (ac87) sites and **(B)** Tn7 site (ac32, ac87, ac110 and ac112/113). The ampicillin marker is flanked by the site-specific recombination recognition sites for its subsequent removal by the corresponding recombinase. **(C)** 1% agarose gel of PCR amplified fragments utilised in the Red/ET reaction: ac20VloxP (1237 bp, ac20VloxPFor + Rev primers), ac87RoxP (1261 bp, ac87RoxPFor + Rev primers) and the four (ac32, ac87, ac110 and ac112/113) Tn7 site fragments (2199 bp, Tn7FragFor + Rev primers).

5.3.2 VloxP and RoxP Insertion into SynBac1 with Original Tn7 Site

5.3.2.1 Ac20VloxP Insertion

The first modification undertaken was the insertion of the VloxP site at the ac20 position in the baculoviral genome. Multiple clones were screened by PCR to confirm correct size of the fragment after modification, with the original and intermediate genome validated as well. Figure 5.3 (A) displays an example of the bands obtained by PCR validation from one of the clones including the original SynBac1 genome (1042 bp), the intermediate VloxP and AmpR at ac20 genome (1944 bp) and the final SynBac1 ac20VloxP genome (866 bp). Several clones were included in the standard testing of virus kinetics and protein expression by transforming AmCyan into the Tn7 site and measuring fluorescence during V1 amplification. The AmCyan fluorescence data (Figure 5.3 (B)) from samples taken at 48 and 72 hpi displays the baculovirus mutant clones containing the VloxP site instead of the ac20 gene perform as well, if not slightly better, especially at 48 hpi, than the control genome, SynBac1. Although most of the clones were functional and could be taken for the next stage of the modification, clone #5 was chosen, and the correct DNA sequence present containing the VloxP site at ac20 was validated by sequencing the PCR fragments.

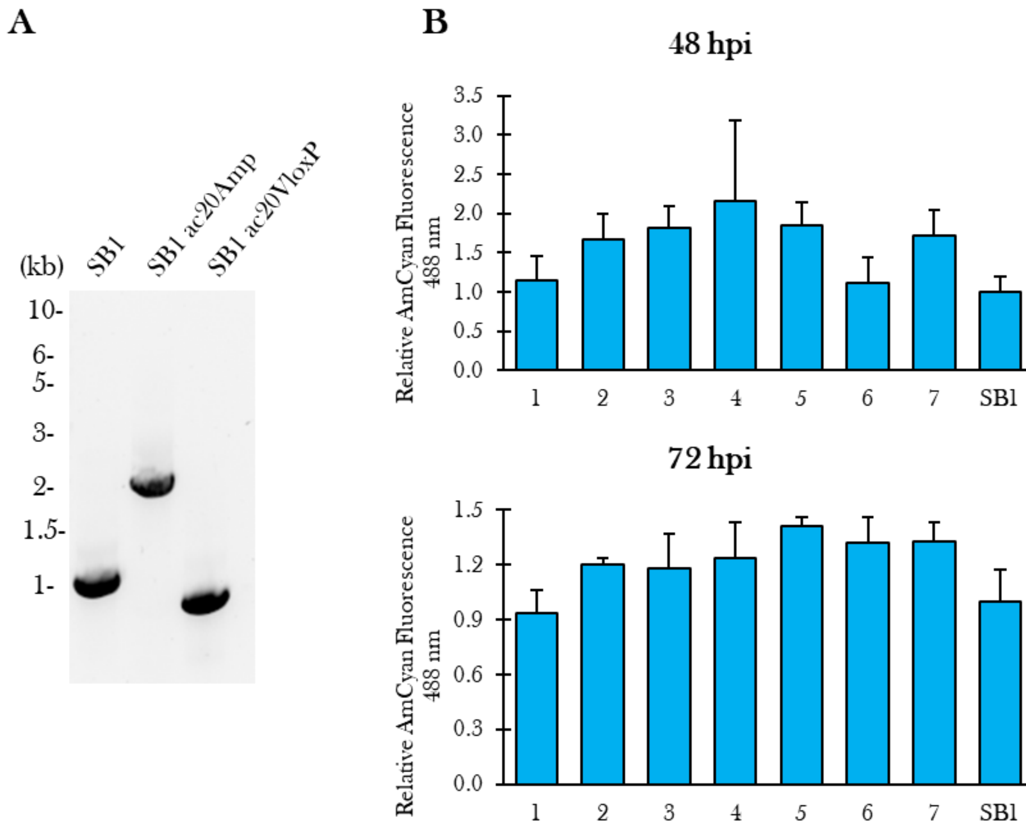


Figure 5.3 – Ac20VloxP insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac20 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 (1042 bp), the intermediate genome, SynBac1 ac20Amp (1944 bp), and the final genome, SynBac1 ac20VloxP (866 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#5) for further proceedings. (10/11CheckFor and Check10Rev primers). **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.2.2 Ac87RoxP Insertion

The genome containing the VloxP site at ac20 (Clone #5) was then taken to insert the RoxP site at ac87. Like previously multiple clones were screened by PCR, with Figure 5.4 (A) displaying the expected PCR sizes of the bands acquired from the original genome, SynBac1 ac20VloxP (1255 bp), the intermediate genome containing the RoxP and AmpR at ac87 (1982 bp), and the final genome, SynBac1 ac20VloxP ac87RoxP (906 bp). AmCyan expression and fluorescence during V1 amplification was recorded for several clones, including the precursor genome containing the VloxP site at ac20 and the original SynBac1 variant. The fluorescence data (Figure

5.4 (B)) for the 48 and 72 hpi time points displayed most of the clones had similar fluorescence intensities to the control genomes. One of the clones did not exhibit any fluorescence, which may be due to a faulty white colony picked for the bacmid preparation and thus no AmCyan expression. Nevertheless, all the remining clones showed comparable protein expression to the controls and after confirmation of the DNA sequence at the ac87 location, Clone #6 was chosen for further testing.

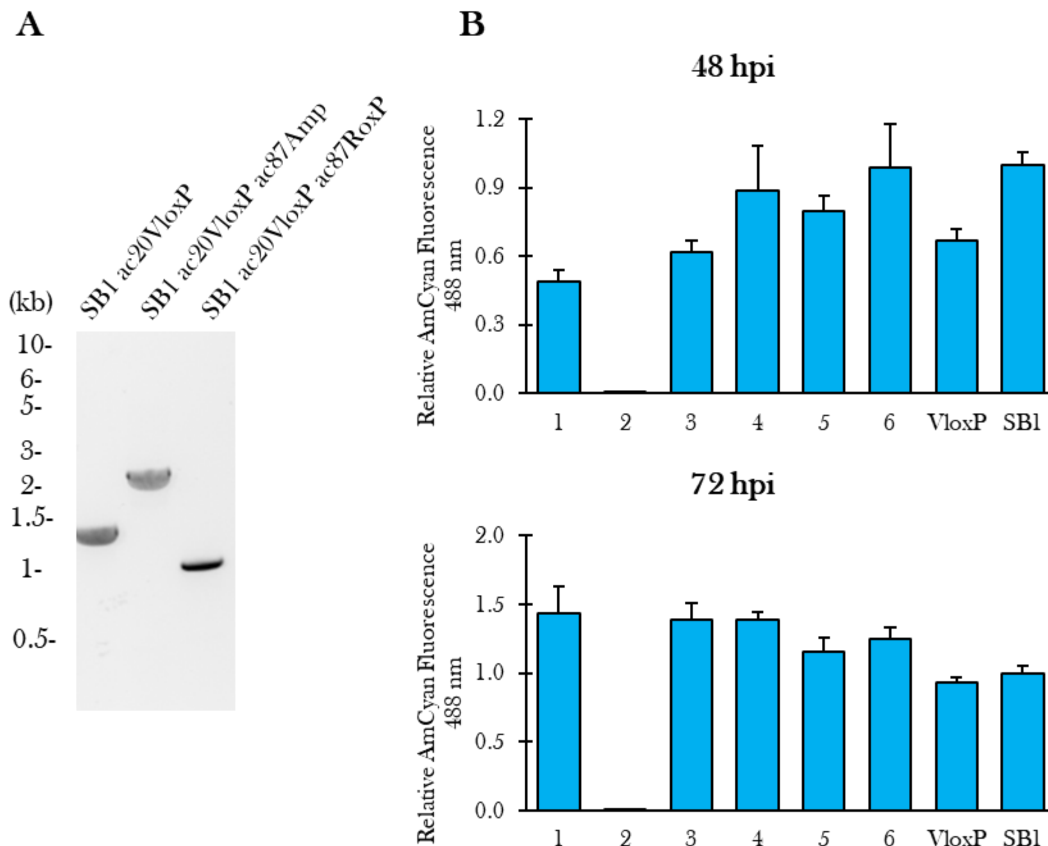


Figure 5.4 – Ac87RoxP insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac87 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 ac20VloxP (1255 bp), the intermediate genome, SynBac1 ac20VloxP ac87Amp (1982 bp), and the final genome, SynBac1 ac20VloxP ac87RoxP (906 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#6) for further proceedings. (38/39CheckFor and ac87Tn7CheckRev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3), VloxP is the precursor genome, SynBac1 ac20VloxP clone 5.

5.3.2.3 Reassessment of Best SynBac1 Ac20VloxP Ac87RoxP Clone

As the nature of the virus preparations can differ between separate experiments, the SynBac1 genome with the VloxP at ac20 and RoxP at ac87 (Clone 6) was tested again. Re-examining the virus kinetics and protein expression capabilities during V1 amplification confirms the values of fluorescence measured are still within the same levels, compared to the control genome. Figure 5.5 displays the AmCyan fluorescence data measured from 48 and 72 hpi samples during V1 amplification of the SynBac1 ac20VloxP ac87RoxP virus and the control virus, SynBac1. Although the levels of fluorescence are lower for the SynBac ac20VloxP ac87RoxP virus at 48 hpi, the values reach near levels of fluorescence at 72 hpi compared to the control genome. The newly generated baculoviral genome with the VloxP and RoxP sites at ac20 and ac87, respectively, was confirmed in terms of protein expression levels and was utilised afterwards for the EYFP insertion.

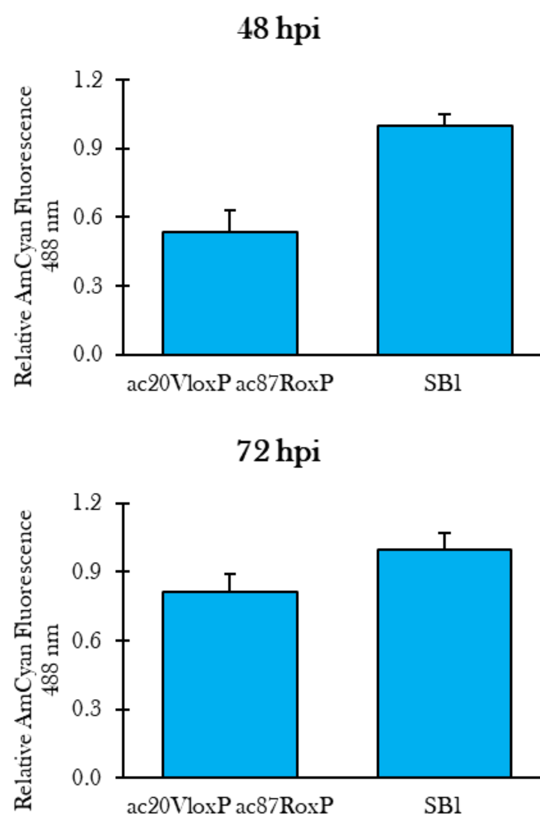


Figure 5.5 – Reassessment of SynBac1 ac20VloxP ac87RoxP.

AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac20VloxP ac87RoxP clone 6. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.3 Tn7 Site Relocation

Alongside the insertion of the site-specific sites to the SynBac1 genome with the original Tn7, the baculoviral genome containing no Tn7 was used for the relocation of the Tn7 transposition site to four different locations in the genome. Multiple clones (or as many as available that were confirmed by PCR), were tested by measuring the AmCyan fluorescence levels during V1 amplification of the viruses. The following sections display the fluorescence results gained and the PCR validation, for the virus genomes with the four different Tn7 locations, ac32, ac87, ac110 and ac112/113. Sequencing was performed to confirm the modification at the four locations and to check no mutations were present within the lacZ_attTn7 cassette.

5.3.3.1 Ac32Tn7

Figure 5.6 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 (1666 bp), the intermediate genome containing the Tn7 site and AmpR at ac32 (3170 bp), and the final genome, SynBac1 ac32Tn7 (2094 bp). The fluorescence data (Figure 5.6 (B)) for the 48 and 72 hpi time points display both the clones tested exhibit similar levels of AmCyan fluorescence compared to the SynBac1 control genome. As Clone #1 showed a slightly higher level of fluorescence, it was saved for further work including re-testing of the protein expression and EYFP insertion.

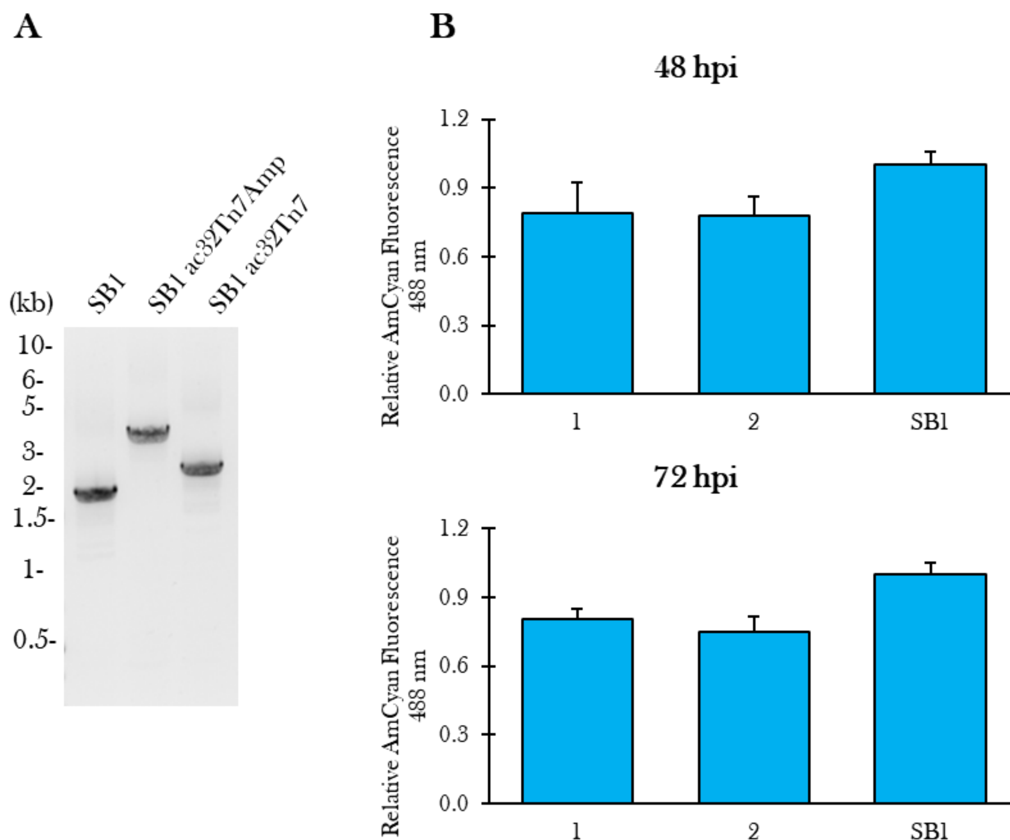


Figure 5.6 – Ac32Tn7 insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac32 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 (1666 bp), the intermediate genome, SynBac1 ac32Tn7Amp (3170 bp), and the final genome, SynBac1 ac32Tn7 (2094 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#1) for further proceedings. (14CheckFor and ac32Tn7CheckRev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac32Tn7 clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.3.2 Ac87Tn7

Figure 5.7 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 (1255 bp), the intermediate genome containing the Tn7 site and AmpR at ac87 (2924 bp), and the final genome, SynBac1 ac87Tn7 (1848 bp). The fluorescence data (Figure 5.7 (B)) for the 48 and 72 hpi time points display there was a greater variance of the AmCyan fluorescence measured between the clones. Clones #1, #2 and #4 showed fluorescence levels of half or less compared to the control virus. The fluorescence measured for clone #3 was slightly higher, but just below the 80% threshold, which was set as a

guideline for a functional virus. Clone #5 however displayed AmCyan fluorescence that was slightly higher compared to the control, and as a result was chosen for future work.

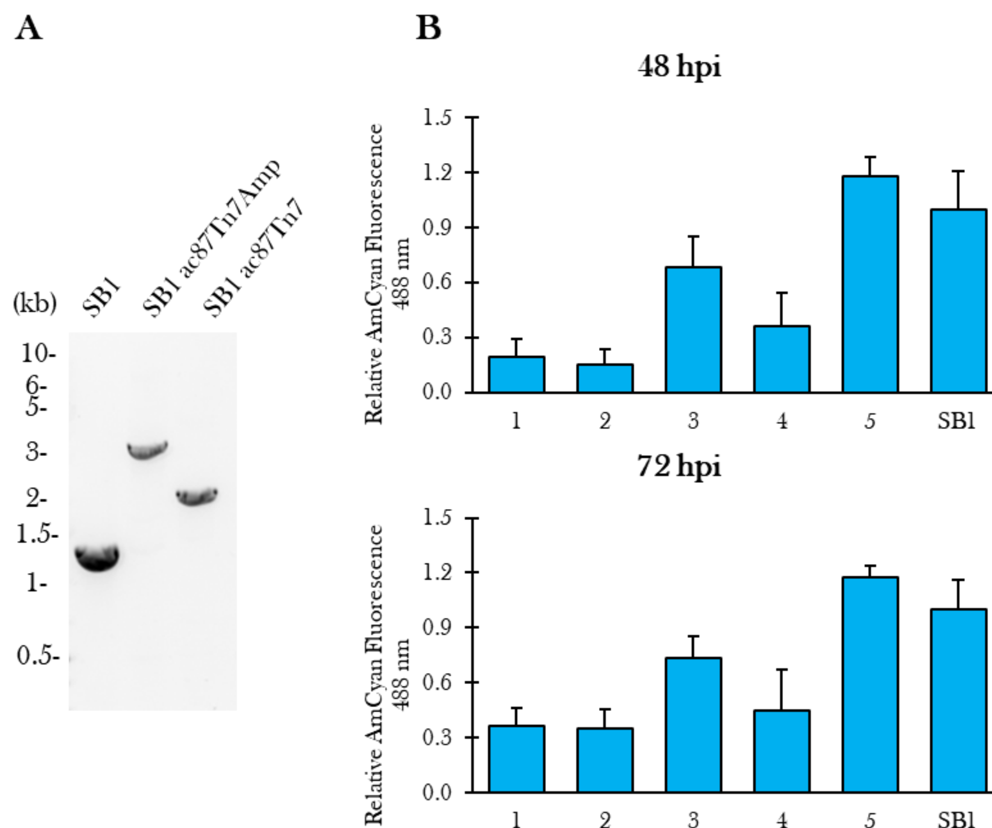


Figure 5.7 – Ac87Tn7 insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac87 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 (1255 bp), the intermediate genome, SynBac1 ac87Tn7Amp (2924 bp), and the final genome, SynBac1 ac87Tn7 (1848 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#5) for further proceedings. (38/39CheckFor and ac87Tn7CheckRev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac87Tn7 clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.3.3 Ac110Tn7

Figure 5.8 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 (795 bp), the intermediate genome containing the Tn7 site and AmpR at ac110 (2674 bp), and the final genome, SynBac1 ac110Tn7 (1598 bp). The AmCyan fluorescence data (Figure 5.8 (B)) from samples taken at 48 and 72 hpi displays the

baculovirus clones containing the Tn7 site at position ac110. All but one of the clones (#3) display positive AmCyan fluorescence levels, which all reached similar levels to the control virus at 72 hpi. Clone #3 was likely the result of a faulty white colony, which was indeed blue and did not contain the AmCyan cassette in the Tn7 transposition site. As clone #7 displayed a slightly higher fluorescence value compared to the other two best performing clones, #5 and #6, it was saved for further proceedings.

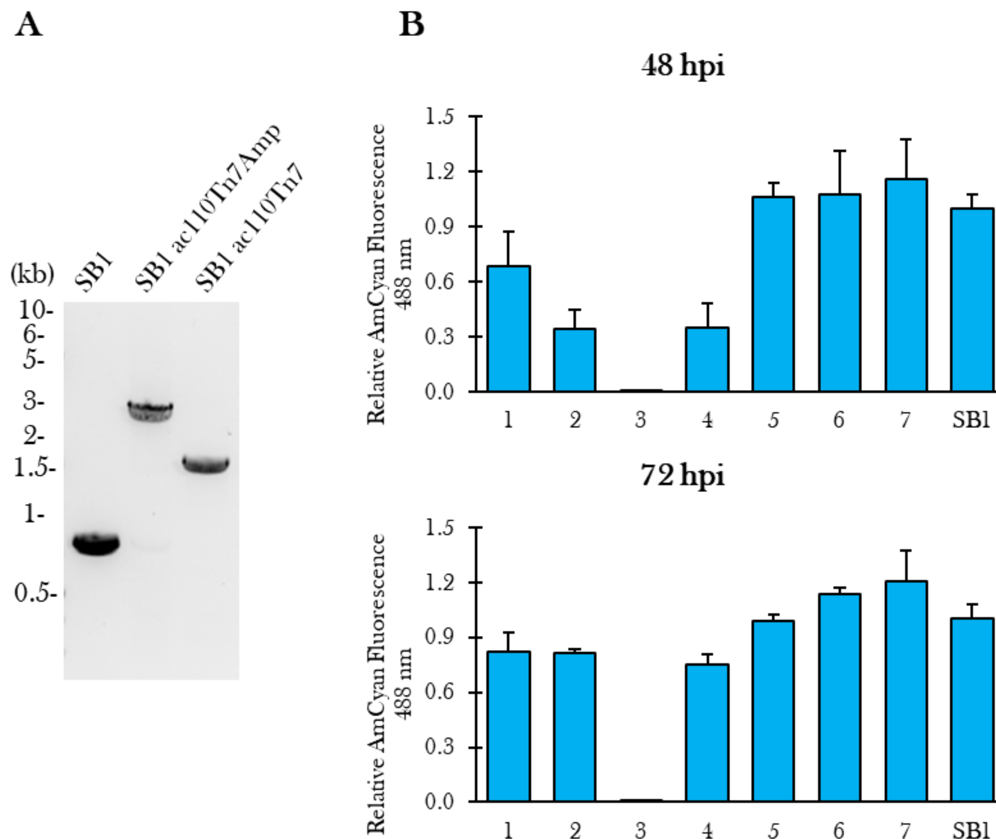


Figure 5.8 – Ac110Tn7 insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac110 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 (795 bp), the intermediate genome, SynBac1 ac110Tn7Amp (2674 bp), and the final genome, SynBac1 ac110Tn7 (1598 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#7) for further proceedings. (65CheckFor and ac110Tn7CheckRev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac110Tn7 clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.3.4 Ac112/113Tn7

Figure 5.9 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 (1986 bp), the intermediate genome containing the Tn7 site and AmpR at ac112/113 (3258 bp), and the final genome, SynBac1 ac112/113Tn7 (2182 bp). The fluorescence data (Figure 5.9 (B)) for the 48 and 72 hpi time points display most of the clones are at a similar level in AmCyan expression compared to the SynBac1 control genome. Most of the clones show higher fluorescence values at 48 hpi compared to the control, however the values even out to a similar level as the control at 72 hpi. The higher values at 48 hpi may be a result of a virus with faster virus kinetics, hence quicker protein expression. Another reason could be the control virus exhibiting a slightly slower start in the V1 amplification; however, the control triplicates are all within the same range of fluorescence. The highest performing clone, #5, was chosen for further work.

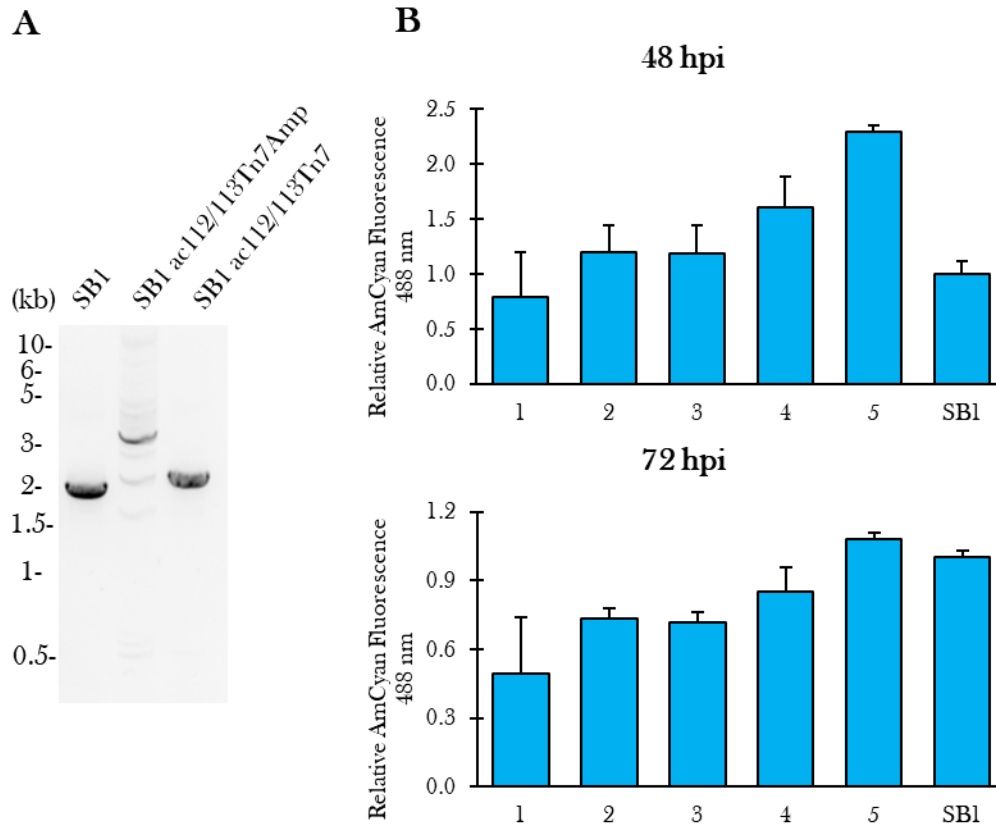


Figure 5.9 – Ac112/113Tn7 insertion.

(A) 1% agarose gel showing a representation of the PCR validation of the ac112/113 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 (1986 bp), the intermediate genome, SynBac1 ac112/113Tn7Amp (3258 bp), and the final genome, SynBac1 ac112/113Tn7 (2182 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#5) for further proceedings. (67CheckFor and 67CheckRev2 primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac112/113Tn7 clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.3.5 Reassessment of Best Clones for New Tn7 Sites

The best performing clones from the previous examinations of the baculoviral genomes with four differently located Tn7 sites were re-tested to confirm the protein expression capabilities. The genome with a fifth location of the Tn7 site at ac45, which was generated by a colleague, was additionally tested alongside the newly generated Tn7 variants. The ac45Tn7 virus was included to illustrate the lower levels of protein expression achieved, hence the need to test the four other locations. The AmCyan fluorescence data (Figure 5.10) from samples taken at 48 and 72 hpi

displays there was more variance in fluorescence measured compared to the first time the clones were tested. Although, the AmCyan values were lower, the four Tn7 mutants generated in this thesis reached 70-90% fluorescence compared to the SynBac1 control at 72 hpi, with the ac32Tn7 genome equal to the control at 48 hpi. The ac45Tn7 virus displayed half the fluorescence compared to the control at both the time points, with a wider variation (error bars) in measured values at 48 hpi.

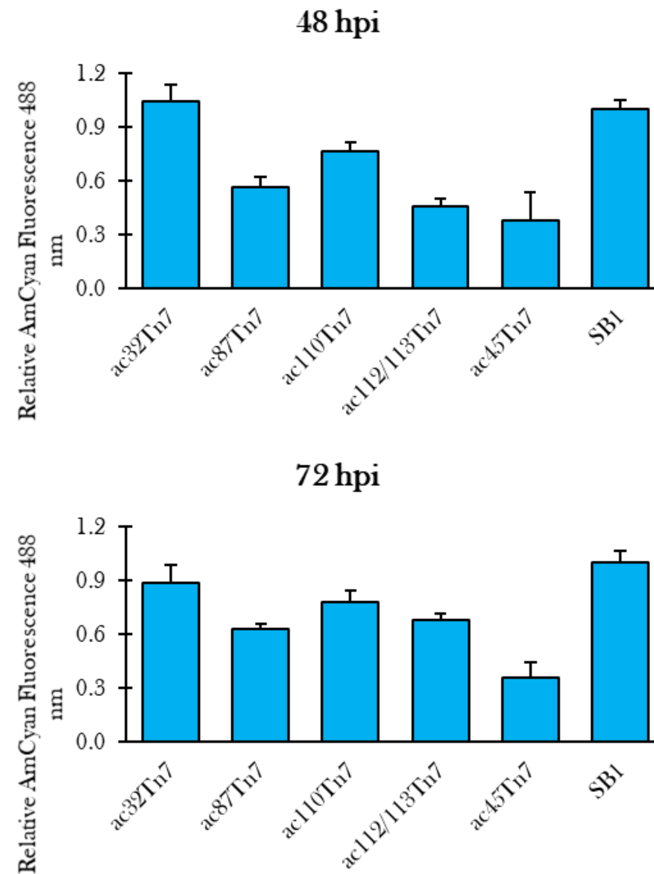


Figure 5.10 – Reassessment of the different SynBac1 Tn7 genomes.

AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac32, ac87, ac110, ac112/113 and ac45 (generated by H. Crocker) Tn7 genomes. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.4 VloxP Insertion into SynBac1 containing New Tn7 Sites

The next step involved taking the genomes with the four different Tn7 sites and inserting the VloxP site at the ac20 position. Four clones for each of the varying Tn7 sites containing the VloxP modification at ac20 were screened by measuring the AmCyan fluorescence during V1 amplification of the virus, as well as PCR validating the ac20 region at each step of the procedure. The following sections exemplify the results acquired for the four Tn7 viruses (ac32, ac87, ac110 and ac112/113) with VloxP at ac20. The modifications were confirmed by sequencing the ac20 region to make sure no mutations were present in the VloxP sequence.

5.3.4.1 Ac32Tn7 Ac20VloxP

Figure 5.11 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac32Tn7 (1042 bp), the intermediate genome containing the VloxP site and AmpR at ac20 (1944 bp), and the final genome, SynBac1 ac32Tn7 ac20VloxP (866 bp). The fluorescence data (Figure 5.11 (B)) for the 48 and 72 hpi time points display some variation between the clones for fluorescence values at 48 hpi, however the values increase to a level close to the control virus at 72 hpi. Clone #2 displayed the highest AmCyan fluorescence, with values close to the control even at 48 hpi, hence this clone was saved for further testing.

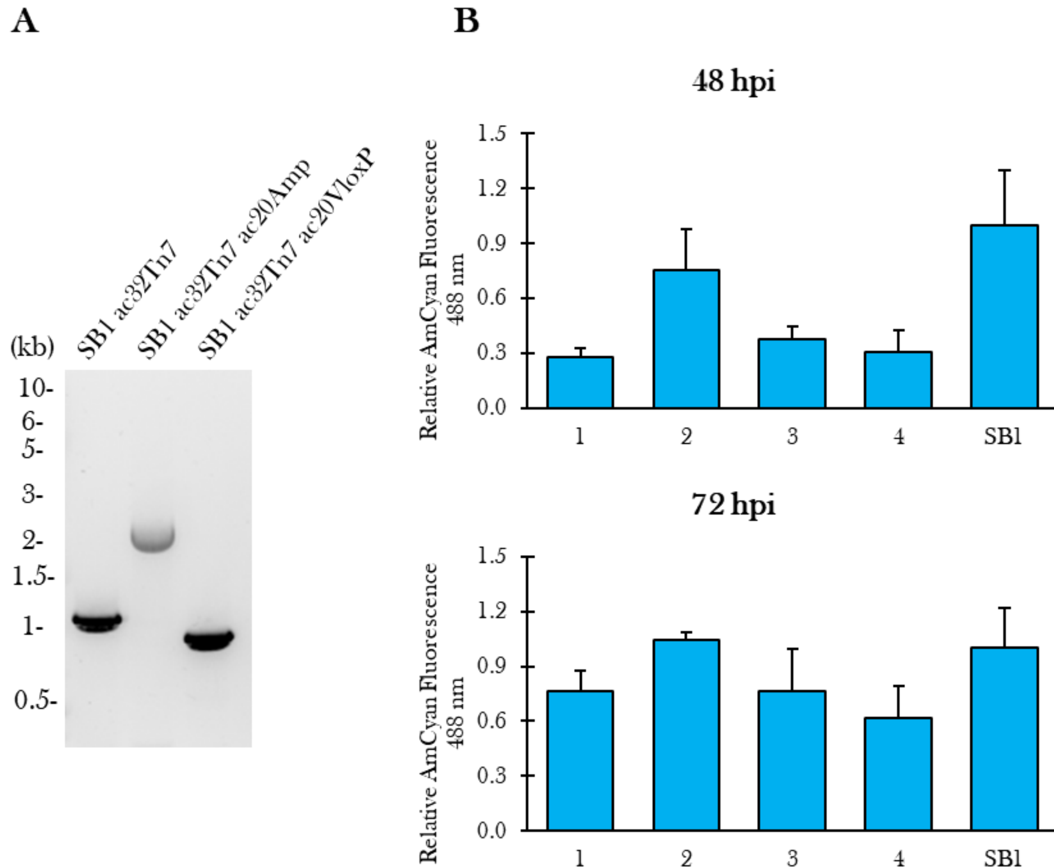


Figure 5.11 – Ac20VloxP insertion into SynBac1 Ac32Tn7.

(A) 1% agarose gel showing a representation of the PCR validation of the ac20 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 ac32Tn7 (1042 bp), the intermediate genome, SynBac1 ac32Tn7 ac20Amp (1944 bp), and the final genome, SynBac1 ac32Tn7 ac20VloxP (866 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#2) for further proceedings. (10/11CheckFor and Check10Rev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac32Tn7 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.4.2 Ac87Tn7 Ac20VloxP

Figure 5.12 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac87Tn7 (1042 bp), the intermediate genome containing the VloxP site and AmpR at ac20 (1944 bp), and the final genome, SynBac1 ac87Tn7 ac20VloxP (866 bp). The AmCyan fluorescence data (Figure 5.12 (B)) from samples taken at 48 and 72 hpi show the ac87Tn7 virus with VloxP at ac20 has lower fluorescence values at 48 hpi compared to the control virus, however three of the clones reach the same levels at 72 hpi. Clone

#2 displayed no fluorescence, which once again might have been a result of a false positive white colony. As clone #1 displayed the highest amounts of AmCyan expression at 72 hpi, it was chosen to proceed with future work.

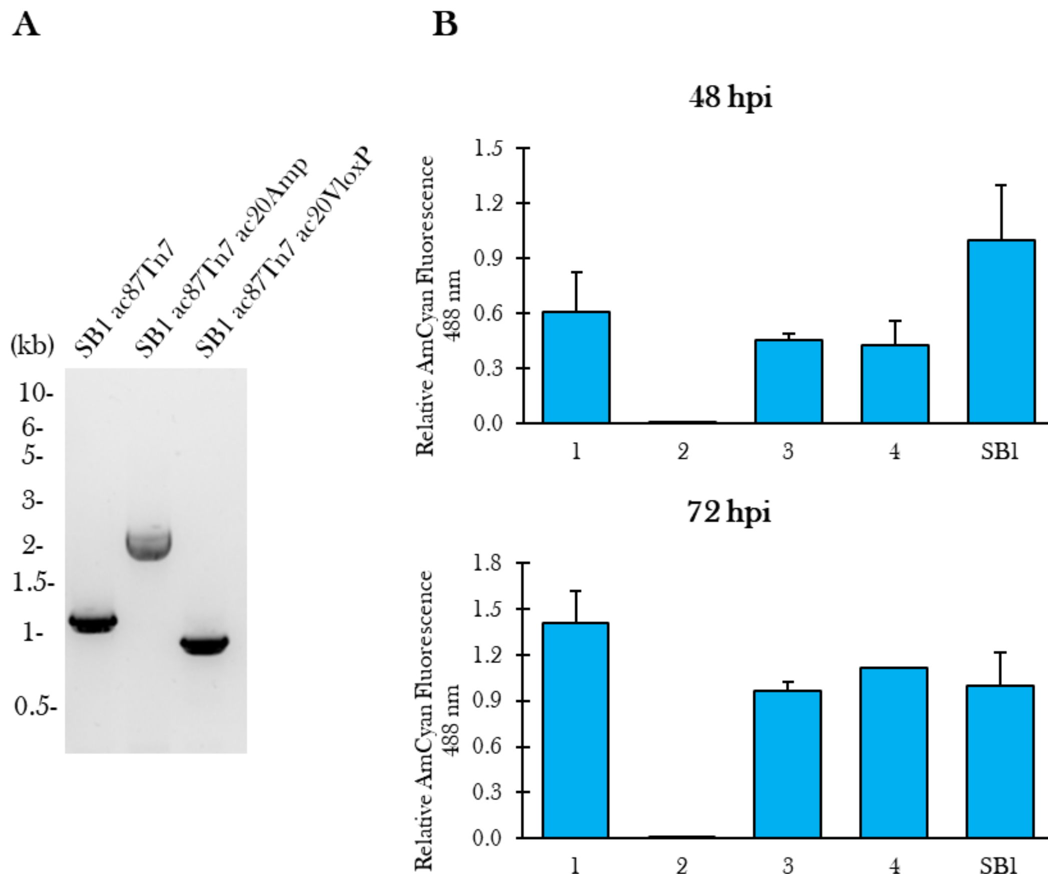


Figure 5.12 – Ac20VloxP insertion into SynBac1 Ac87Tn7.

(A) 1% agarose gel showing a representation of the PCR validation of the ac20 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 ac87Tn7 (1042 bp), the intermediate genome, SynBac1 ac87Tn7 ac20Amp (1944 bp), and the final genome, SynBac1 ac87Tn7 ac20VloxP (866 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#1) for further proceedings. (10/11CheckFor and Check10Rev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac87Tn7 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.4.3 Ac110Tn7 Ac20VloxP

Figure 5.13 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac110Tn7 (1042 bp), the intermediate genome

containing the VloxP site and AmpR at ac20 (1944 bp), and the final genome, SynBac1 ac110Tn7 ac20VloxP (866 bp). The fluorescence data (Figure 5.13 (B)) for the 48 and 72 hpi time points show all four of the clones reach equal if not higher levels of AmCyan expression at 72 hpi, although are lower at 48 hpi compared to the SynBac1 control genome. Clone #2 displayed the highest fluorescence at 48 hpi, therefore was chosen as the clone to proceed with further work.

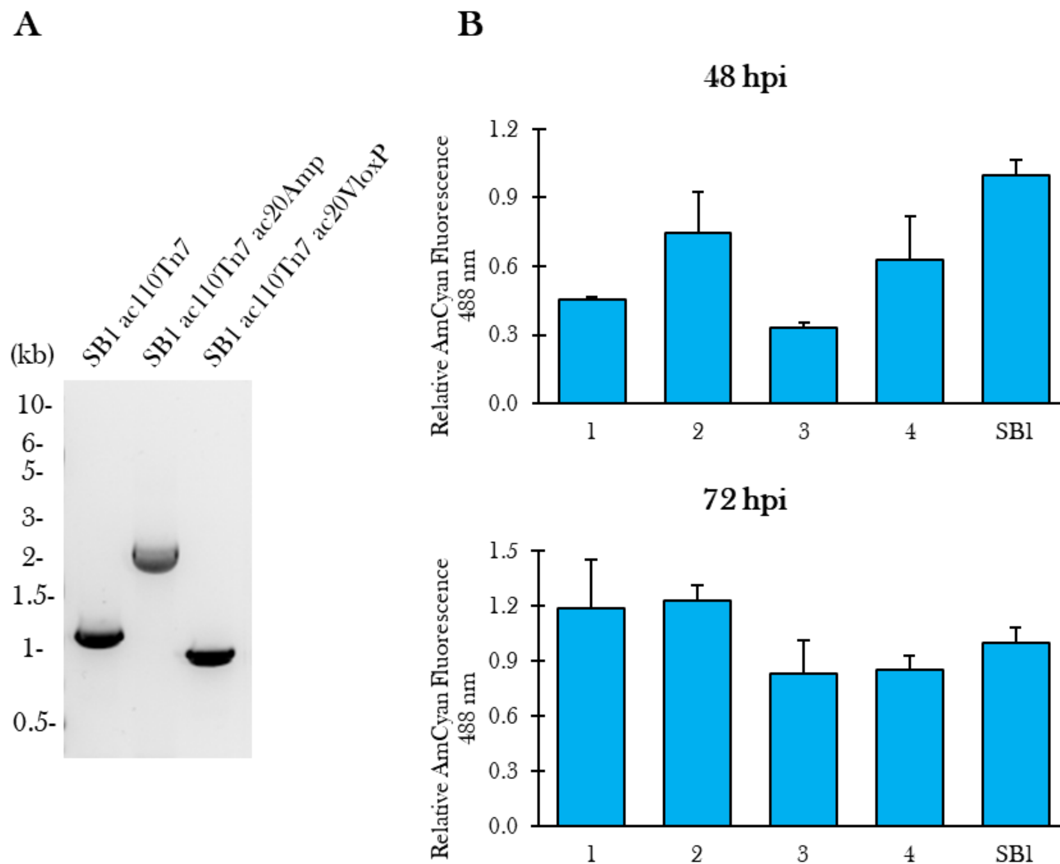


Figure 5.13 – Ac20VloxP insertion into SynBac1 Ac110Tn7.

(A) 1% agarose gel showing a representation of the PCR validation of the ac20 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 ac110Tn7 (1042 bp), the intermediate genome, SynBac1 ac110Tn7 ac20Amp (1944 bp), and the final genome, SynBac1 ac110Tn7 ac20VloxP (866 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#2) for further proceedings. (10/11CheckFor and Check10Rev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac110Tn7 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.4.4 Ac112/113Tn7 Ac20VloxP

Figure 5.14 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac112/113Tn7 (1042 bp), the intermediate genome containing the VloxP site and AmpR at ac20 (1944 bp), and the final genome, SynBac1 ac112/113Tn7 ac20VloxP (866 bp). The AmCyan fluorescence data measured from samples taken at 48 and 72 hpi during V1 amplification of the virus is displayed in Figure 5.14 (B). Similarly to previous tests, one clone (#4) displayed no measurable fluorescence, most likely due to picking incorrect colony for the bacmid preparation. The remaining three clones displayed lower values for AmCyan at 48 hpi however the values increased at 72 hpi, with clone #1 reaching the 80% threshold compared to the control genome and taken further for re-examination.

Figure 5.14 – Ac20VloxP insertion into SynBac1 Ac112/113Tn7.

(A) 1% agarose gel showing a representation of the PCR validation of the ac20 region in the baculoviral genome. The PCR products show the expected fragment sizes at each step of the modification, beginning with the starting genome, SynBac1 ac112/113Tn7 (1042 bp), the intermediate genome, SynBac1 ac112/113Tn7 ac20Amp (1944 bp), and the final genome, SynBac1 ac112/113Tn7 ac20VloxP (866 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#1) for further proceedings. (10/11CheckFor and Check10Rev primers) **(B)** AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac112/113Tn7 ac20VloxP clones under inspection. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.4.5 Reassessment of Best Clones for VloxP Insertion into New Tn7 Site Viruses

The best performing clones of the VloxP site insertion at the ac20 position in the four different Tn7 containing viruses were re-examined to verify the protein expression achieved. The AmCyan fluorescence data acquired for samples taken at 48 and 72 hpi during V1 amplification is displayed in Figure 5.15. Unlike some of the Tn7 variants performing slightly worse at 48 hpi in the initial screen, the 48 hpi values measured in this reassessment showed more similarity to the SynBac1 control genome. All four variants reached comparable AmCyan expression at 72 hpi, with ac110Tn7 virus showing marginally lower fluorescence, however the ac32Tn7 virus displayed higher fluorescence at both time points compared to the control.

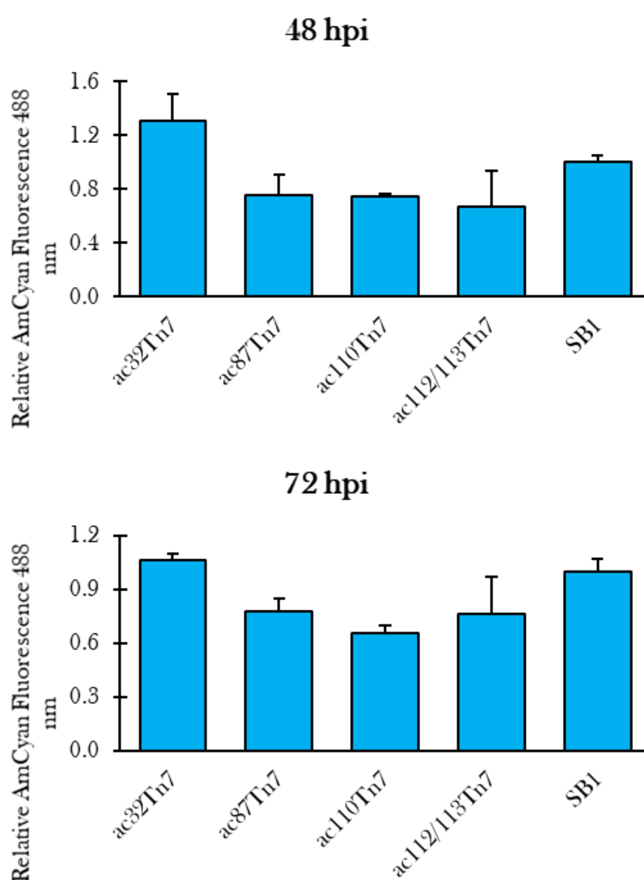


Figure 5.15 – Reassessment of the different SynBac1 Tn7 ac20VloxP genomes.

AmCyan production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac32, ac87, ac110, ac112/113 Tn7 ac20VloxP genomes. Values represent averages normalized to SynBac1 (SB1) = 1; error bars indicate SEM (n=3).

5.3.5 EYFP Insertion

As the protein expression capabilities were validated for all four of the Tn7 VloxP containing viruses, the work to insert the EYFP cassette could proceed confidently. The EYFP would be inserted into the four new Tn7(ac32, ac87, ac110 and ac112/113) ac20VloxP viruses and the virus with the original Tn7 containing VloxP and RoxP at ac20 and ac87, respectively, as well as the precursor control virus SynBac1 with original Tn7 site. Prior to starting the insertions, the plasmid containing the EYFP cassette was modified by cutting out the unused p10 promoter, as it has been shown the polh promoter is more efficient without the competition of the p10 promoter nearby.[321-323] The pUCDM_YFP plasmid was cut with NheI and SpeI, which removed the region containing the p10 promoter, and the plasmid ligated back together with the compatible sticky ends. The sequence for the plasmid containing no p10 promoter (pUCDM_YFP_polh) can be found in the appendix in section 8.2.3.7.

The modified plasmid was then used in fusion reactions by incubating baculoviral DNA with the plasmid, Cre recombinase and the Cre buffer as per methods section. The optimal amount of baculoviral DNA to acquire enough colonies needed to be optimised, and ranged from 3 to 6 µg, with concentrations above 8 µg yielding less colonies than reactions containing the same DNA preparation of lesser concentration. The successful insertion of the pUCDM_YFP_polh plasmid into the LoxP site located on the baculoviral genome was confirmed by colony PCR. Three clones for each of the variants were taken for testing by measuring the EYFP fluorescence during V1 amplification of the viruses. As the EYFP would provide a readout about the virus functionality, the AmCyan cassette was not transposed into the Tn7 site to save time by not performing transformation reactions into multiple clones. EMBacY was the control virus used in the testing of the clones as the virus contains the EYFP cassette under the control of the polh promoter. The following sections show the EYFP data gathered from testing three clones for each of the six different baculoviral genomes:

1. SynBac1 original Tn7
2. SynBac1 original Tn7 ac20VloxP ac87RoxP
3. SynBac1 ac32Tn7 ac20VloxP
4. SynBac1 ac87Tn7 ac20VloxP
5. SynBac1 ac110Tn7 ac20VloxP
6. SynBac1 ac112/113Tn7 ac20VloxP

5.3.5.1 SynBac1 Original Tn7

The bands visualised in Figure 5.16 (A) display the expected products of the PCR validation for the starting genome, SynBac1 (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP production at 48 and 72 hpi during V1 amplification is displayed (Figure 5.16 (B)). The EYFP fluorescence at 48 hpi was under half the amount compared to the EMBacY control virus, however the values increased at 72 hpi. Although the fluorescence was under the 80% threshold, the best performing clone (#2) was chosen and the EYFP values re-tested later to verify the amounts of protein expression detected.

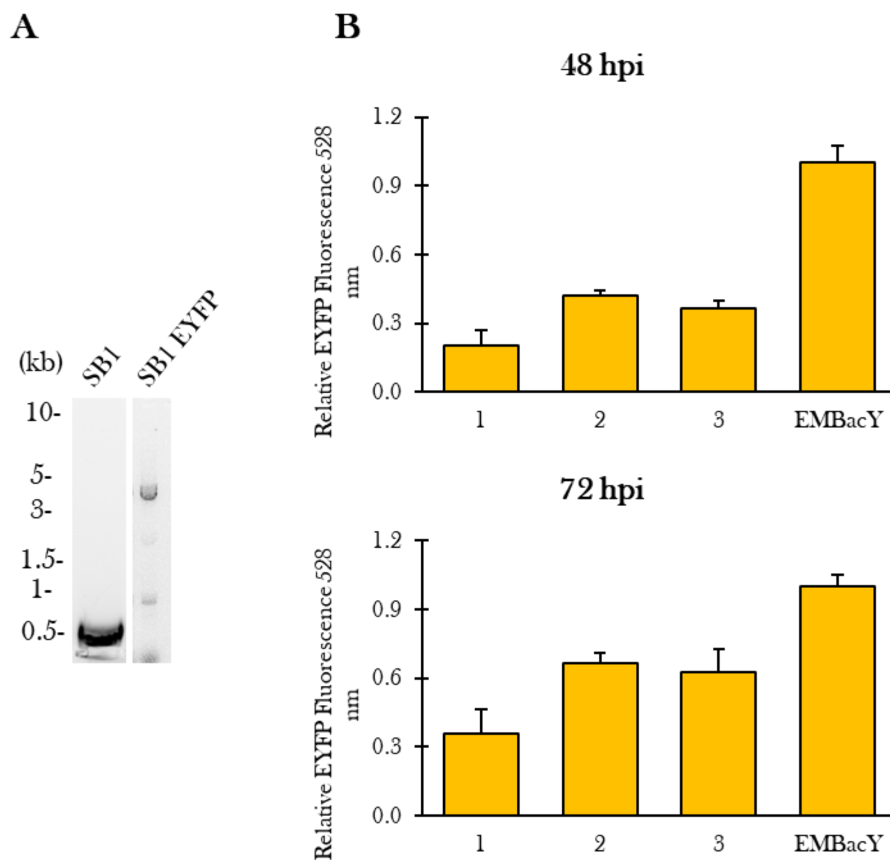


Figure 5.16– EYFP insertion into SynBac1.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#2) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.2 SynBac1 Original Tn7 Ac20VloxP Ac87RoxP

Figure 5.17 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac20VloxP ac87RoxP (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP fluorescence data (Figure 5.17 (B)) for the 48 and 72 hpi time points show one clone (#3) reached the same levels of EYFP expression at both time points compared to the control, thus was saved for further validation. The other two clones displayed half the amount of EYFP fluorescence compared to the EMBacY virus.

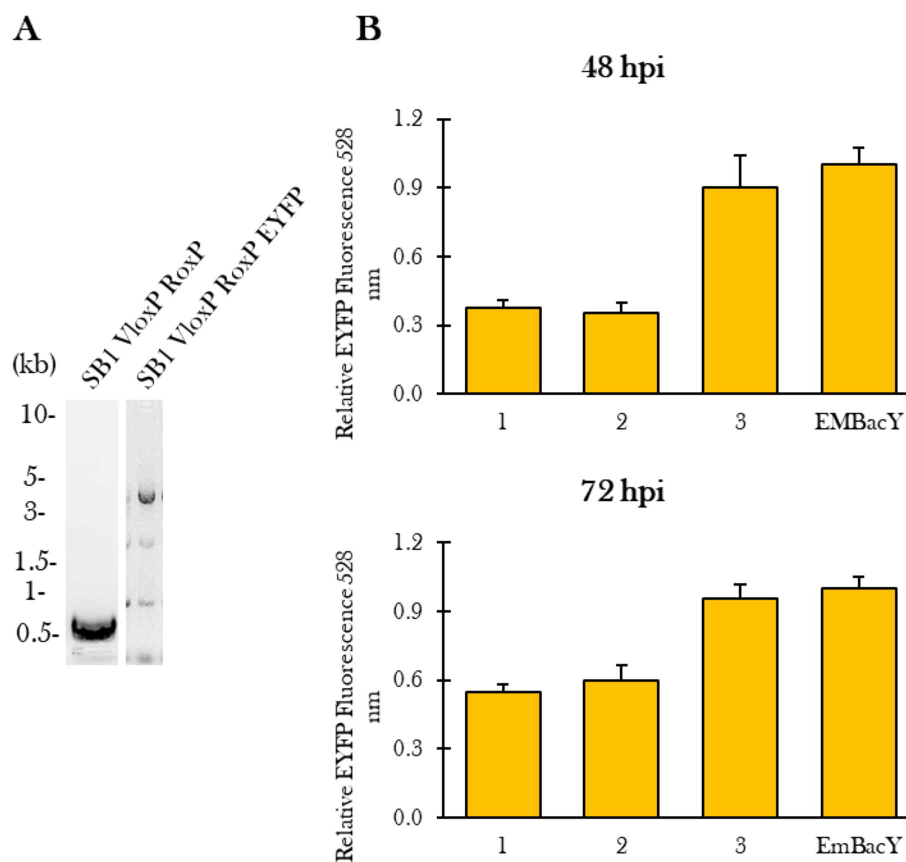


Figure 5.17– EYFP insertion into SynBac1 ac20VloxP ac87RoxP.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 ac20VloxP ac87RoxP EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#3) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac20VloxP ac87RoxP EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.3 SynBac1 Ac32Tn7 Ac20VloxP

The bands visualised in Figure 5.18 (A) display the expected products of the PCR validation for the starting genome, SynBac1 ac32Tn7 ac20VloxP (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP production at 48 and 72 hpi during V1 amplification is displayed in Figure 5.18 (B). All three clones showed similar EYFP fluorescence compared to the EMBacY control virus at both time points. As clone #2 reached the highest EYFP levels at 72 hpi, it was chosen for further processing.

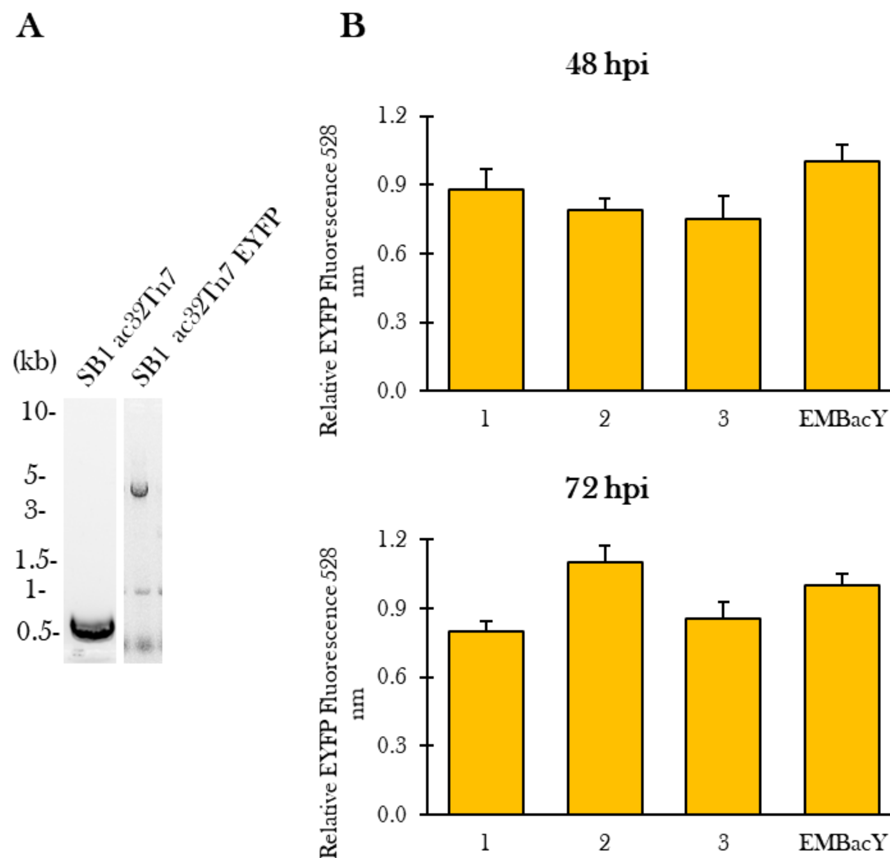


Figure 5.18– EYFP insertion into SynBac1 ac32Tn7 ac20VloxP.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 ac32Tn7 ac20VloxP EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#2) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac32Tn7 ac20VloxP EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.4 SynBac1 Ac87Tn7 Ac20VloxP

Figure 5.19 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac87Tn7 ac20VloxP (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP fluorescence data (Figure 5.19 (B)) for the 48 and 72 hpi time points show the three clones chosen for the testing did not perform well due to the very low EYFP fluorescence detected. As the re-testing of the other baculovirus genome variants was ongoing, the best performing clone (#3) was taken for re-examination nevertheless, while new clones were prepared for testing.

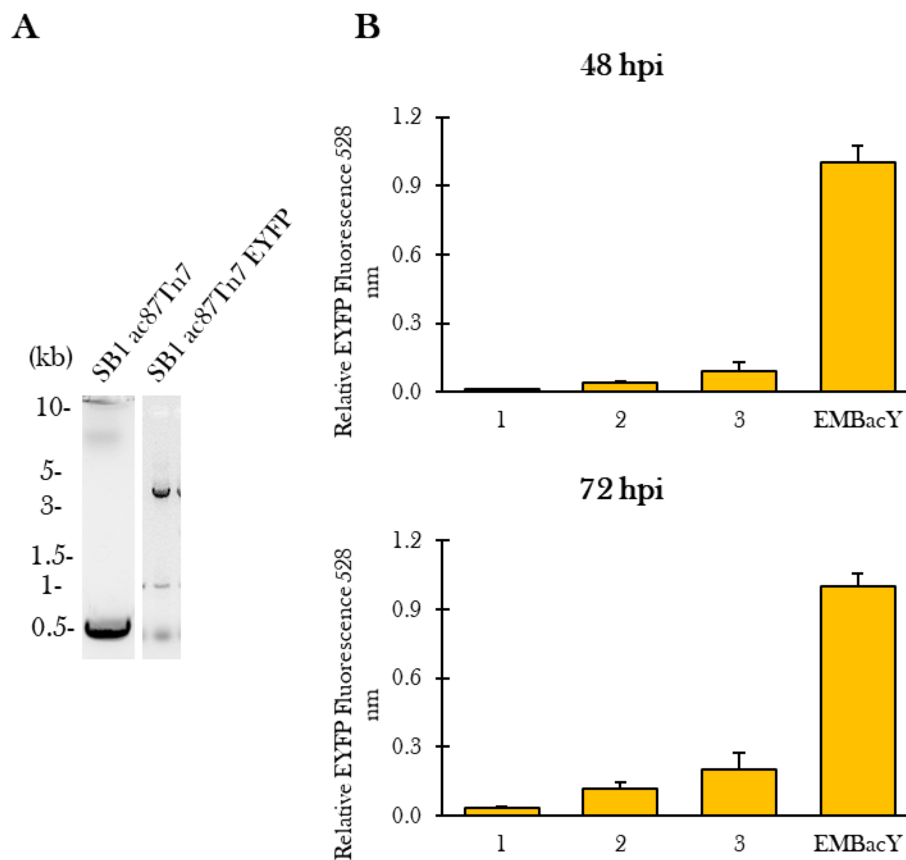


Figure 5.19– EYFP insertion into SynBac1 ac87Tn7 ac20VloxP.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 ac87Tn7 ac20VloxP EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#3) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac87Tn7 ac20VloxP EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.5 SynBac1 Ac110Tn7 Ac20VloxP

The bands visualised in Figure 5.20 (A) display the expected products of the PCR validation for the starting genome, SynBac1 ac110Tn7 ac20VloxP (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP production at 48 and 72 hpi during V1 amplification is shown in Figure 5.20 (B). The three clones showed low EYFP fluorescence at 48 hpi compared to the control virus, however the values increased slightly at 72 hpi. Although the EYFP expression was lower than hoped for, clone #3 that displayed around 60% EYFP value to the control was taken for further analysis while new clones were prepped for testing.

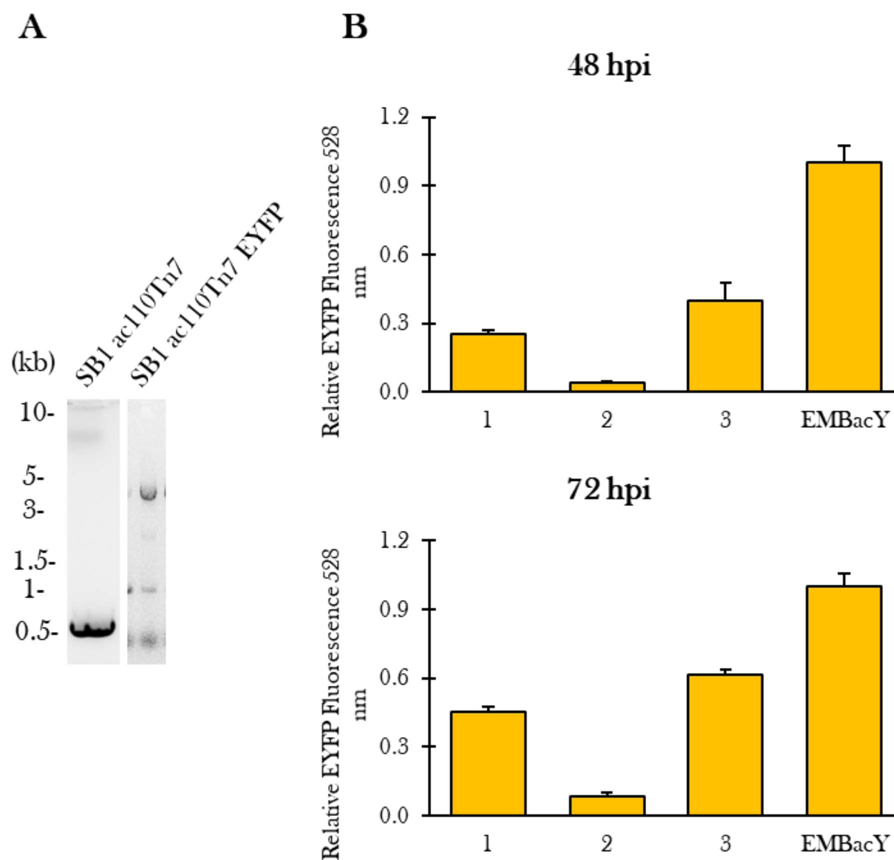


Figure 5.20– EYFP insertion into SynBac1 ac110Tn7 ac20VloxP.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 ac110Tn7 ac20VloxP EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#3) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac110Tn7 ac20VloxP EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.6 SynBac1 Ac112/113Tn7 Ac20VloxP

Figure 5.21 (A) displays an example of the expected PCR band sizes for one of the clones acquired from the original genome, SynBac1 ac112/113Tn7 ac20VloxP (542 bp), and the final genome containing the inserted EYFP cassette (4088 bp). The EYFP fluorescence data (Figure 5.21 (B)) for the 48 and 72 hpi time points showed the clones similarly experienced lower EYFP levels at 48 hpi as did the ac110Tn7 virus and increased slightly at 72 hpi. Overall, the three clones reached around 60% fluorescence compared to the EMBacY virus at 72 hpi, and clone #2 was taken for re-evaluation of the EYFP levels.

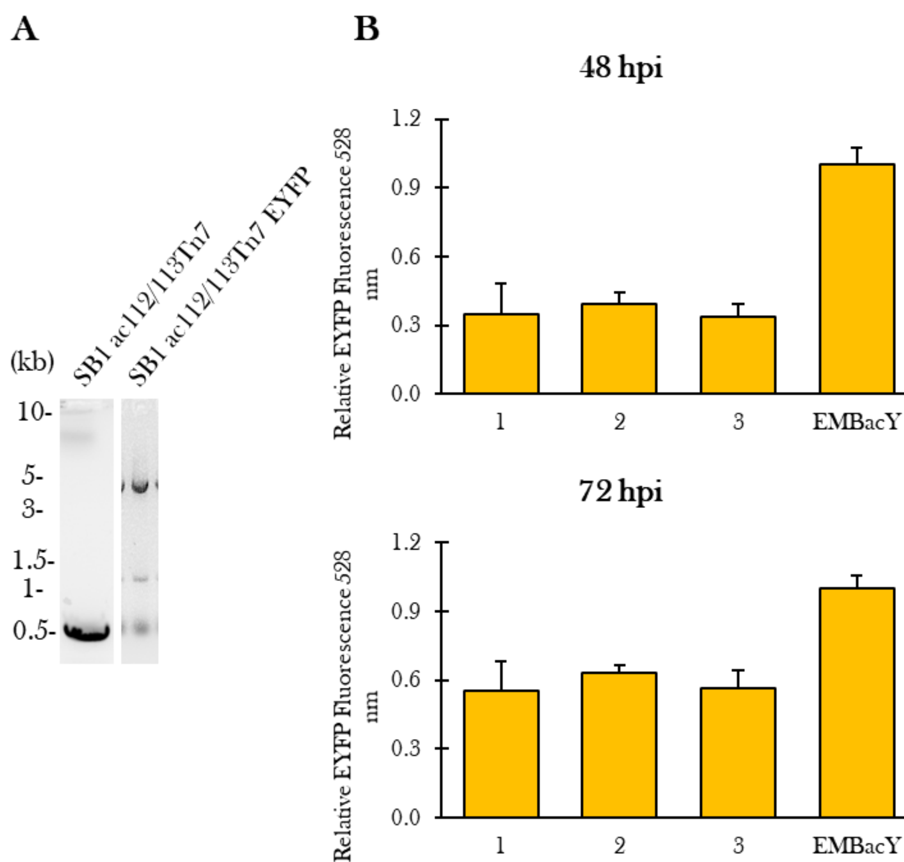


Figure 5.21 – EYFP insertion into SynBac1 ac112/113Tn7 ac20VloxP.

(A) 1% agarose gel showing a representation of the PCR validation of the EYFP insertion region in the baculoviral genome. The PCR products show the expected fragment sizes for the starting genome, SynBac1 (542 bp) and the final genome, SynBac1 ac112/113Tn7 ac20VloxP EYFP (4088 bp). Each clone under examination was verified by PCR, with the band displayed above for the final genome, representing the chosen clone (#2) for further proceedings. (YFPCheckFor and YFPCheckRev primers) **(B)** EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 ac112/113Tn7 ac20VloxP EYFP clones under inspection and the control genome, EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.3.5.7 Best EYFP Insertion Clones

The best performing clones from the initial screen of the various baculoviral genomes after EYFP insertion were examined again to confirm the results. The genome variants and clones selected were:

1. SynBac1 original Tn7 – Clone 2
2. SynBac1 original Tn7 ac20VloxP ac87RoxP - Clone 3
3. SynBac1 ac32Tn7 ac20VloxP – Clone 2
4. SynBac1 ac87Tn7 ac20VloxP – Clone 3
5. SynBac1 ac110Tn7 ac20VloxP – Clone 3
6. SynBac1 ac112/113Tn7 ac20VloxP – Clone 2

For this evaluation, the AmCyan plasmid was transformed into the Tn7 site and the AmCyan fluorescence measured as well as the EYFP during V1 amplification of the viruses. Figure 5.22 displays the fluorescence data acquired from 48 and 72 hpi samples for EYFP (A) and AmCyan (B) for the six viruses and the control, EMBacY.

The EYFP values for SynBac1 with original Tn7 (SB1) were around 70% previously, however only reached 40% compared to EMBacY in this screen. The lower value seen was a result of a higher variation between the triplicates (error bars), where one of the triplicate viruses performed worse than the other two. Nevertheless, the AmCyan value at 72 hpi for SynBac1 with original Tn7 still reached around 70% compared to the control.

The SynBac1 with original Tn7 and VloxP and RoxP at ac20 and ac87, respectively, (SB1 VloxP RoxP) performed worse than previously, as the values for EYFP fluorescence were the same as EMBacY in the first screen, however only around 60% in this screen. The AmCyan measured at 72 hpi did however reach similar values compared to the control virus.

Similarly to last time the SynBac1 with Tn7 site at ac32 and VloxP at ac20 (SB1 ac32Tn7), reached equal values of EYFP fluorescence compared to EMBacY. Additionally, the AmCyan protein expression was around 80 and 100% at 48 and 72 hpi, respectively, compared to the control virus.

The SynBac1 ac87Tn7 ac20VloxP virus (SB1 ac87Tn7) did not perform well in the initial screen, however the EYFP values reached 60 and 70% at 48 and 72 hpi, respectively, compared to the control virus in this re-evaluation. Although the AmCyan fluorescence measured at 48 hpi was half the amount of EMBacY, the value increased at 72 hpi to around 70%.

While both the viruses with Tn7 sites at ac110 and ac112/113 (SB1 ac110Tn7 and SB1 ac112/113Tn7) reached EYFP levels of around 60% at 72 hpi compared to EMBacY in the previous screen, the values were equal to EMBacY at 48 hpi. The amounts then slightly decreased at 72 hpi, which might have been due to the faster virus kinetics experienced thus reaching maximum EYFP expression earlier. Both ac110 and ac112/113 Tn7 viruses displayed equal values for AmCyan fluorescence at 48 and 72 hpi compared to EMBacY.

The re-examination of the EYFP fluorescence measured from the six baculovirus variants, with the additional data of the AmCyan protein expression capabilities was sufficient for progressing to the next stage of validating the baculovirus genomes. The need to test more clones for the EYFP insertion stage, especially for SB1 ac87Tn7 was not necessary as the virus performed better than last time, however the variations between the EYFP values measured for future experiments should be recorded. Overall, the six newly generated SynBac1 baculoviral genomes were comparable to EMBacY (subjective to a statistical analysis), with ac32Tn7, ac110Tn7 and ac112/113Tn7 at the top of the group.

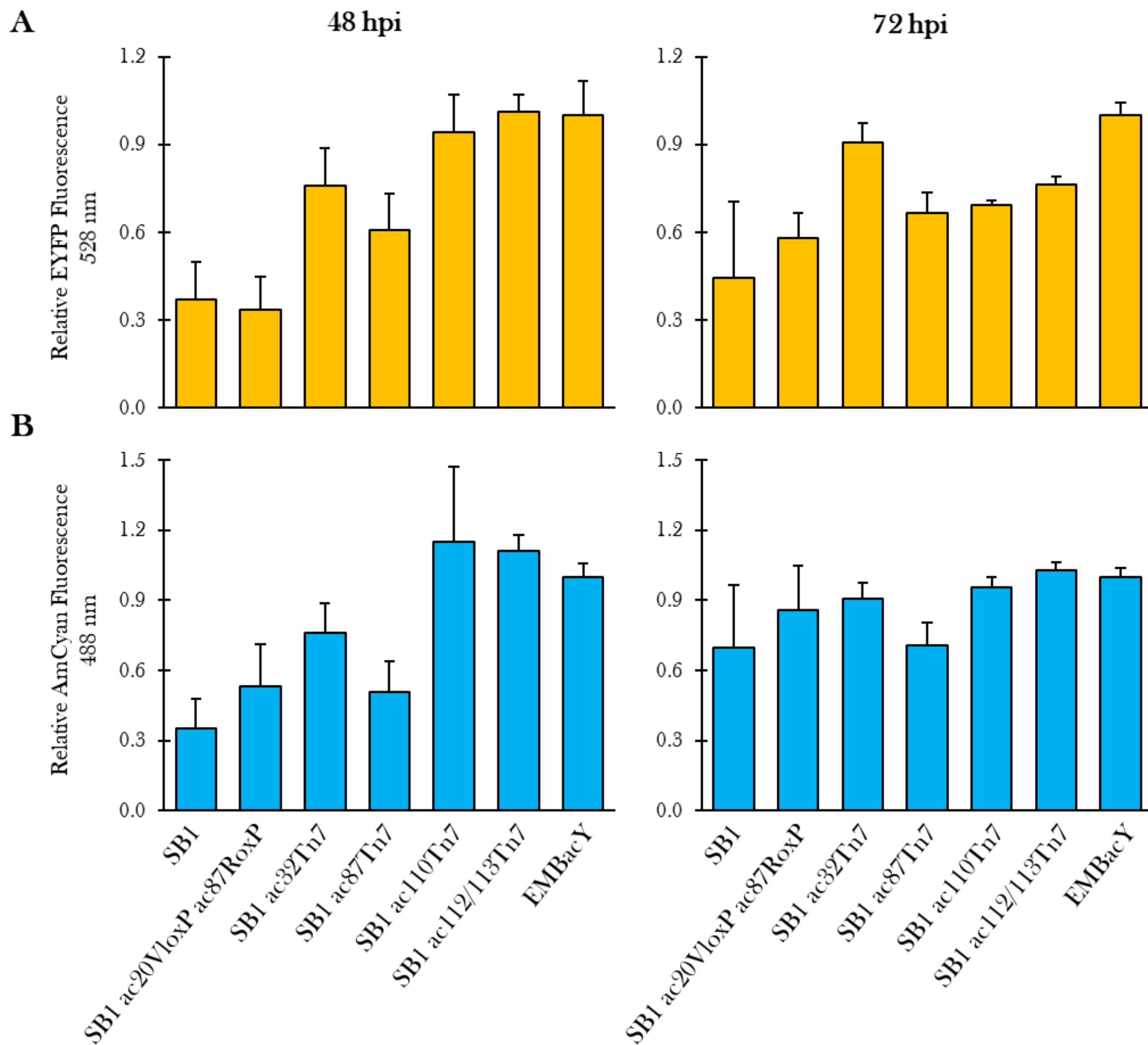


Figure 5.22 – Reassessment of the different SynBac1 genomes.

EYFP **(A)** and AmCyan **(B)** production from 500,000 cells taken at 48 and 72 hpi during V1 amplification for SynBac1 (SB1), SynBac1 ac20VloxP ac87RoxP and SynBac1 ac32, ac87, ac110, ac112/113 Tn7 ac20VloxP genomes. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3).

5.4 Discussion

In this chapter, the generation and validation of SynBac1 baculovirus variants containing differently located Tn7 transposition sites and inserted VloxP/RoxP sites was detailed. The five here generated genomes and the original SynBac1 genome were utilised for the re-insertion of a EYFP cassette, driven by the same promoter that was also utilised in the well established EMBacY genome. The six SynBac1 genomes were examined in terms of EYFP and AmCyan expression, benchmarked to the EMBacY virus and could proceed to the final stages of investigating potential gains in genome stability.

5.4.1 Generation of SynBac1 Variants

The methods to relocate the Tn7 transposition site and insert the site-specific recognition sites, VloxP and RoxP, were efficient and seamless. The Red/ET homologous recombination, previously used throughout this thesis, encountered no issues in terms of efficiency of reactions and the number of colonies acquired to screen. As the homology arms for the target genes were already used in the single gene deletion screen in Chapter 4, the design of the engineering was simple and guaranteed to succeed.

The use of the site-specific recombination systems, Dre/RoxP and VCre/VloxP, in the removal of the ampicillin marker posed no major issues, apart from the number of colonies that needed to be screened if multiple clones were required for testing. The work carried out in this chapter confirms the functionality of these systems (explored in Chapter 3) as tools for genome engineering, especially for excision of antibiotic markers, thus enabling the genome editing to be near scar less.

5.4.2 Tn7 and VloxP/RoxP Localisation

The work carried out in this chapter compellingly validated the previously performed single gene deletion screen in Chapter 4. The genes chosen for removal and insertion of the Tn7 site or the VloxP/RoxP sites were rightly selected, as the measured fluorescence intensities in the generated genomes were all comparable to the control genomes. Due to the single gene deletion screen performed in TB media, because of the presence of the blasticidin cassette, the quality of the bacmid and virus preparations were slightly different. A question about the reliability of the results was emerging, however the work completed in this chapter proved the chosen genes; ac20, ac32, ac87, ac110 and ac112/113; all could be deleted from the baculovirus genomes without major repercussions. The SynBac1 genome with the VloxP and RoxP at ac20 and ac87,

respectively, showed comparable levels of EYFP and AmCyan expression, (Figure 5.5 and 5.22). All the SynBac1 genomes with the Tn7 transposition site relocated to four sites were similarly functional even with removing ac20 and inserting the VloxP site on top of the Tn7 modification (Figure 5.15), thus deleting two genes in total.

The only Tn7 location that did not result in a completely unaffected virus was ac45, which the work to generate the virus with the relocated Tn7 site was performed by a colleague (H. Crocker). The virus was however tested alongside the four different Tn7 viruses generated in this chapter. The AmCyan fluorescence recorded was only half of the amount compared to the SynBac1 control. The ac45 single gene deletion screening in Chapter 4 showed the virus performed better with values around 90% of the control. The fluorescence was however lower than the control especially when comparing the other five single gene deletions (ac20, ac32, ac87, ac110 and ac112/113), which displayed higher fluorescence than the control. Out of the selected genes for the Tn7 relocation, ac45 did show the lowest potential and therefore did generate a virus with lower capacity.

5.4.3 EYFP Insertion - an Optimised Approach

Although the re-insertion of the EYFP took place as one of the last modifications, having the EYFP in place prior to the site-specific recognition sites addition and Tn7 site relocations would have resulted in less work needed to insert the EYFP (two genomes, SynBac1 original Tn7 and SynBac1 no Tn7, instead of the six generated versions). However, due to ambiguities encountered by another lab member while trying to insert the polh driven EYFP and obtain clones producing similar EYFP levels as the control EMBacY genome, the modifications to insert the site-specific recognition sites and relocate the Tn7 sites went ahead first.

While the work itself to insert the EYFP via the Cre/LoxP fusion would have been on a similar timeline (multiple reactions set up in parallel), the optimisation of the procedure while having six different baculovirus genomes involved more hands-on time. Particularly, optimising the amount of total bacmid DNA used meant numerous reactions had to be set up before finding a range where the reaction was successful, and colonies were present. Firstly, putting a concentration of around 200 ng, an amount normally used for plasmid fusions, did not generate any colonies. It was thought the amount was too low for a fusion reaction involving the bacmid, as the size of the bacmid is vastly larger. Due to the size, less copies of the bacmid would be present in the reaction, thus less chance of the LoxP sites locating together for the fusion reaction. On the other hand, reactions involving over 8 μg resulted in less colonies present on plates than reactions set up with

the same bacmid preparation but lesser bacmid concentration. This was most likely a result of saturating the enzyme, thus not performing as efficiently, however using more enzyme would not be cost effective. As colonies were present on the reactions using less bacmid (3-7 μg), the need to use more enzyme was unnecessary. Another point to highlight is the preparation of the bacmid DNA and resulting DNA concentrations measured. Unlike plasmid preparations using standard purifications kits, bacmid DNA is prepared by isopropanol precipitation where there is a higher probability of *E. coli* genomic DNA contamination. The values recorded for the bacmid concentrations may in fact be lower, which may explain the generally high bacmid DNA used compared to amounts used in plasmid-plasmid fusions or other cloning ligations. Lastly, screening clones for two genomes instead of six would have allowed the possibility to examine more clones, thus making sure the best possible clone was selected in terms of the EYFP fluorescence output.

5.4.4 Variations between Fluorescence Readout

The fluorescence data for the generated baculovirus genomes presented throughout this chapter displayed variations within the same genome. For example, both SynBac1 with Tn7 at ac110 and ac112/113 showed EYFP fluorescence of around 60% at 72 hpi (Figure 5.20 and 5.21) compared to the control, EMBacY, in the first screen of EYFP insertion clones, but reached levels above 100% on the retest of the clones and already at 48 hpi (Figure 5.22). The AmCyan levels of these two variants were fluctuating less between 70-100% in the re-evaluation of clones after Tn7 and VloxP insertion (Figure 5.10 and 5.15). Another example was SynBac1 with Tn7 at ac87, where the highest EYFP insertion clone reached only 20% EYFP (Figure 5.19) compared to EMBacY, however when re-examined the EYFP fluorescence increased to 70% at 72 hpi (Figure 5.22). On the other side original SynBac1 and SynBac1 with VloxP and RoxP inserted showed EYFP fluorescence at 70% and 100% compared to EMBacY in the first EYFP screen, but when retested the values decreased to only 40% and 60%, respectively. The fluorescence was standardised to the control genome, which can influence the final numbers if the control virus performed extremely well or poorly, however the raw fluorescence data for the control was always checked to make sure it fits within a range that was commonly acquired for the virus. Additionally, any variability between the triplicates were displayed as error bars, where overall the triplicates showed a low value for the standard error of mean (SEM). It remains to be seen whether these variations between the fluorescence intensities are due to the inherent variability of biological systems or are alternatively a result of fluctuations in the viral cycles of the modified SynBac1 variants.

5.4.5 Concluding Remarks and Future Directions

The generation and testing of SynBac1 viruses with relocated Tn7 transposition site and inserted VloxP/RoxP sites has been presented. The engineering of the SynBac1 genome was achieved using homologous recombineering and the modifications confirmed by colony PCR of the baculoviral genome. The viruses were subjected to testing and the fluorescence levels of AmCyan were measured compared to the control, SynBac1. The EYFP cassette under control of the polh promoter was inserted by Cre/LoxP fusion into the following six baculoviral genomes:

1. SynBac1 original Tn7
2. SynBac1 original Tn7 ac20VloxP ac87RoxP
3. SynBac1 ac32Tn7 ac20VloxP
4. SynBac1 ac87Tn7 ac20VloxP
5. SynBac1 ac110Tn7 ac20VloxP
6. SynBac1 ac112/113Tn7 ac20VloxP

The viruses with the inserted EYFP were assessed once again, by measuring the fluorescence levels of the expressed reporter proteins, EYFP and AmCyan, and compared to the routinely utilised EMBacY baculovirus as a benchmark. Overall, the results confirmed the genes were suitable candidates for deletion from the genome, as previously showed in the single gene deletion screen in Chapter 4, as all viruses performed similarly (subjective to a statistical analysis) in terms of fluorescence measured compared to the control virus. Out of the six virus variants tested, SynBac1 ac32Tn7 ac20VloxP displayed the least variations between the fluorescence values obtained, with both EYFP and AmCyan values between 90-120% compared to the controls throughout the various experiments in this chapter. Another two promising genomes were SynBac1 with the Tn7 site at ac110 and ac112/113 with EYFP and AmCyan levels reaching the same values as EMBacY at 48 hpi already.

As a result of this work, the six baculoviral genomes generated in this chapter will be subject to further testing to discover if the SynBac1 variants show improved virus stability compared to the precursor genome, EMBacY, the ultimate goal of this thesis. The stability of the viruses will be determined by measuring the percentage of mCherry and mTagBFP cells (expressed from Tn7 site) within the percentage of cells that are EYFP positive (viral backbone). The lower the number of mCherry and mTagBFP positive cells within an EYFP positive cell pool will signify lower stability due to the loss/inactivation of the Tn7 region. The stability of the viruses will be

measured over a course of five viral passages to be able to show any increasing instability of the genome.

Additionally, the most stable baculoviral genomes will be subject to a heterologous protein expression benchmarking study with an array of single and multi-peptide complex proteins. This will enrich the information about the properties and the suitability of the new viruses towards a better system. Lastly, the site-specific recognition sites inserted in the genomes, VloxP and RoxP, will be validated by fusing the colour donor plasmids generated in Chapter 3 and visualising the fluorescence by microscope analysis. Being able to insert multiple modalities into the baculoviral genome and different loci will add further functionality, modularity and flexibility to the next generation baculovirus genome aspired to in this thesis.

Chapter 6: SynBac Variants Validation

6.1 Introduction

The work carried out previously identified a set of potential genes for deletion in the framework of SynBac1. The information was used for the relocation of the attTn7 site to four different locations, as well as the insertion of the RoxP and VloxP sites to augment the utility of the baculoviral genome. The EYFP reporter gene was successfully re-inserted into these mutant SynBac genomes under the control of the polh promoter, the same as in the established and widely used benchmark EMBacY genome. The SynBac variants were examined in terms of EYFP and AmCyan fluorescence and results indicated similar values for most of the variants as compared to the EMBacY. The six generated SynBac genomes can therefore be assessed to establish how far they present a step-change over and above the state-of-the-art. Firstly, the new RoxP and VloxP sites will be validated by fusion of the previously established colour donor plasmids outfitted with the respective site-specific sites. Critically, the SynBac variants and EMBacY will be subjected to a serial passaging experiment to probe the stability of the genomes by measuring the loss of fluorescence signal across multiple virus amplifications. Moreover, the propensity for heterologous target protein production capabilities of the SynBac variants will be contrasted to EMBacY as a benchmark by means of expressing an array of proteins including a protein complex.

6.1.1 RoxP and VloxP Site Validation

The inserted site-specific recognition sites, RoxP and VloxP, in the SynBac variants need to be tested by fusing, mediated by the action of specific recombinase, a plasmid comprising the same site into the baculoviral genome. The colour donor plasmids generated earlier, (Section 3.3.1.3.2) containing the RoxP and VloxP sites are pUCDM_RoxP_AmCyan and pUPDM_VloxP_EGFP, respectively. The fluorescence emission spectrum of EGFP partly overlaps with the fluorescence emission spectrum of EYFP (in bacmid backbone), thus this pair is not ideal when deployed concomitantly due to proximity of the emission wavelengths that can be measured. Therefore, EGFP in the pUPDM_VloxP_EGFP donor plasmid was exchanged for mCherry. The spectral separation between excitation and emission wavelengths of EYFP, mCherry and AmCyan render these a suitable choice for concomitant detection of the fluorescent reporters during microscopy analysis.

Briefly, the two donor plasmids will be inserted into the SynBac genomes containing the VloxP and RoxP sites by incubating the plasmid and bacmid preparation in the presence of the corresponding recombinase, VCre and Dre. SynBac genomes with single inserted colour

plasmid, pUADM_RoxP_AmCyan in ac87 and pUPDM_VloxP_mCherry in ac20, as well as a double insertion will be generated. After screening for successful fusion by PCR, bacmid preparations would be transfected and virus amplified in insect cells. Cells from the V1 amplification will be analysed under the confocal microscope by checking for AmCyan and mCherry fluorescence, which would be expressed from the inserted donor plasmids in the RoxP and VloxP sites.

6.1.2 Genome Stability

The expression from large heterologous DNAs, encoding for instance, for multi-subunit protein complexes and optional accessory factors such as chaperones or modifying enzymes (kinases, phosphatases) is starkly sensitive to detrimental events. These events eliminate pieces of DNA from the baculoviral genome during replication, as occurrence that can only be partly addressed by optimising amplification protocols.[160] The overall effect of this genomic instability is a substantial decrease of productive yield during serial passaging, in particular for fermenter-scale applications where large volumes of baculovirus are required. These observations support the serial passaging effect where the viral genome undergoes deletions during successive virus passages.[218,219] To enable the study of the genome stability of the new viral specimens generated here, an appropriate multi-subunit reporter system was required. The system needed to have a clear measurable read-out, such as fluorescence, which can be followed over the course of serial virus amplifications. The chosen reporter system included multiple fluorescent components to enable analysis through cytometry, fluorimetry and microscopy.

The reporter system was previously designed and generated by another lab member (H. Crocker), which included the genes for mCherry, mTagBFP with a nuclear localisation sequence and a SNAP-tag attached to the baculoviral membrane protein, GP64. The individual components were combined together by Cre/LoxP and ligation of acceptor and donor plasmids resulting in the following construct: pKL_NLSmTagBFP x pUCDM_SNAPTag_GP64 x pSPL_mCherry. Due to the SNAP-tag involving additional steps for its detection (binding of fluorescent or chemiluminescence probes) and the protocol not optimised previously, the SNAP-tag was not utilised in this study and the plasmid utilised is referred to as Dual Colour. The remaining two fluorescent proteins and the EYFP reporter located in the baculoviral genome are however adequate to measure the virus stability.

The Dual Colour plasmid was transformed into the Tn7 site of the different SynBac variants and the EMBacY virus. To acquire interpretable information regarding genome stability, the viruses

were passaged for 5 generations and the fluorescence tracked by flow cytometry, microscopy and fluorimetry of the samples collected at each virus amplification. All three of the fluorescent proteins are under the control of the same promoter, polh, and thus should be expressed at comparable levels. However, as previously described in Section 1.3.3, the Tn7 site is one of the main regions lost during serial passaging of the virus. The percentage of cells displaying the two fluorophores found in the Tn7 site, mCherry and mTagBFP, (Figure 6.1) in comparison to the percentage of EYFP (found elsewhere in the genome) positive cells will facilitate the analysis of the genome stability. Thus, upon successive virus amplification, the decrease of mCherry and mTagBFP within EYFP positive cells would signify the loss/inactivation of the Tn7 region, and the amount can be compared among the different SynBac variants and EMBacY to determine the most stable genome.

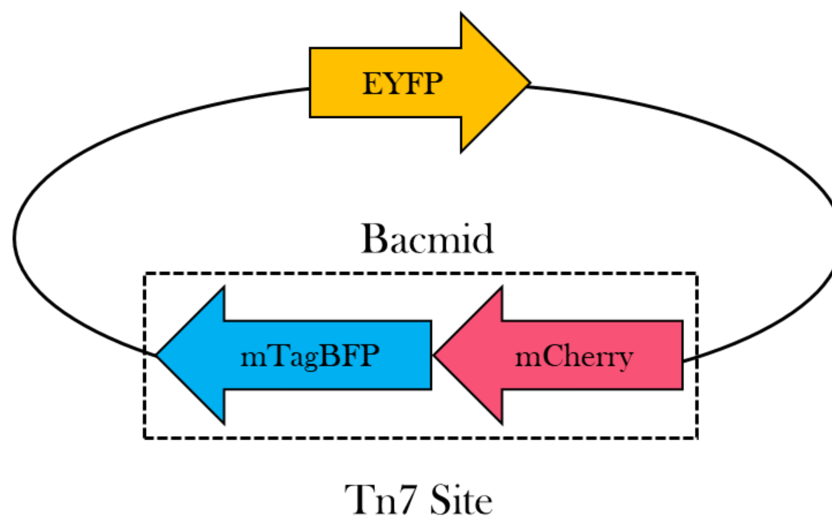


Figure 6.1 - Fluorophore location in genome stability analysis.

A schematic illustration of the bacmid containing the EYFP reporter genome and the two fluorescent proteins found in the Tn7 transposition site from the transformed Dual Colour plasmid. All three genes are under the control of the polh promoter.

6.1.3 Protein Expression

The ultimate validation procedure included assessing the most stable SynBac genomes in terms of their protein production ability. The selected SynBac variants, based on the genome stability results, are to be tested alongside EMBacY, the current best protein expression system. A selection of single proteins varying in molecular size (56 to 110 kDa) and a multi-protein transcription factor complex, composed of the TATA-box binding protein (TBP) associated factors TAF5, TAF6 and TAF9.[324] This complex, henceforth denominated as TAF569 in this

thesis, is a subcomplex of human general transcription factor TFIID, a cornerstone of gene expression, studied in the Berger laboratory. The single proteins (Keap1, Lck, PPAR and hPI3Kg) were commonly used beforehand as test proteins at AstraZeneca whenever new protocols needed optimising and exhibit easily observed protein expression already in cell lysate without the need for purification. Selecting targets that are detected easily and do not need to be purified is highly desirable to reduce experimental inconsistencies when comparing multiple viruses. The TAF569 complex was selected based on the differing molecular sizes of the individual proteins (TAF5 - 92 kDa, TAF6 - 73 kDa and TAF9 - 30 kDa). All the protein expression validation samples will be collected from the same number of Sf21 insect cells during V1 amplification and the fluorescence of the EYFP reporter compared. Some targets will be amplified at the V2 stage to check if the protein production differs across the two virus generations. The samples will be examined by resolving proteins on SDS-PAGE gels, and where necessary carrying out Western blot analyses.

6.2 Chapter Aims

The aim of this chapter is to validate the newly generated SynBac variants in terms of functionality of the site-specific recognition sites, VloxP and RoxP, stability of the genomes and heterologous target protein expression capability. This will be explored through the following approaches:

1. Fuse colour donors into the VloxP and RoxP sites located in the bacmid and analyse fluorescence by microscopy.
2. Determine stability of the SynBac genomes compared to EMBacY across serial passages, by calculating number of mCherry and mTagBFP cells within a EYFP positive population.
3. Analyse the protein production of SynBac variants compared to EMBacY through expression of multiple single protein candidates and a multi-subunit protein.

6.3 Results

6.3.1 RoxP and VloxP Site Validation

The first validation experiment of the SynBac variants included checking the functionality of the earlier inserted RoxP and VloxP sites, near the Tn7 site and at ac20, respectively. The colour donors generated in Chapter 3 were utilised for the following work.

6.3.1.1 Redesign of VloxP Donor Plasmid

The donors containing the RoxP and VloxP sites are pUCDM_RoxP_AmCyan and pUPDM_VloxP_EGFP, respectively, however the EGFP was exchanged for mCherry in pUPDM_VloxP to enable better separation between the fluorescent proteins used. Briefly, the mCherry was PCR amplified from the pUZDM_Vox_mCherry plasmid and gel extracted. The mCherry fragment and the pUPDM_VloxP plasmid were digested with BamHI and XbaI, gel extracted and cleaned up. A ligation reaction of equimolar concentrations of backbone and insert was set up, incubated at room temperature for 2 hours, cleaned up, electroporated into appropriate cells, recovered and plated on puromycin plates. A correct plasmid was acquired by isolating the DNA, checking by restriction digestion and sequencing. The newly generated colour donor plasmid, pUPDM_VloxP_mCherry was fused as per methods section to the corresponding acceptor, pACEBac1_VloxP, and recovered cells plated on puromycin and gentamicin plates. The clones were miniprepped and checked by restriction digestion for correct fusion product. As previously illustrated in Chapter 3, (Figure 3.8 (B)), the donors (~3.6 kb), acceptors (~2.9 kb) and fusions (~6.5 kb) for the RoxP and VloxP sites were cut once to illustrate the size of the linear products, (Figure 6.2 (B)), with a schematic representation of the colour donors, pUCDM_RoxP_AmCyan and pUPDM_VloxP_mCherry (Figure 6.2 (A)).

Following the generation of the new donor and fusion plasmid, the ability to express the fluorescent protein needed to be tested in EMBacY again. The pUPDM_VloxP_mCherry fusion, as well as the previously checked pUCDM_RoxP_AmCyan were transformed into EMBacY containing cells by Tn7 transposition and triplicates of the two different fluorescent expressing fusions were transfected into Sf21 cells as per methods section. Samples were taken daily during V1 amplification of the virus to track the EYFP reporter and fluorescent protein expression. The fluorescence values of the samples taken 3 days post infection of V1 were plotted to show their relative fluorescence units, (Figure 6.2 (C)), EYFP-yellow, AmCyan-light blue and mCherry-pink. Images of the Sf21 cells were taken on the DMI6000 Leica confocal microscope

48 hpi during V1 amplification of the colour fusion plasmids to illustrate expression of the fluorescent markers, (Figure 6.2 (C)). The newly generated colour fusion plasmid, pUPDM_VloxP_mCherry demonstrated successful fluorescent marker expression in the EMBacY system.

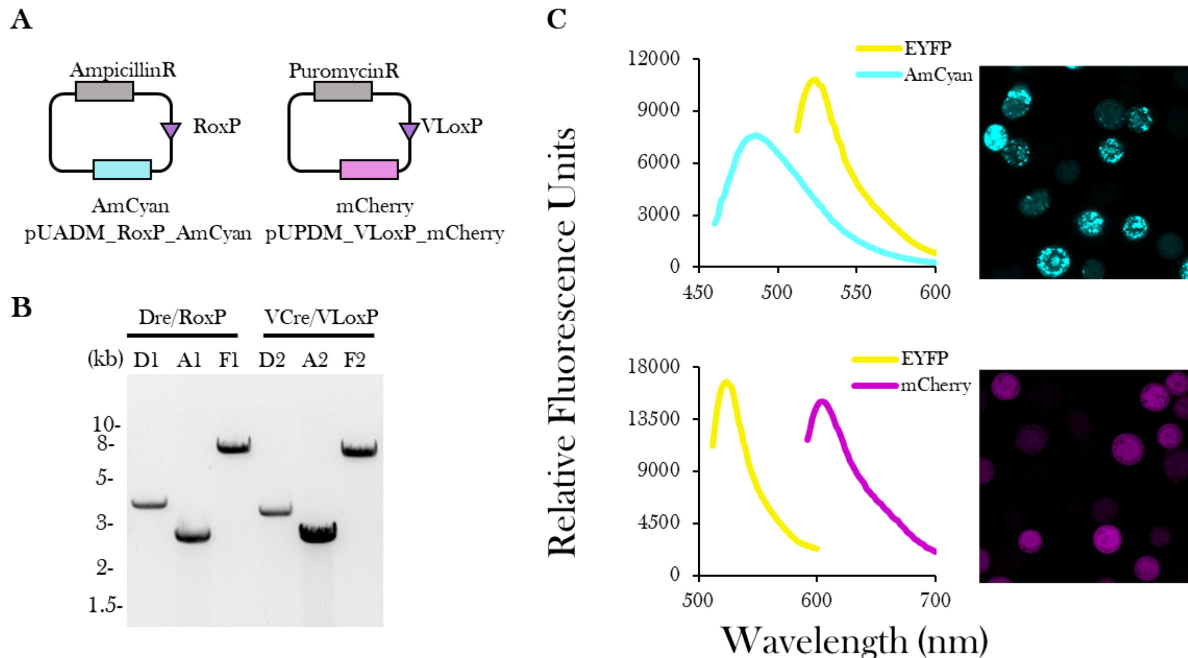


Figure 6.2 - Fluorescent donor plasmids and protein expression in EMBacY transfected Sf21 cells.

(A) Schematic diagram showing the generated fluorescent donor plasmids, containing a different resistance marker (ampicillin and puromycin), site-specific recombination site (RoxP and VloxP) and fluorescent protein (AmCyan and mCherry). **(B)** 1% agarose gel of the fusions of donor and acceptor plasmids *in vitro* via different site-specific recombinases and their recognition sites. The donor (D) and fusion (F) plasmids were analysed by NheI restriction, and acceptor (A) plasmids by BamHI restriction. **(C)** EMBacY containing cells were transformed with two colour plasmid fusions, transfected and V1 amplified according to the methods section. EYFP (yellow plot), AmCyan (light blue) and mCherry (pink) values were measured from 1 million cell samples collected 3 days post infection of V1, according to their excitation and emission spectra, (Table 2.4, Materials and Methods). At 48 hpi 70,000 cells per well in 300 µl total volume were transferred to 8 chamber Nunc™ Lab-Tek™ II chamber slides (Thermo Fisher) and left overnight at 27 °C. Images of the Sf21 cells were taken at 72 hpi on the DMI6000 Leica confocal microscope. 63X magnification, oil immersion objective used. pUADM_RoxP_AmCyan*pACEBac1_RoxP (top) and pUPDM_VLoxP_mCherry*pACEBac1_VloxP (bottom).

6.3.1.2 Integration Site Validation

The RoxP and VloxP sites inserted into the new SynBac variants (SynBac1 original Tn7 ac20VloxP ac87RoxP, SynBac1 ac32Tn7 ac20VloxP, SynBac1 ac87Tn7 ac20VloxP, SynBac1 ac110Tn7 ac20VloxP, SynBac1 ac112/113Tn7 ac20VloxP) could now be tested utilising the colour donors. The two donor plasmids, pUCDM_RoxP_AmCyan and pUPDM_VloxP_mCherry, were used in fusion reactions by incubating baculoviral DNA with the plasmid, Dre or VCre recombinase, respectively, and the 3.1 buffer as per methods section. Previous work for insertion of the EYFP reporter into the LoxP site provided information about the optimal amount of baculoviral DNA (3 to 6 μg) to acquire enough colonies. However, due to the use of self-expressed and purified Dre and VCre recombinases, the reactions needed to be repeated several times to obtain a colony displaying a successful insertion of the donor plasmid. For illustration purposes, one of the SynBac variants (ac32Tn7 ac20VloxP) was selected to show the correct insertion of the donors and fluorescence analysis. The insertion of the plasmids was confirmed by colony PCR, (Figure 6.3 (A)) with pUPDM_VloxP_mCherry at ac20 (4433 bp), pUADM_RoxP_AmCyan insertion at ac32RoxP (5990 bp) and pUPDM_VloxP_mCherry insertion at ac20VloxP (4433 bp) on top of the clone already containing the ac32RoxP insertion.

The SynBac1 ac32Tn7 ac20VloxP variant containing the inserted donors, pUCDM_RoxP_AmCyan and pUPDM_VloxP_mCherry, as well as the double insertion were prepped and transfected in triplicates into Sf21 cells as per methods section. As controls for the AmCyan and mCherry detection, the fluorescent donors were transformed into the Tn7 site of the SynBac ac32Tn7 ac20VloxP genome (AmCyan in Tn7 site, mCherry in Tn7 site), as well as an empty virus (blue colony) for EYFP detection and V1 amplified alongside. At 48 hpi 70,000 cells per well in 300 μl total volume were transferred to 8 chamber Nunc™ Lab-Tek™ II chamber slides (Thermo Fisher) and left overnight at 27 °C. Images of the Sf21 cells were taken at 72 hpi on the DMI6000 Leica confocal microscope, (Figure 6.3 (B)), where the fluorescent proteins from the donor plasmids were visualised. The variants containing the pUADM_RoxP_AmCyan and pUPDM_VloxP_mCherry fusions displays AmCyan and mCherry fluorescence, respectively. The double insertion variant shows the presence of both AmCyan and mCherry fluorescence. The controls of the two donors transformed into the Tn7 site also present the correct fluorescence. The empty virus displays only EYFP fluorescence, which can be visualised in all the samples.

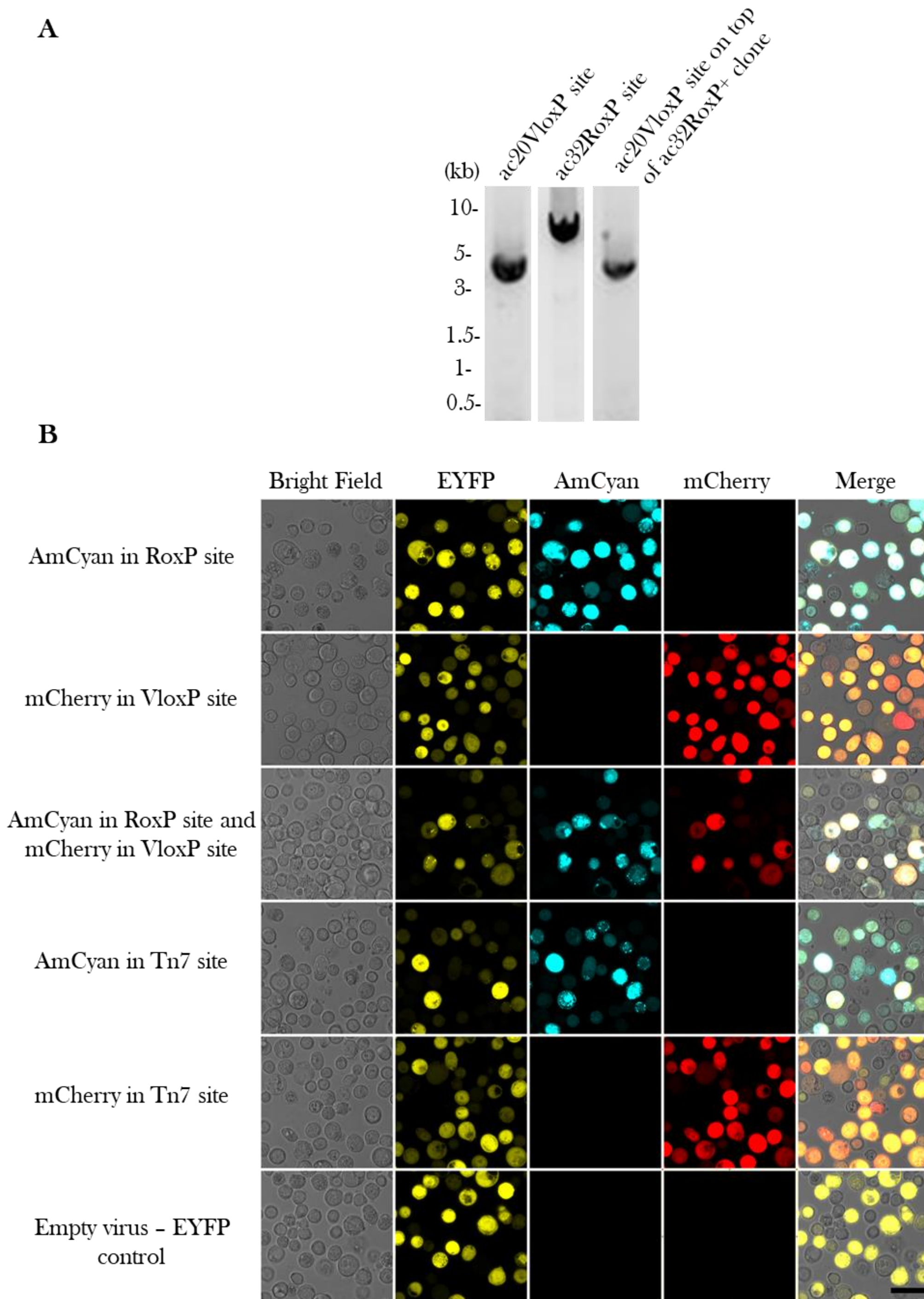


Figure 6.3 – VloxP and RoxP site validation.

(A) 1% agarose gel showing the PCR validation of the ac20VloxP and ac32RoxP site in the baculoviral genome. The PCR products show the expected fragment sizes after (left) pUPDM_VloxP_mCherry insertion at ac20VloxP (4433 bp), (middle) pUADM_RoxP_AmCyan insertion at ac32RoxP (5990 bp) and (right) pUPDM_VloxP_mCherry insertion at ac20VloxP (4433 bp) on top of the clone already containing the ac32RoxP insertion. (10/11CheckFor and 10CheckRev (ac20), 14CheckFor and

ac32Tn7CheckRev (ac32) primers) **(B)** V1 amplifications of the SynBac ac32Tn7 ac20VloxP genome with the individual (AmCyan in RoxP site, mCherry in RoxP site) and double fluorescent (AmCyan in RoxP site and mCherry in RoxP site) donor insertions. As controls for the AmCyan and mCherry detection, the fluorescent donors were transformed into the Tn7 site of the SynBac ac32Tn7 ac20VloxP genome (AmCyan in Tn7 site, mCherry in Tn7 site), as well as an empty virus (blue colony) for EYFP detection and V1 amplified alongside. At 48 hpi 70,000 cells per well in 300 µl total volume were transferred to 8 chamber Nunc™ Lab-Tek™ II chamber slides (Thermo Fisher) and left overnight at 27 °C. Images of the Sf21 cells were taken at 72 hpi on the DMI6000 Leica confocal microscope. 63X magnification, oil immersion objective used, scale bar = 50 µm.

6.3.2 Genome Stability

The six SynBac variants generated in the previous chapter were subject to genome stability analysis and benchmarked against the established EMBacY genome. The Dual colour plasmid was transformed into all variants and EMBacY, recombinant bacmids prepared and transfected into Sf21 cells in triplicates. The viruses were passaged from V0 to V5 and microscope, fluorimetry and FACS data collected at each virus passage. The percentage of cells displaying mCherry and mTagBFP expressed from the Dual colour plasmid in the Tn7 site was calculated within the population of EYFP positive cells, the fluorescent reporter found within the baculoviral genome. Lower percentage of mCherry and mTagBFP compared to the EYFP across the virus generations signified lower genome stability.

6.3.2.1 P10 Absence Confirmation

Firstly, the identity of the six different SynBac variants and EMBacY was confirmed by checking for presence of P10, a protein found within the EMBacY genome, but not SynBac, as it was deleted within the first Fragment 1 modification. Samples from V3 amplification were analysed by SDS-PAGE, (Figure 6.4), where the absence of P10 can be seen for the six SynBac variants and present in the EMBacY sample. This confirmation made sure that none of the SynBac viruses were cross contaminated with the EMBacY virus.

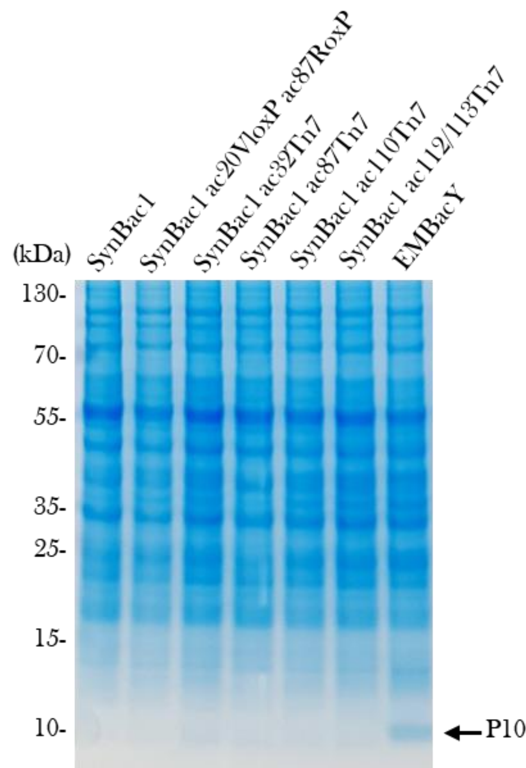
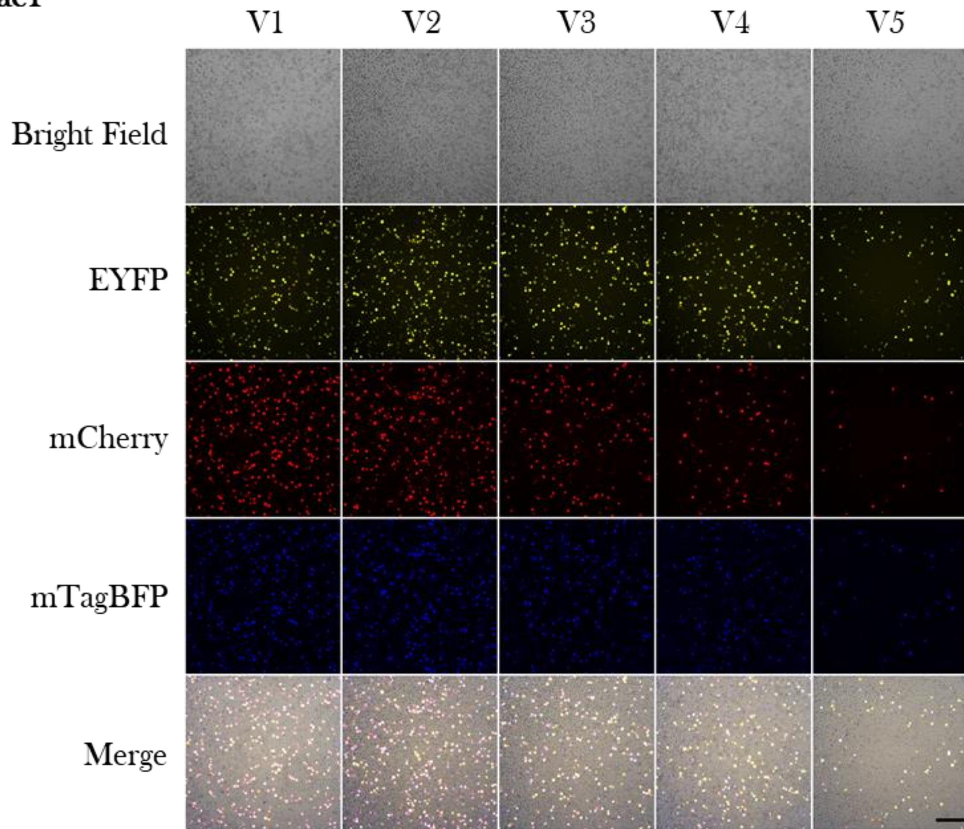


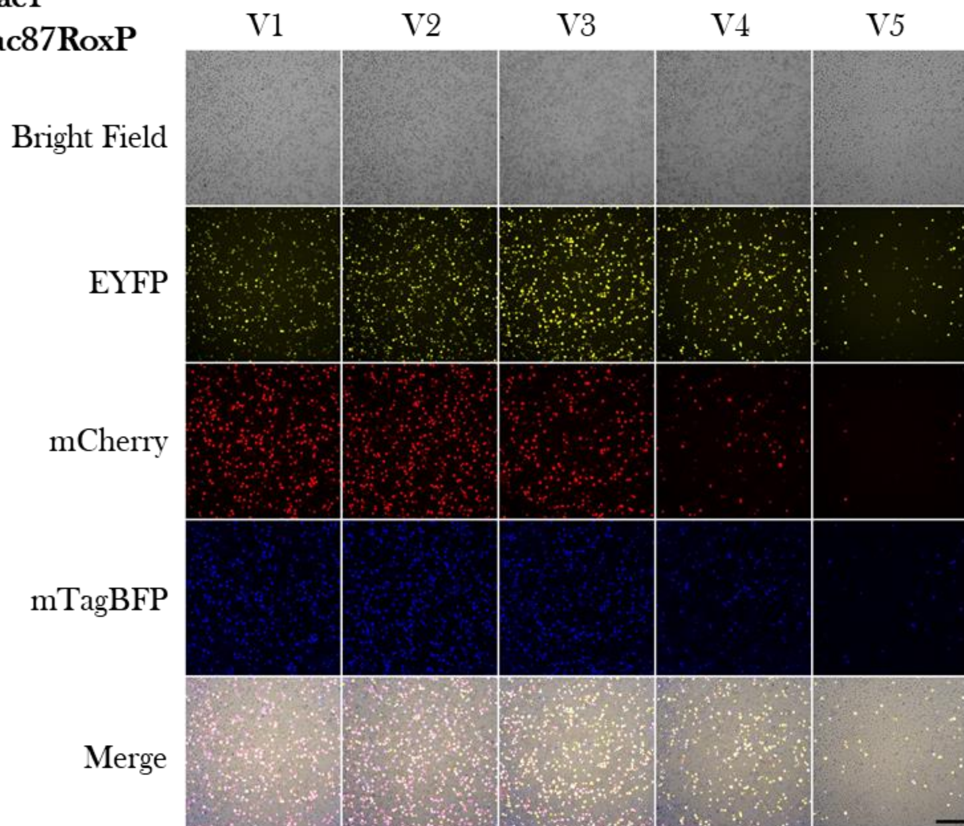
Figure 6.4 – P10 absence confirmation.

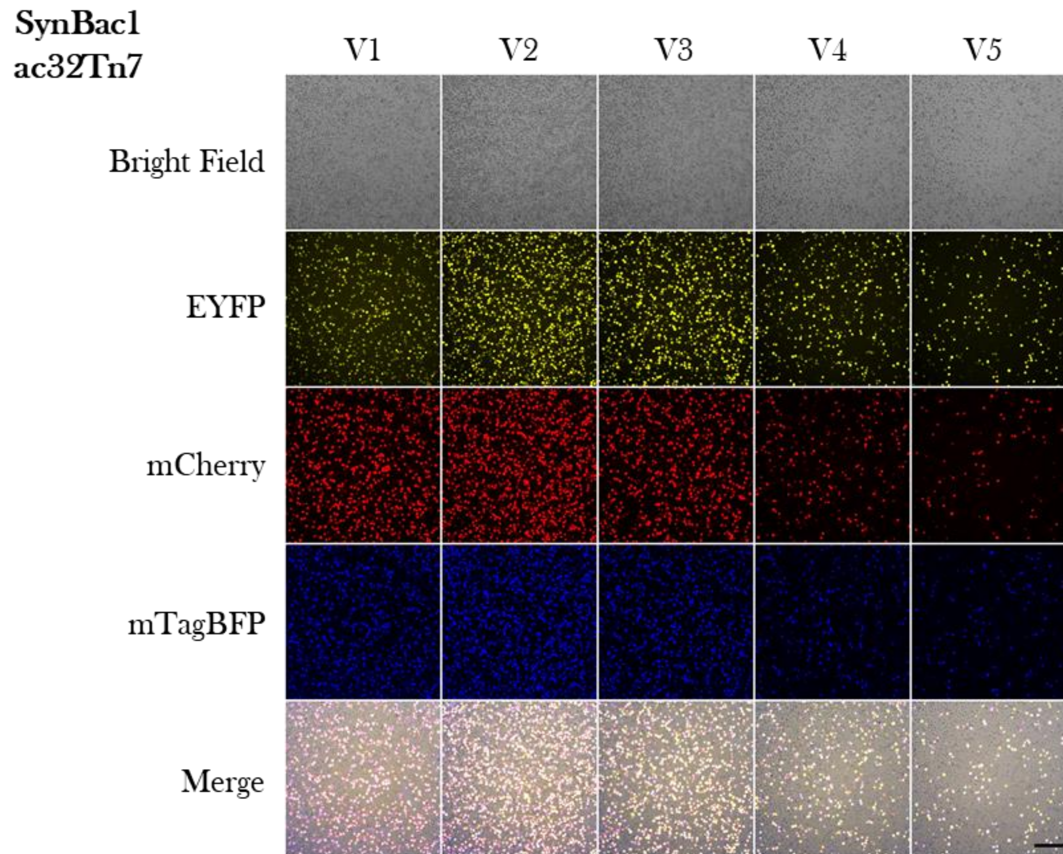
SDS-PAGE analysis of SynBac variants and EMBacY 500,000 cell samples collected at V3 amplification of the stability experiment. 13 µl samples were analysed on a NuPAGE™ 4-12 % Bis-Tris Protein Gel indicating the absence of P10 in the SynBac variants and the presence in the EMBacY sample (last lane).

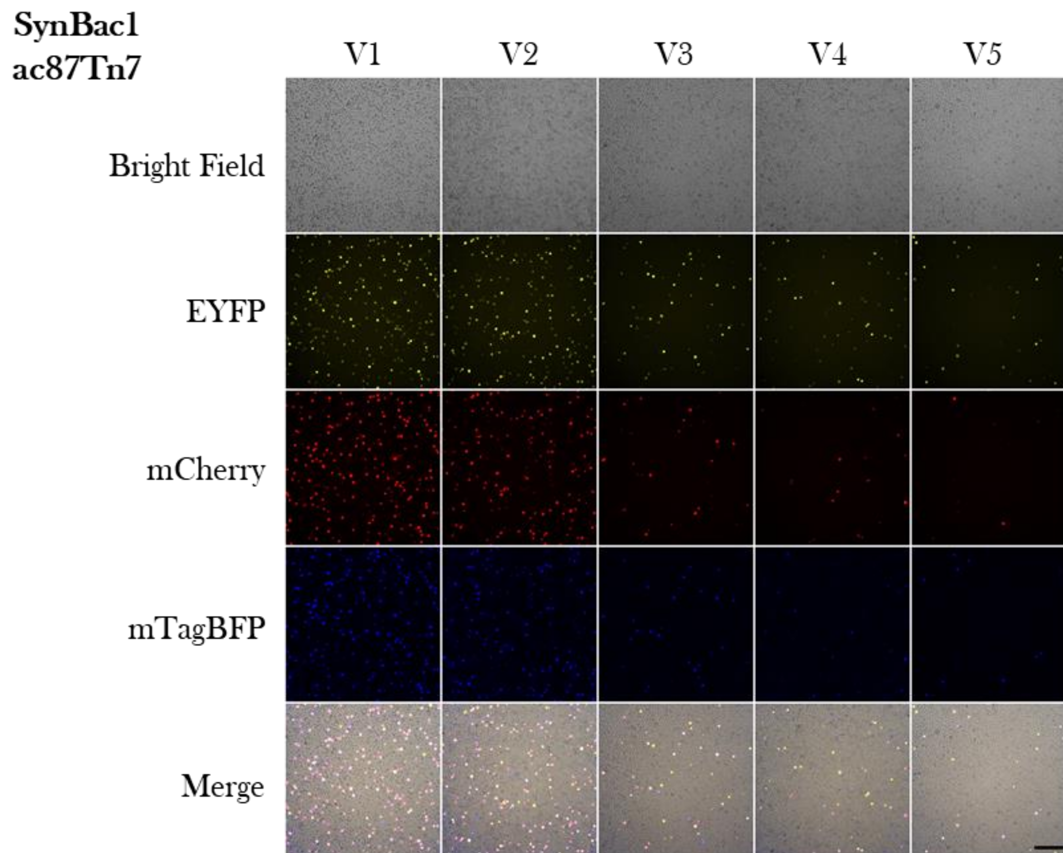
6.3.2.2 Microscopy Analysis during Serial Passaging

At 48 hpi for each virus amplification, approximately 1 million cell samples from one of the triplicates were transferred onto 6 well plates and incubated at 27 °C for at least an hour before microscope analysis. The images in Figure 6.5 display the EYFP, mCherry and mTagBFP fluorescence for six of the SynBac variants and EMBacY from V1 to V5. A general trend of highest fluorescence can be seen for V2 samples, with the fluorescence decreasing with each virus passage. Although the data is only qualitative, the EMBacY virus showed the strongest fluorescence, with only two of the SynBac variants, ac32Tn7 and ac112/113Tn7, displaying similar levels of fluorescence. The SynBac ac110Tn7 variant presented slightly lower fluorescence and the remaining three variants even less.

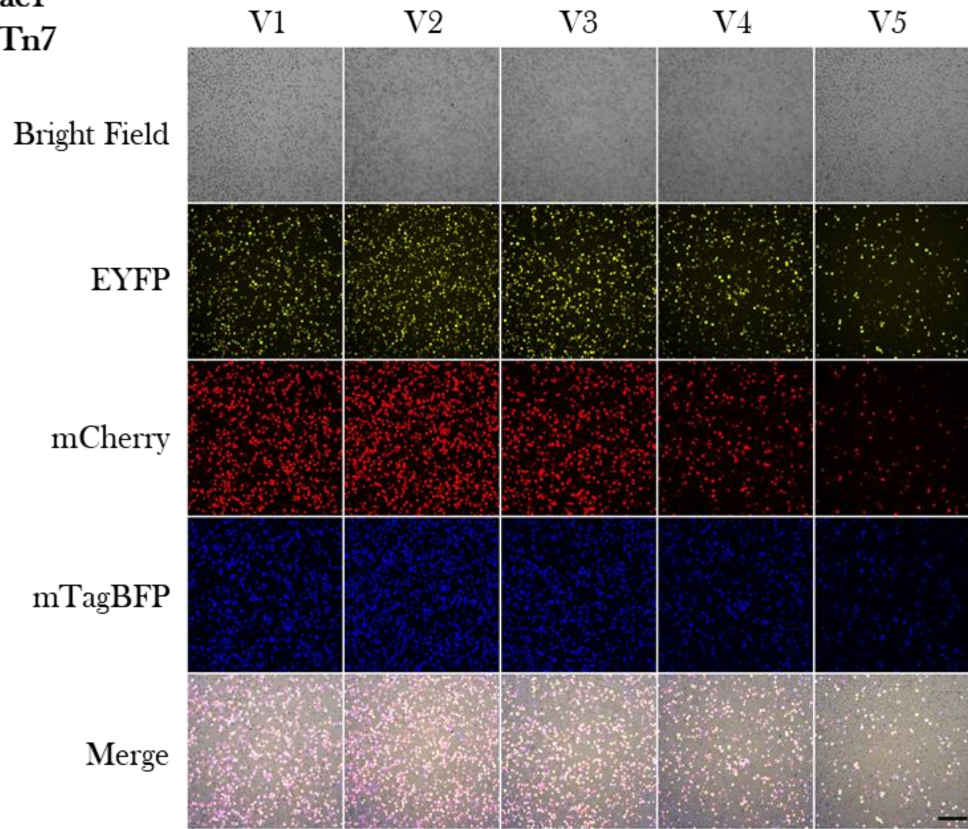

SynBac1
V1
V2
V3
V4
V5
Bright Field
EYFP
mCherry
mTagBFP
Merge


SynBac1
ac20VloxP ac87RoxP
V1
V2
V3
V4
V5
Bright Field
EYFP
mCherry
mTagBFP
Merge

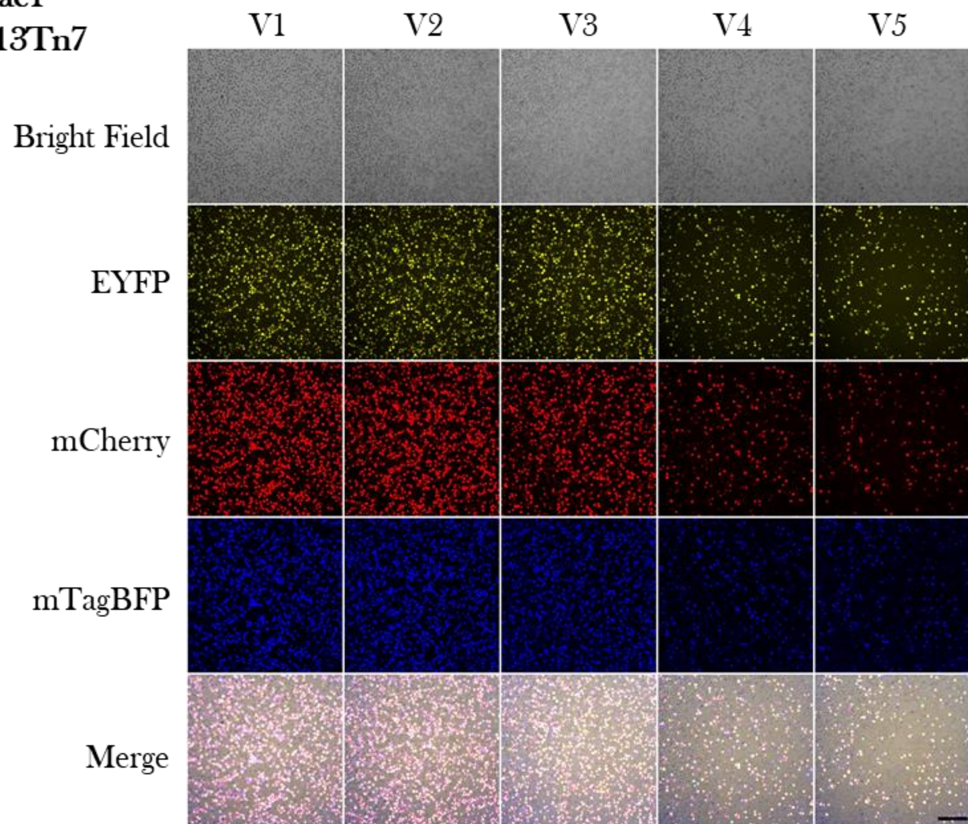

SynBac1
ac32Tn7
V1
V2
V3
V4
V5
Bright Field
EYFP
mCherry
mTagBFP
Merge


SynBac1
ac87Tn7
V1
V2
V3
V4
V5
Bright Field
EYFP
mCherry
mTagBFP
Merge

SynBac1 ac110Tn7



SynBac1 ac112/113Tn7



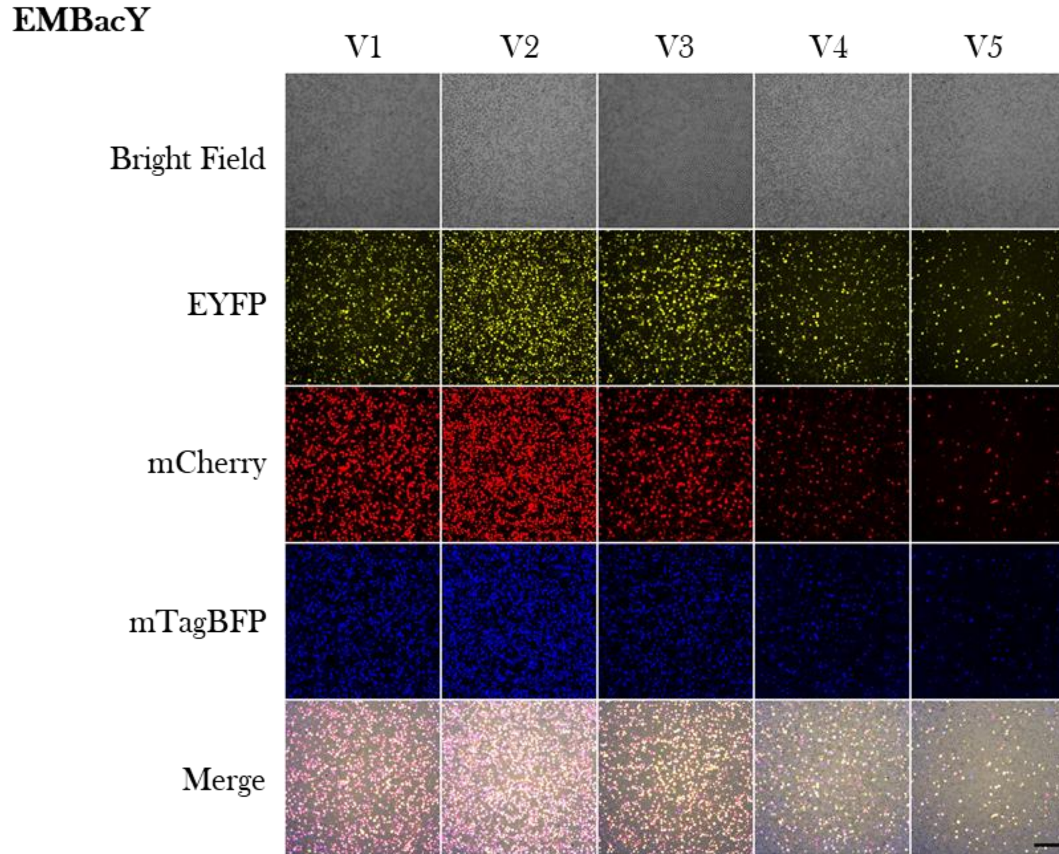


Figure 6.5 – Microscopy analysis during serial passaging.

Widefield microscopy images of SynBac1, SynBac1 ac20VloxP ac87RoxP, SynBac1 ac32Tn7, SynBac1 ac87Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY expressing mTagBFP and mCherry during V1 to V5 virus amplification (left to right). 1 million cells in 3 ml total volume per well were transferred onto 6 well plates (WVR) and left at 27 °C for 1-2 hours before imagining. Displayed from top to bottom are the bright field images, fluorescence channels of EYFP, mCherry and mTagBFP, and the merge of the four channels. Scale bar = 250 µm.

6.3.2.3 Fluorescence Quantification during Serial Passaging

Samples were collected at each virus generation for analysis by fluorimetry, (Figure 6.6), where (A) displays EYFP, (B) mCherry and (C) mTagBFP. The data indicates EMBacY reached maximum protein expression levels at V2, whereas most of the SynBac variants reached their maximum at V1. However, the SynBac1 ac32Tn7 variant showed maximum protein levels at V3. The EYFP fluorescence remains the highest for the EMBacY and SynBac1 ac32Tn7 variant across the five virus passages. Although the mCherry and mTagBFP fluorescence levels are highest for EMBacY, the SynBac1 ac110Tn7 and ac112/113Tn7 variants displayed very similar amounts of fluorescence, whereby V3 the levels are higher than EMBacY. Overall, the fluorescence yields decrease by V3, as seen in the microscope images in Figure 6.5.

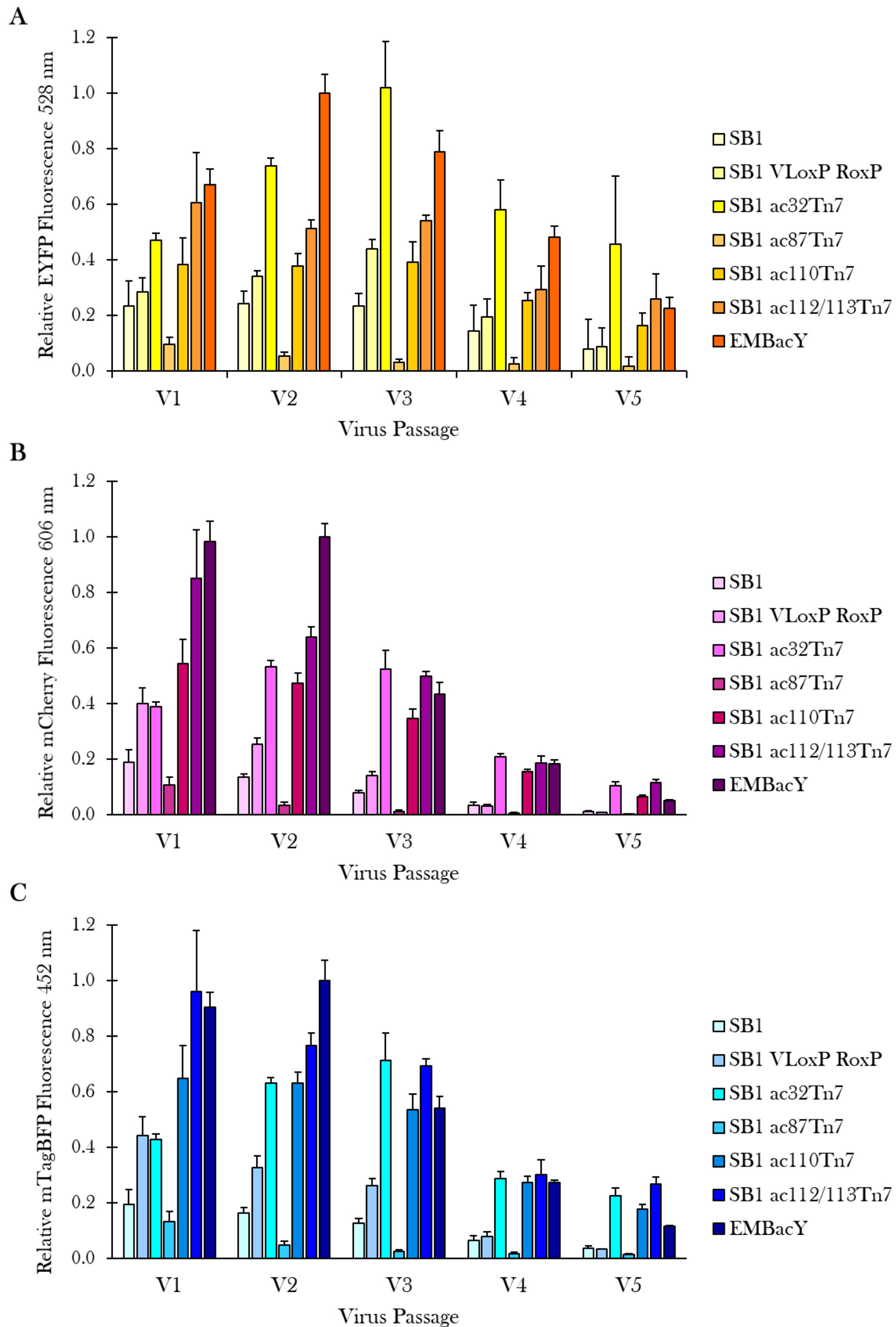


Figure 6.6 – Quantification of fluorescence during serial passaging.

EYFP **(A)**, mCherry **(B)** and mTagBFP **(C)** production from 500,000 cells taken at 48 hpi during V1 to V5 amplification for SynBac1 (SB1), SynBac1 ac20VloxP ac87RoxP, SynBac1 ac32Tn7, SynBac1 ac87Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. Values represent averages normalized to EMBacY = 1 at V2; error bars indicate SEM (n=3).

6.3.2.4 Cytometry Analysis of Genome Stability

The earlier microscope and fluorimetry data provided a good insight into the best virus genomes in terms of protein expression and their potential stability, however it was the cytometry data provided next that was the signifying data due to the measurable percentages of fluorescence cells within a cell population. An appropriate number of cells was taken at each virus passage and DRAQ7 added to distinguish live and dead cells. An example of the gating strategy applied to all of the flow cytometry data is shown in Figure 6.7, where (A) gates the cells from any other cell debris, (B) selects the single cells, (C) gates the live cells, (D) takes the EYFP population and (E) separates the fluorescence of the cells into only EYFP, EYFP and either of the two fluorescent proteins, mCherry and mTagBFP, and all three fluorescent proteins.

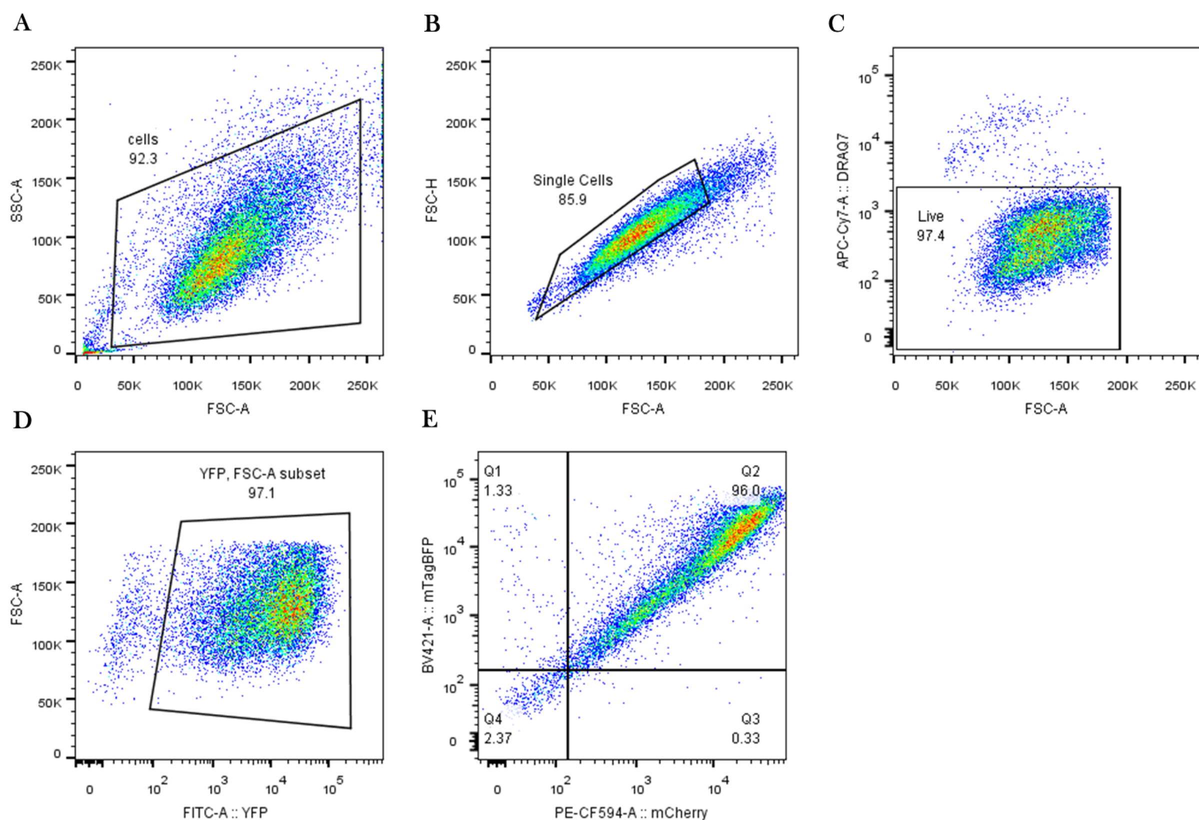


Figure 6.7 – Gating applied to flow cytometry data for fluorescence determination.

Representative FACS plots of Sf21 cells infected by EMBacY expressing mTagBFP and mCherry at 48 hpi from V1 amplification. The same gating was applied across all samples from V1 to V5 amplification. Gating strategy employed: (A) cell, (B) single cell, (C) live cell (DRAQ7 staining), (D) EYFP positive cell and (E) mTagBFP and mCherry positive cells within EYFP population, where Q1= EYFP and mTagBFP, Q2= EYFP, mTagBFP and mCherry, Q3= EYFP and mCherry and Q4= EYFP.

The data acquired from the cytometry analyses were plotted, (Figure 6.8), where the percentage of cells within the whole cell population displayed EYFP only (yellow), EYFP and mCherry (pink), EYFP and mTagBFP (blue) and all three fluorescent proteins (green). The results indicate EMBacY is the strongest virus in terms of percentage of fluorescent cells, with values around 90%, compared to the SynBac variants between 20-60% at V1 and V2. The number of cells showing all three fluorescent proteins does however decrease significantly for EMBacY, with only half the population expressing all three fluorophores by V4 and even less for V5. The SynBac1 ac87Tn7 variant was the worst performing virus, with little to no fluorescence by V3. The SynBac1, SynBac1 ac20VloxP ac87RoxP and SynBac1 ac32Tn7 variants already showed quite a reasonable proportion of cells only displaying EYFP or two of the fluorophores at V1. Although the starting number of fluorescent cells was less for the SynBac ac110Tn7 and ac112/113Tn7 variants compared to EMBacY at V1, the number of cells displaying all three fluorophores looked to be decreasing less.

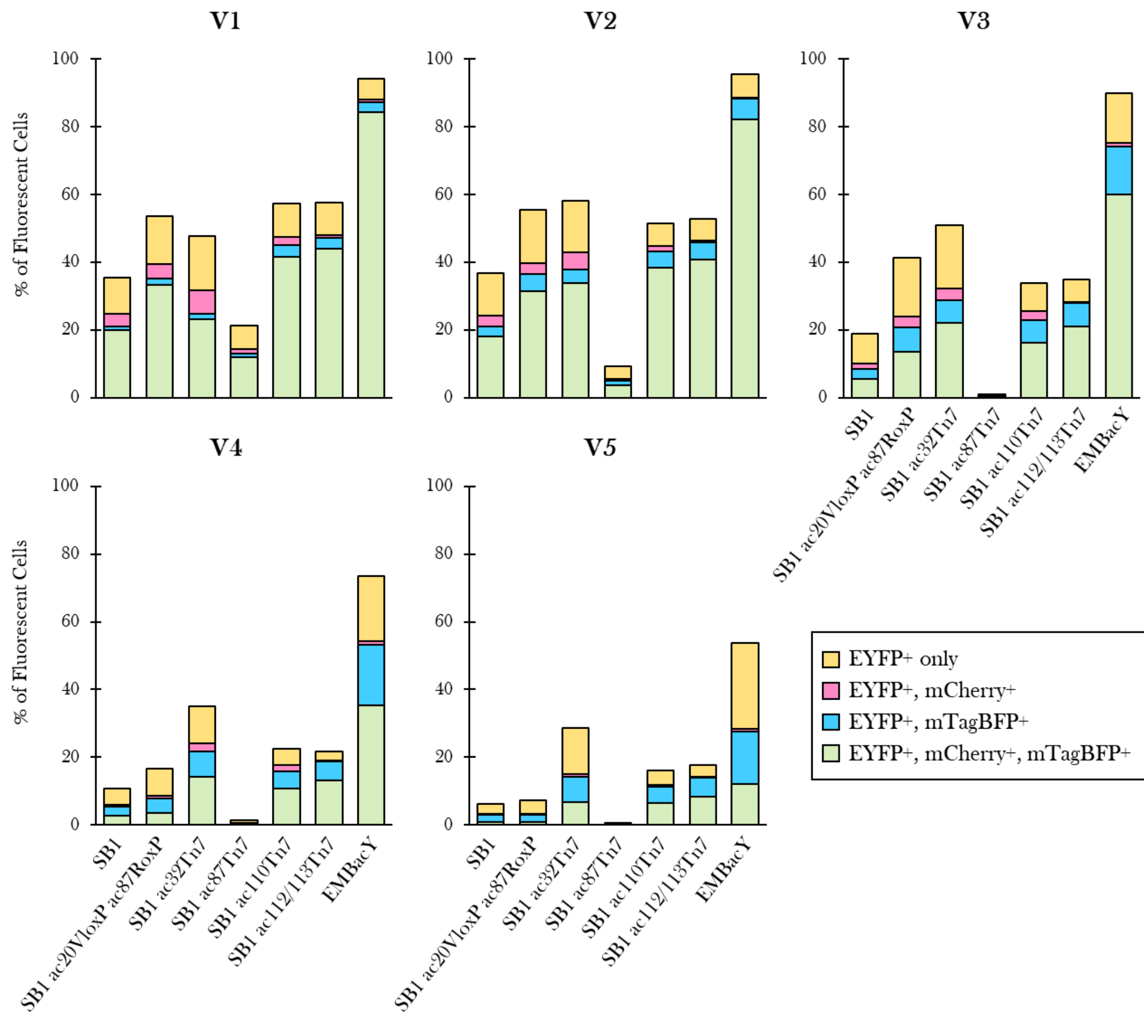


Figure 6.8 – Genome stability analysis.

Data acquired from flow cytometry analysis during V1 to V5 amplification of SynBac1 (SB1), SynBac1 ac20VloxP ac87RoxP, SynBac1 ac32Tn7, SynBac1 ac87Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY expressing mCherry and mTagBFP. Values are averages representing % of total cells analysed (10,000 live cell events). Cells displaying only EYFP fluorescence are shown in yellow, EYFP and mCherry fluorescence are shown in pink, EYFP and mTagBFP fluorescence are shown in blue and cells displaying all the fluorescence, EYFP, mCherry and mTagBFP, are shown in green.

To further analyse the cytometry data in terms of the rate of fluorescence decrease, the percentage of triple fluorescent cells within the EYFP positive population was determined. The data for the V1-V5 generations were plotted, (Figure 6.9), where EMBacY (yellow) shows the biggest decrease in the number of triple fluorescent cells. On the other hand, the rate of decrease for SynBac1 ac110Tn7 and ac112/113Tn7 variants (light blue and pink, respectively) is smaller, whereby V4 and V5 the percentage of triple fluorescent cells within the EYFP positive population is higher compared to the EMBacY.

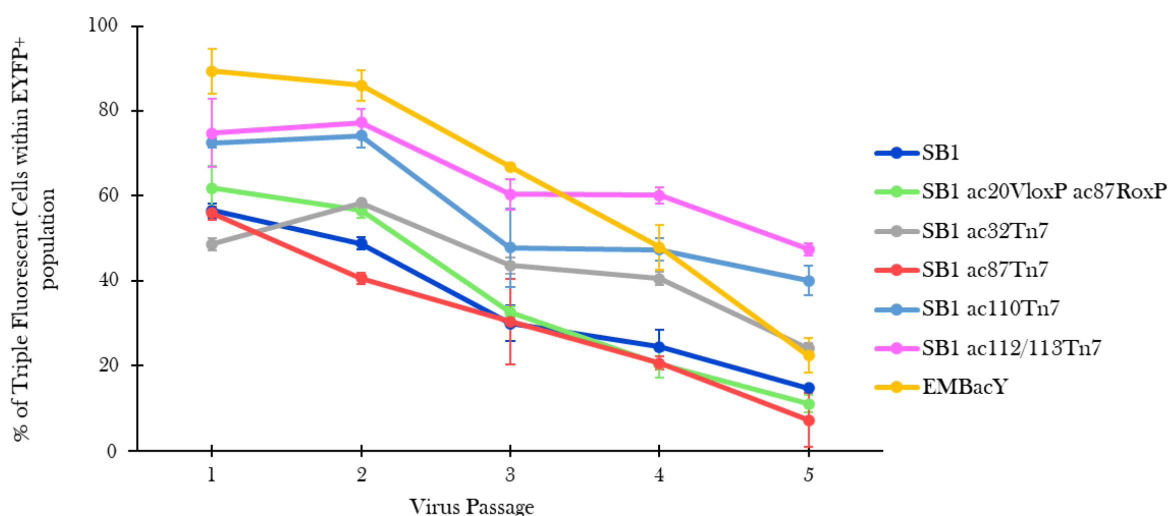


Figure 6.9 – Rate of fluorescence decrease.

Data acquired from flow cytometry analysis during V1 to V5 amplification of SynBac1 (SB1), SynBac1 ac20VloxP ac87RoxP, SynBac1 ac32Tn7, SynBac1 ac87Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. Values represent % of triple fluorescent cells (EYFP, mCherry and mTagBFP) within the EYFP fluorescent cell population. Error bars indicate STDEV (n=3).

6.3.3 Heterologous Target Protein Expression

The last of the validations included testing the new SynBac genome variants in terms of their heterologous target protein expression capabilities compared to the EMBacY genome. A selection of four single protein targets; Keap1, Lck, PI3Kg and PPAR and a protein complex, TAF569, were selected. The recombinant bacmids were generated and transfected into Sf21 cells in triplicates and V1 amplified as per methods section. Samples of the same number of cells were taken at 48 and 72 hpi for EYFP fluorescence measurements and SDS-PAGE analysis. Both the time points of the SDS-PAGE samples were firstly run and the 72 hpi samples were chosen onwards, as they displayed higher yields of the protein compared to the 48 hpi samples. The triplicates of each protein were analysed by SDS-PAGE and the best replicate chosen for the final comparison gel.

6.3.3.1 Single Protein Expression Tests

As a result of the genome stability experiments, only three of the SynBac1 variants were chosen for testing. SynBac1 ac32Tn7, ac110Tn7 and ac112/113Tn7 were tested alongside EMBacY as earlier they showed highest fluorescent protein production. EYFP fluorescence values (Figure 6.10) from 48 and 72 hpi samples collected from the expression of (A) Keap1- 73 kDa, (B) Lck- 56 kDa, (C) PPAR- 58 kDa and (D) PI3Kg- 110 kDa. On the right side, the corresponding SDS-

PAGE gels of the whole cell extract (SNP) and soluble fraction (SN) 72 hpi samples are displayed for one of the triplicates. Overall, all three of the SynBac1 variants showed similar EYFP fluorescence at both 48 and 72 hpi and protein yields compared to the EMBacY system. For Keap1 expression, the SynBac1 ac112/113Tn7 variant displayed lower EYFP values at 72 hpi, yet the protein band detected is comparable to the EMBacY sample. The protein expression for PI3Kg EMBacY sample looks slightly stronger, however the EYFP fluorescence was lower for the SynBac1 ac110Tn7 and ac112/113Tn7 variants that may correlate to the strength of the virus and amount of recombinant protein expressed.

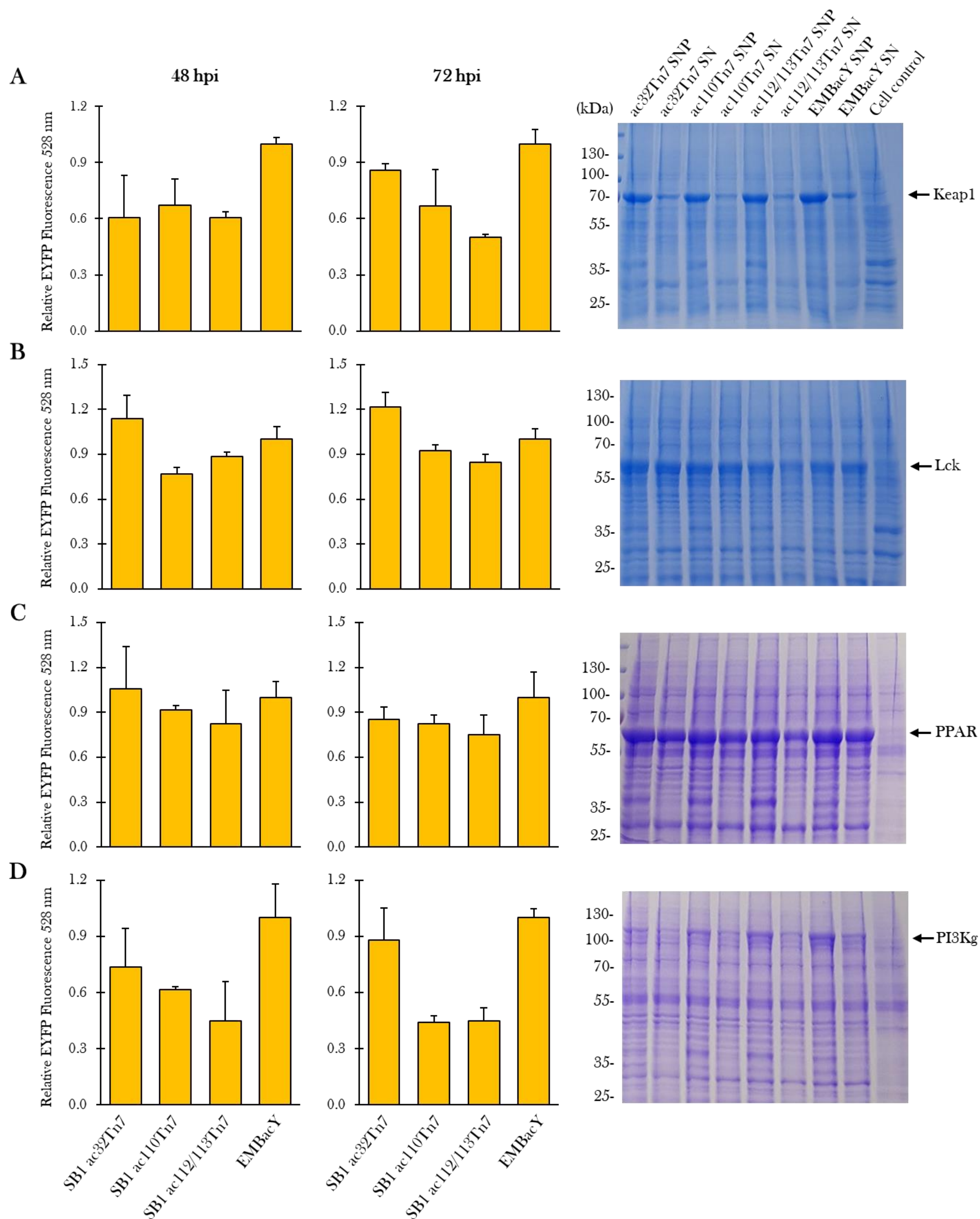


Figure 6.10 – Single protein expression test.

EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification of Keap1 **(A)**, Lck **(B)**, PPAR **(C)** and PI3Kg **(D)** for SynBac1 (SB1) ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3). On the right, 12 % SDS-PAGE gels of Keap1 73 kDa **(A)**, Lck 56 kDa **(B)**, PPAR 58 kDa **(C)** and PI3Kg 110 kDa **(D)** samples collected at 72 hpi from V1 amplification of SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. 14 µl of whole cell extract (SNP), soluble fraction (SN) and cell only control were loaded.

6.3.3.2 Single Protein Expression Retest

As the protein expression of the three SynBac1 variants was similar to the EMBacY virus, even though the fluorescence percentages were lower in the genome stability experiment, we decided to include the unmodified SynBac1 genome in the next trial. The version with the new EYFP under the polh promoter (SynBac1 new EYFP) and the original starting genome with the EYFP under the ie1 promoter (SynBac1 old EYFP) were included. The reasoning behind this is to acquire information about the protein expression capabilities of the SynBac1 genome containing the original Tn7 site. As more viruses needed to be tested, only two of the previous protein candidates were chosen, Lck and PPAR, due to their good protein expression. Additionally, the previously used fluorescent reporter AmCyan was utilised as it is known to be highly over expressed, thus clearly visible on the SDS-PAGE gel. It should be noted the levels for EYFP fluorescence are much lower for the SynBac1 old EYFP virus due to the weaker promoter and although plotted on the graphs, should not be compared to the other samples. Figure 6.11 (A) displays the EYFP and AmCyan fluorescence data for AmCyan expression at 48 and 72 hpi, and the corresponding SDS-PAGE gel of the best 72 hpi replicate. The fluorescence values are comparable across all viruses, where the AmCyan fluorescence confirms the SynBac1 old EYFP variant performs just as well as the other viruses. The protein band at 28 kDa for AmCyan shows equivalent protein expression across all SynBac1 variants and EMBacY. Figure 6.11 (B) and (C) displays the EYFP data and corresponding SDS-PAGE gels for Lck (56 kDa) and PPAR (58 kDa), respectively. Both the EYFP fluorescence and the protein amounts looked essentially the same for all the viruses tested.

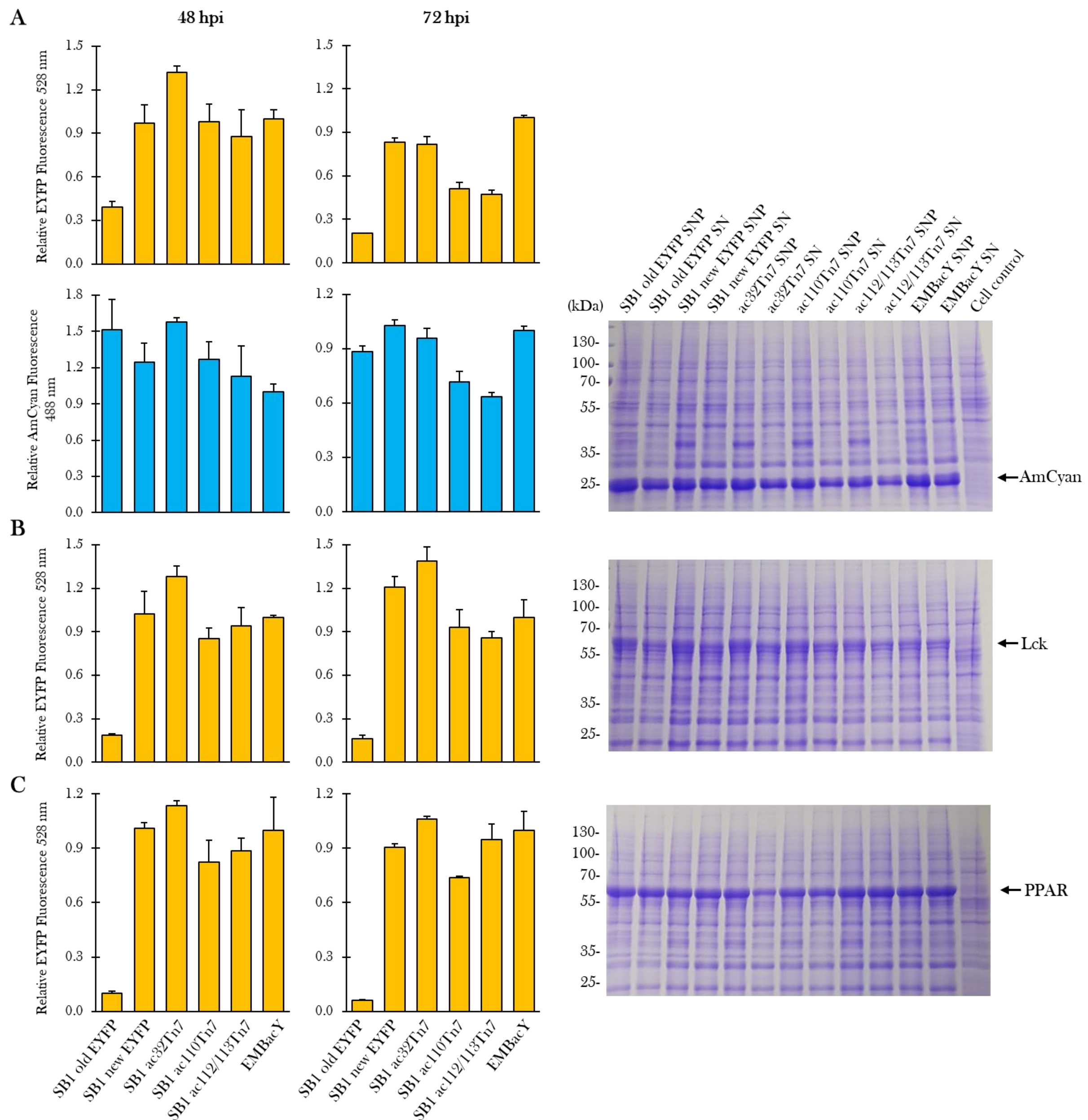


Figure 6.11 – Single protein expression retest.

EYFP production from 500,000 cells taken at 48 and 72 hpi during V1 amplification of AmCyan **(A)** (yellow), Lck **(B)** and PPAR **(C)** for SynBac1 (SB1) old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. AmCyan fluorescence visualised in blue **(A)**. Values represent averages normalized to EMBacY = 1; error bars indicate SEM (n=3). On the right, 12 % SDS-PAGE gels of AmCyan 28 kDa **(A)**, Lck 56 kDa **(B)** and PPAR 58 kDa **(C)** samples collected at 72 hpi from V1 amplification of SynBac1 old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. 14 µl of whole cell extract (SNP), soluble fraction (SN) and cell only control were loaded.

6.3.3.3 Protein Expression at V2 Amplification

Most protein expression for harvesting and purification occurs at V2 amplification, thus the best replicate of the Lck and PPAR V1 virus from the previous experiment was utilised to start a V2 culture as per methods section. Figure 6.12 (A) and (B) illustrates the results from the EYFP fluorescence of 48 and 72 hpi samples and the corresponding SDS-PAGE gels of the 72 hpi samples for Lck (56 kDa) and PPAR (58 kDa), respectively. Once again, the EYFP values were comparable across all viruses (exempting the SynBac1 old EYFP variants), where the SynBac1 ac110Tn7 and ac112/113Tn7 variants showed lower EYFP levels at 48 hpi but reached similar levels at 72 hpi compared to the EMBacY system. The SynBac1 ac32Tn7 variant achieved higher EYFP fluorescence at 72 hpi than EMBacY for Lck expression, which the slightly stronger protein band on the SDS-PAGE may correlate to. Overall, the protein amounts obtained for both, Lck and PPAR, are of similar intensities across all the SynBac1 variants and EMBacY within the resolution of the experiment.

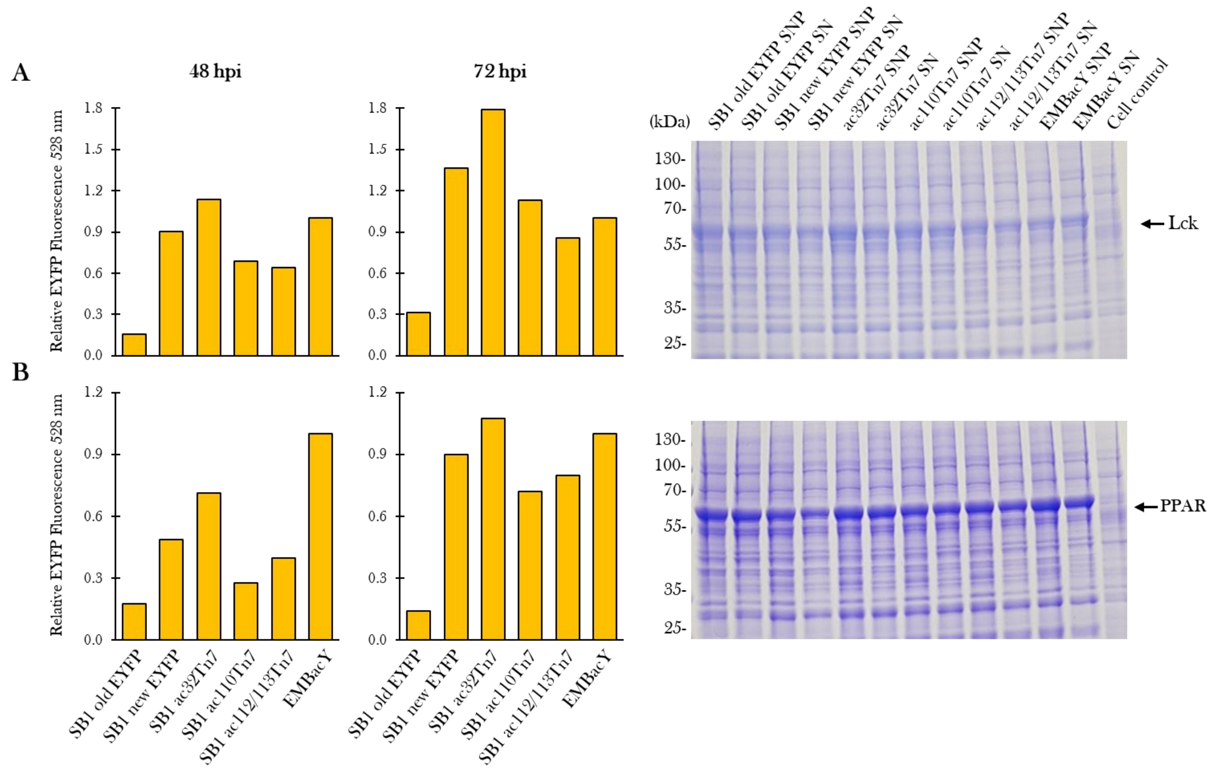


Figure 6.12 – Protein expression at V2 amplification.

EYFP production from 500,000 cells taken at 48 and 72 hpi during V2 amplification of Lck **(A)** and PPAR **(B)** for SynBac1 (SB1) old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. Values normalized to EMBacY = 1. On the right, 12 % SDS-PAGE gels of Lck 56 kDa **(A)** and PPAR 58 kDa **(B)** samples collected at 72 hpi from V2 amplification of SynBac1 old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. 14 µl of whole cell extract (SNP), soluble fraction (SN) and cell only control were loaded.

6.3.3.3 TAF569 – Multi-subunit Protein Complex Expression

A test was performed utilising a protein complex, TAF569. Figure 6.13 (A) and (B) displays the EYFP and mCherry, respectively, fluorescence data for the 48 and 72 hpi samples. Altogether, most of the SynBac1 variants, apart from SynBac1 new EYFP, show similar levels of mCherry fluorescence at 48 hpi, with the values increasing at 72 hpi compared to the EMBacY virus. The three SynBac1 variants with newly located Tn7 sites show higher EYFP values at both time points compared to the EMBacY system. As the TAF569 complex is known to be not as highly expressed as the previous single protein candidates, a western blot analysis, (Figure 6.13 (D)) was performed alongside the SDS-PAGE (C) of the 72 hpi samples. The bands for TAF5 – 92 kDa, TAF6 – 73 kDa and TAF9 – 30 kDa were not easily distinguishable on the SDS-PAGE, where the band between the 35 and 25 kDa marker corresponds to the mCherry expression and not TAF9. The TAF5 and TAF6 subunits are known in our lab to run higher on the SDS-PAGE gels than their predicted molecular weights, with bands around 110-120 kDa for TAF5 and ~80 kDa for TAF6. The TAF6 band is seen, notably in the EMBacY samples, but faint bands were visualised for all the remaining SynBac1 variants. The 6His tagged TAF6 and TAF9 subunits were successfully identified at the predicted location, with a positive His tagged control, TEV protease, visible around the 17 kDa height. The western confirmed higher protein production of the TAF6 subunit for EMBacY samples compared to the SynBac1 variants. The CBP (calmodulin binding protein) tagged TAF5 subunit showed a higher degree of background during western development, however a band around the 110-120 kDa mark was visible, which was not present in the cell control sample.

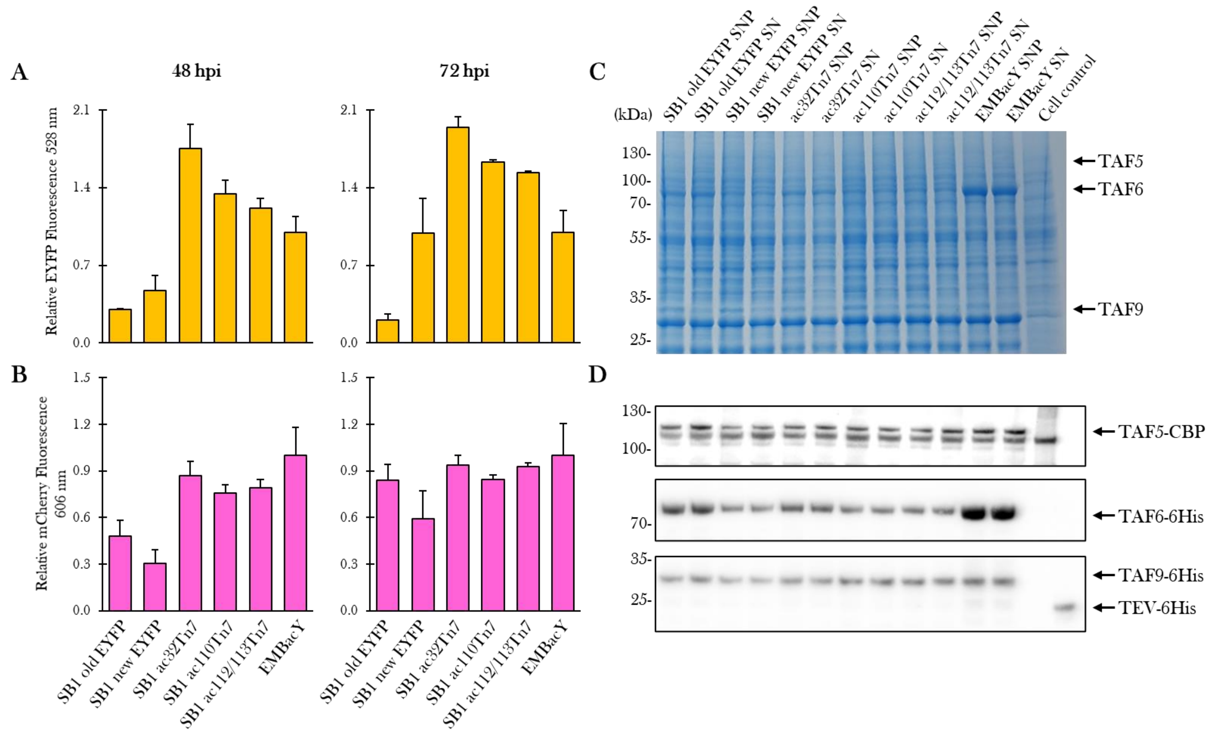


Figure 6.13 – TAF569 – multi-subunit protein complex expression.

(A) EYFP and **(B)** mCherry production from 500,000 cells taken at 48 and 72 hpi during V1 amplification of TAF569 for SynBac1 (SB1) old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. Values normalized to EMBacY = 1, error bars indicate SEM (n=3). **(C)** 12 % SDS-PAGE gel and **(D)** Western blot analysis of Sf21 cell samples collected at 72 hpi from V1 amplification of TAF569 (TAF5- 92 kDa CBP tagged, TAF6- 73 kDa and TAF9- 30 kDa His tagged) for SynBac1 old EYFP, SynBac1 new EYFP, SynBac1 ac32Tn7, SynBac1 ac110Tn7, SynBac1 ac112/113Tn7 and EMBacY. 14 µl of whole cell extract (SNP), soluble fraction (SN) and cell only control were loaded, as well as a positive His tagged control, TEV protease- 17 kDa. His tag detection utilised Anti Penta Histidine HRP conjugate and CBP detection utilised primary Rabbit Polyclonal CBP tag antibody, and secondary Goat anti-Rabbit HRP antibody.

6.4 Discussion

In this chapter, the validation of the six new SynBac variants was presented. Firstly, the site-specific recognition sites, VloxP and RoxP, inserted in the SynBac genomes, were validated by detection of fluorescence from colour donor plasmids fused by the cognate recombinases. Moreover, the stability of the new SynBac genomes, benchmarked against EMBacY, was explored by monitoring the number of EYFP, mCherry and mTagBFP fluorescent cells over a period of five virus generations during serial passaging. Lastly, the protein production capability of the SynBac variants was compared to EMBacY by expressing an array of heterologous target proteins including a protein complex.

6.4.1 VloxP and RoxP Sites are Functional

The site-specific recognition sites, VloxP and RoxP, within the newly generated SynBac variants were validated. The colour donor plasmids with corresponding recombination sites, pUADM_RoxP_AmCyan and pUPDM_VloxP_mCherry, were fused into the SynBac genome and the fluorescence of proteins, AmCyan and mCherry, detected by confocal microscopy. Insertion of donor plasmids was successful validating the RoxP and VloxP sites in the SynBac genomes, albeit the efficiency of the reaction was low. Many attempts and adjustments to the protocol were required before a positive clone was identified. Previously, for the EYFP insertion into the LoxP site, a range of bacmid concentrations was tested and concentrations between 3-7 μg yielded the highest number of colonies for a given site-specific recombination reaction. Of note, commercially sourced Cre recombinase was used for this Cre/LoxP fusion. For the RoxP and VloxP recombination the amounts of the cognate recombinase enzymes, Dre and VCre, needed to be increased. This was likely due to inferior preparation quality due to the Dre and VCre enzymes being purified in-house and the frozen protein stocks having relatively low concentration. It was found that increasing the incubation time of the reaction to 1 hour 15 minutes, as opposed to 45 minutes or an hour, resulted in a higher number of colonies. To increase the efficiency of the recombination reactions, the purity and yield of the recombinases likely needs to be improved. The purification protocol utilised for Dre and VCre only included a single polyhistidine tag affinity purification step. Adding ion exchange and size exclusion will likely improve protein quality. The expression protocol of the Dre and VCre could also be optimised. A recent study exploring diverse expression and purification protocols for Cre recombinase yielded best results when cultures were grown at 37°C after inoculation and harvested after 4 hours.[325] The enzymes generated in this thesis were purified from cultures

incubated at 18°C after inoculation and harvested after only 2 hours. Thus, there is ample room for improvements. Lastly, the recombinase activity of the enzymes can be assessed in a cleavage assay,[326] similar to the validation of the recombinases in Section 3.3.1.3.1. Incubating the enzymes with a plasmid containing a fragment flanked by the site-specific sites would result in the excision of the fragment, which can be conveniently visualised by DNA gel electrophoresis. The enzymes utilised in the validation of the RoxP and VloxP sites in this chapter were not freshly purified but stored frozen over time, therefore the activity of the enzymes may also have decreased, for instance by the freezing procedure. For future recombination reactions, enzyme activity should be tested routinely via the cleavage assay prior to usage. While the insertion of the RoxP and VloxP donors was managed, further optimisation of expression methods and routine quality control needs to be performed to improve the overall procedure.

6.4.2 Determination of Genome Stabilities

Using fluorescent proteins as a proxy to follow viral infection over five generations, we estimated the genome stabilities of the SynBac variants relative to EMBacY, (Figure 6.8). Overall, the highest percentage of fluorescent cells, ~90% between V1 and V3, was detected for the EMBacY virus, implying best virus transfection and replication. The number of triple fluorescent cells was however declining rapidly. At generation V3, 60% of the cells were EYFP, mCherry and mTagBFP compared to 90% EYFP positive cells. By V4, around half of the EYFP positive cells and only 25% by V5, displayed all three fluorescent proteins. All the SynBac variants exhibited lower total percentages of EYFP positive cells as compared to EMBacY, suggesting the SynBac viruses have lower titers when prepared according to the MultiBac standard operating procedure (SOP) used here, which was optimised for EMBacY. The SynBac viruses reached a maximum of only approximately 60% fluorescence at V1 and V2 and the best variants, ac110Tn7 and ac112/113Tn7, displayed approximately 40% fluorescence at V3, over half compared to the EMBacY virus. Although the total number of fluorescent cells was low, the number of triple fluorescent cells was significantly better maintained as compared to EMBacY. V4 and V5 generation SynBac1 ac110Tn7 and ac112/113Tn7 exhibit around 20% EYFP positive cells, but the number of triple fluorescent cells at V5 in total was similar to EMBacY, even though EMBacY showed over 50% of EYFP positive cells. This could be interpreted as a better maintenance of all heterologous cassettes within these SynBac1 variants as compared to EMBacY, which could indicate superior stability, although further work is required to corroborate this interpretation

To investigate the rate of decrease of the mCherry and mTagBFP cells over the five virus generations, a percentage of triple fluorescent cells within the EYFP positive population was calculated, (Figure 6.9). The plotted data confirmed the EMBacY virus was losing the mCherry and mTagBFP fluorescence at a faster rate as compared to the SynBac variants, consistent with EMBacY representing a less stable genome. For instance, the experiment performed followed the SOP established for MultiBac system and specifically for the EMBacY virus, but the SynBac viruses are genetically different specimens. Because the MultiBac SOP was followed, the same volumes of individual virus were administered to start the next passage, however the viral titer between the viruses could differ – the SynBac virus titers may quite possibly be lower resulting in fewer EYFP positive cells. Due to time constrains, it was decided not to determine the viral titer for all the variants before each virus amplification, as the workload involved would have been excessive. The higher stability of the SynBac variants compared to the EMBacY system may therefore also derive in part of their lower starting titers. It has been shown starting a virus amplification with a higher titer virus is not beneficial in terms of virus production and protein expression.[76,77,213] This may be a consequence of oversaturating the cells replication machinery due to the higher numbers of present viruses. The increased replication demand and higher chance of homologous recombination between the entering viruses may lead to replication of interfering virus particles.[327,328] However, these observations come to the fore when multiplicities of infection (MOI), i.e. the number of infectious virions per cell in an infected culture, are significantly above 1. The MultiBac SOP has been standardised, exactly for this reason, to guarantee operation at a MOI below one and this was the protocol followed in the experiments presented. The higher stability of the SynBac genomes may therefore be a result of a lower starting viral titer, which did not overwhelm the infected cells. In summary, the results are suggestive, however, in order to arrive at a final conclusion about the relative genome stabilities of the SynBac variants relative to EMBacY, further work would need to be performed, notably including viral titer determination. Also, an adjusted SOP customised specifically for SynBac variants may need to be established.

Several interesting observations from the genome stability experiments emerged. For instance, SynBac1 with original Tn7 site displayed a lower number of fluorescent cells compared to the three SynBac1's with new Tn7 locations. The rate at which the triple fluorescent cells were decreasing seemed similar though, suggesting the relocation of the Tn7 site did not overly impact on genome stability under the conditions studied. Again, these results need to be revisited in an experiment establishing the exact titers of the viruses analysed to make a conclusive comparison. Another observation was the cytometry data indicating a lower percentage of fluorescent cells for

the SynBac1 variants, whereas the differences in fluorescence intensities collected by fluorimetry, (Figure 6.6), were not as substantial compared to the EMBacY. For example, at V1 for SynBac1 ac112/113Tn7, fluorimetry revealed close to equal values of fluorescence, even though the cytometry data indicated the ac112/113Tn7 variant displayed 2/3 of the number of fluorescent cells compared to EMBacY. Variations between the fluorescence intensities may be more common due to the sampling. Nonetheless, overall fluorescence values for all SynBac1 viruses were consistently better than the cytometry results. Thus, this observation implies that although a lower number of cells are infected, higher levels of heterologous protein are reached, consistent with better protein productivity of the SynBac viruses.

6.4.3 SynBac Variants Protein Yields Match EMBacY

The protein expression capabilities of the best performing SynBac1 variant genomes in the genome stability experiment were investigated and compared to EMBacY. A selection of single protein candidates and a multi-protein complex were expressed. EYFP fluorescence was tracked for the different viruses, as well as protein amounts analysed by SDS-PAGE. As a whole, the EYFP fluorescence and the levels of heterologous target protein expression were similar in between the different SynBac1 variants and matched the levels of EMBacY, (Figure 6.10). Retesting of two proteins, Lck and PPAR, resulted in equivalent protein amounts detected, corroborating the original observation, (Figure 6.11). The protein production equally remained unchanged during V2 amplification of the viruses, (Figure 6.12). The expression of TAF569 by the SynBac variants also matched EMBacY, except one subunit TAF6, which EMBacY curiously seemed to express at a higher level in this stoichiometric (1:1:1) complex (Figure 6.13). The new SynBac1 EYFP variant with the original Tn7 attachment site located in the 'classical' position performed worse than the other three SynBac1 variants with relocated Tn7 attachments sites in the cytometry analysis, (Figure 6.8). However, the protein amounts were not reduced for SynBac1 with original Tn7 site. The likelihood that the SynBac1 comprising the original Tn7 site arrangement produces the same protein quantities, although a lower percentage of fluorescent cells were previously detected, is somewhat perplexing. However, it could indicate SynBac1 with the original Tn7 site simply may have underperformed, for unclear reasons, in the genome stability experiment.

6.4.4 Concluding Remarks and Future Directions

The insertion of the VloxP and RoxP colour donors into the corresponding sites found in the SynBac1 genomes was successful. The expression of the reporter proteins, mCherry and

AmCyan was confirmed by fluorescence confocal microscopy. For the Dre/RoxP and VCre/VloxP systems to be used in fusion reactions of plasmid and bacmid, further improvements to the method and reagents would be advisable. Specifically, optimising the purification of the recombinases to reach high purity and high yields of the enzymes would be recommended.

The number of EYFP, mCherry and mTagBFP cells was monitored over the course of five virus generations for the new SynBac variants and also EMBacY. The resulting cytometry data indicated the EMBacY virus showed highest initial percentage of fluorescent cells (i.e. infected), but also displayed the steepest loss of triply fluorescent cells as compared to EYFP positive cells. The SynBac viruses exhibited lower initial percentage of fluorescent cells, however the relative number of mCherry and mTagBFP positive cells was better maintained as compared to EMBacY. The results suggest the SynBac variants exhibit higher genome stability, which would resonate with the original hypothesis underlying this work. However, the ambiguity remains that the virus titers were not determined in the experiment, rather, a SOP developed previously for MultiBac/EMBacY was followed to safeguard working at a MOI below 1. Future work would involve repeating the experiment and determining virus titers before starting the next viral amplification and adjusting the virus volumes deployed accordingly. With identical viral titers, the experimental results will be more conclusive, and enable to truly define the relative genome stabilities, ideally over even more virus passages.

Finally, the SynBac variants and EMBacY were tested for heterologous protein expression using an array of target proteins. The protein amounts detected by SDS-PAGE were virtually the same across all viruses. These results validated that all the SynBac variants tested matched EMBacY in terms of heterologous target protein production. Thus, the sum of genomic interventions did not impact detrimentally on the functionality of the baculovirus genome as a recombinant protein production tool in a cell culture setting.

Chapter 7: Conclusions and Outlook

7.1 Thesis Overview

The baculovirus expression vector system (BEVS) is among the most widely used eukaryotic expression methodologies for producing heterologous target proteins in academia and industry. Originally conceived as a biological pesticide to alleviate crop damage by the fall army worm, it is now accelerating R&D projects and pharmacological production of biologics world-wide. Since the initial discovery and utilisation of the polh locus for heterologous expression, the BEVS underwent numerous improvements to render it more user-friendly and economic, increasing its utility. A significant advance included the insertion of a bacterial origin of replication and a Tn7 attachment site embedded in a LacZα cassette, considerably simplifying foreign gene insertion and the propagation and handling of baculovirus genomes for a range of applications. The introduction of the MultiBac system and derivatives thereof, provided a step-change expanding the scope of BEVS for multiprotein complex structure/function research, at a time when it was increasingly realised that these molecular machines are the predominant catalysts of cellular processes.[160,329] Other exciting developments include the use of BEVS as a large capacity DNA delivery toolkit for transducing mammalian cells, tissues and potentially, whole organisms, ultimately in the context of gene therapy.[330] Recent advances in synthetic biology, including efficient recombineering methods and steep decline of the costs for custom DNA synthesis, unlocked the baculoviral genome to rewiring and enhancing, opening up the possibility for erasing wild-type code that may not be required, or even detrimental for a desired application of choice.

Deletion of several genes had already been accomplished previously and shown to be beneficial. ChiA and v-cath are deleted in MultiBac and p26, p10 and p74 were removed in flashBAC and improved heterologous target protein expression, setting the stage for investigating further gene deletions. The majority of the proteins encoded by AcMNPV genes and their homologs from other baculoviruses have been studied in terms of function in diverse contexts, usually from an entomological point of view (Tables 8.2 - 8.10 in appendix). Deep-mining and compiling the outcomes of these investigations uncovered many gene products that were likely not essential for virus replication and propagation in cell culture, such as proteins involved in *per os* infectivity. This compiled information has been used to design a radical blueprint for a theoretically minimal genome, SynBac, which could be synthesised *de novo*. In practice, a strategy with intervening validation steps was designed, in which the entire genome was partitioned into seven fragments that could be pasted individually on a progenitor genome, replacing wild-type sequences with condensed, minimal synthetic ones. EMBacY, a thoroughly validated and widely used genome

was chosen as the progenitor.[160,318] This approach gave rise to hybrid- part wild-type, part synthetic- genomes that could be validated individually. Ultimately, the modifications installed in this first generation of functional minimised genomes could be further combined progressively to produce a minimised genome. The genome minimisation aimed to uncover the minimal operational genome of AcMNPV without affecting the virus amplification and heterologous protein expression functionalities while generally stabilising the viral genome that, at present, suffers from genomic instability.[314]

The instability of the baculovirus genome was explained, in particular the generation of defective interfering viruses in cell culture and large-scale protein production. The accumulation of DIPs was credited to spontaneous deletions during serial passaging of the virus. A particular deletion prone area in the baculoviral genome was found to encompass the swath of bacterial DNA containing the replication origin, kanamycin resistance and the Tn7 site in the LacZα cassette. Due to the smaller genome size and higher density of oris in the DIPs, as compared to the parental, complete virus, the replication of DIPs is preferred in cell culture. As a consequence, DIPs can predominate, sometimes vastly. Because of the concomitant loss of regions encoding the heterologous target proteins, the overall target protein yield decreases over virus generations, which is highly unfavourable, especially for fermenter-scale application. Put simply, until enough baculovirus is amplified to load an industrial fermenter, the number of viruses that support heterologous target gene expression can have decreased to such a minority that the target protein yield per litre culture is severely compromised. The underlying hypothesis of the work presented in this thesis is that a minimal, functional genome devoid of non-essential DNA elements in the laboratory, and with the heterologous target gene insertion sites relocated into area previously verified as 'safe', would reduce, or wholesale alleviate, this problem, enabling BEVS to release its full potential.

Furthermore, novel site-specific recombination systems akin Cre/LoxP but relying on orthogonal DNA site and cognate enzymes have been described. These more recent systems and their precise recognition sites show high specificity, enabling parallel use for DNA engineering applications. Four different systems were investigated in this thesis, both for use as tools to engineer the baculovirus genome itself, and as added functionalities to facilitate heterologous insertion into diverse loci in the final SynBac genome that was aspired to in this work.

The ultimate aim of this project is to create a minimal synthetic baculovirus genome, SynBac, pursuing the following approaches and leading to the following outcomes and conclusions:

1. Generation of minimal baculoviral genome by replacing wild-type segment with a total of seven synthetic fragments in which multiple non-essential genes were deleted and the DNA maximally condensed. This approach proved to be successful for one of seven fragments (Fragment 1) resulting in SynBac1. In the case of the six other fragments, implemented within the validated SynBac1 framework, the resulting viruses all exhibited significantly compromised virus replication kinetics. In conclusion, the fragments seemingly contained DNA elements which, contrary to the expectation, were necessary for proper virus function, either individually or synthetically with other DNAs removed concomitantly, or already absent in SynBac1.
2. Implementing a set of different recombinases and the development of robust methods resulted in an efficient, streamlined protocol for precision engineering of the baculoviral genome. Room for improvement still exists regarding to purification method of the recombinases used. Moreover, activity testing should be performed as a batch-by-batch quality control. Nonetheless, the procedures developed in this thesis, relying on alternative Cre/LoxP-like recombination reactions can be readily and successfully used in combination with the Cre/LoxP system as a SOP for efficient recombineering of this and other genomes in the future.
3. 67 single gene deletions within the framework of the SynBac1 genome were performed, providing critical information into the essential nature of individual genes. Specifically, these results update and rectify conclusions drawn from data mining and literature searches previously, exemplifying those conclusions arrived at in different contexts (e.g. entomology) may not be directly transferrable to a synthetic biology approach to improve a genome by minimisation within the boundary parameters of laboratory work and industrial application. Unlike previously thought, only a relatively small set of genes could be wholesale deleted without negatively affecting the replication and protein expression of the virus.
4. The relocation of the Tn7 site and insertion of VloxP and RoxP sites into these now verified non-essential gene loci generated fully functional viruses, further corroborating the dispensability of these genes for the purposes here pursued.

5. Intriguingly, the validation of the newly generated SynBac variants by stability analysis suggested the SynBac genomes are in fact more stable compared to the progenitor EMBacY, a notion which, if correct, would compellingly validate the approach taken. At the same time, the SynBac variant viruses appeared to be producing lower titers, however this data is not conclusive due to lack of precise virus titer information. The protein expression capability of SynBac variants was however at least equal to EMBacY. This suggests that the SynBac viruses may exhibit higher protein productivity, as they produce at least the same protein levels at lower viral titers.

7.2 Concluding Remarks and Future Directions

7.2.1 Baculovirus Genome Minimisation Strategy: SynBac1

The baculovirus expression system has been a key player among protein production methods. The reported instability of the virus upon serial passaging currently poses limits to the potential of the baculoviral systems for large scale protein production. Over several viral amplifications, deletions within the genome accumulate leading to the generation of non-functional baculoviruses that can dominate the virus population. These viruses commonly display the loss of DNA regions, and preferably those that contain the recombinant target protein expression cassettes, which obviously are not essential for survival of the virus. The net result is a reduction in target protein yields, which can at times be substantial. The serial passaging effect and potential ways to reduce this is a keen object of study and activity, including the initiation of this present project.[97,159,211,331] A large body of existing data strongly indicates that many baculoviral genes are not required for successful virus propagation in cell culture. There are several examples of gene deletions that even improve protein production. A fundamental promise of synthetic biology is that optimisation by minimisation can result in superior designer genomes for a given purpose as compared to natural genomes. A central hypothesis of this thesis thus is that genome minimisations is a potential path to improve genome stability and heterologous target protein expression yields with a baculoviral genome optimised by size reduction. This minimised genome would contain only the required genes for virus replication and protein production in the defined boundaries of cell culture and up-scaling. The properties of this minimised genome, such as lack of transcription and replication of dispensable genes and a higher ratio of replication origins to genome size may confer a replication advantage and, thus higher stability culminating in superior productivity.

The genome minimization project, SynBac, was initiated in the team previously by Deepak Balaji, setting the stage for this thesis work. Seven designer DNA fragments were originally synthesised representing rewired, maximally condensed segments of the progenitor genome (EMBacY) based on deep data-mining and comprehensive literature search. A baculovirus genome comprising all seven fragments replacing the wild-type sequence would only contain the genes and DNA elements absolutely required for virus amplification and heterologous protein production in cell culture. The implementation of Fragment 1, prior to commencement of this work, yielded a fully functional prototype, SynBac1, with heterologous protein production levels matching the precursor genome, inciting optimism. However, additional fragment implementations, iterative or individual, resulted in viruses exhibiting compromised virus amplification and target protein production. These data indicate that at least one or more of the deleted genes or other DNA elements within the respective fragments were, contrary to expectations, not dispensable for proper virus replication in cell culture. Although previous evidence (Tables 8.2 - 8.10 in appendix) suggested the selected genes were non-essential, the combination of DNA deletions within the fragment, or the removal of these genes on top of the already deleted genes within Fragment 1, or both, may have synergised to bring about the deleterious outcome.[332-334] The studies reporting the dispensability of the genes usually focus on an individual gene, thus the removal of a gene alongside other genes may conceivably have negatively impacted the baculovirus. The negative impact may be due to the protein products of these genes performing similar functions. A multiple gene knockout analysis of the non-essential genes in the BmNPV illustrated intricate genetic interactions between the genes, thus the AcMNPV genome is likely to display similar complexity.[335]

Irrespective of the further deletions, the rewiring of Fragment 1 and generation of SynBac1 lead to a functional baculovirus. The protein expression matched the progenitor, and therefore a first minimised genomic framework was in place and a functional baculovirus expression system was created. The success of Fragment 1 rewiring was probably the most likely, as it contained the five genes (ChiA, v-cath, p10, p26 and p74) previously reported to increase protein production in the flashBAC system,[336] as well as several genes strictly involved in the viral infection of insect larvae that are expected to be dispensable in cell culture. Interestingly, the SynBac1 virus displayed a slightly faster virus kinetics accompanied by a quicker decline in cell viability. Deletion of the five main genes to increase protein production may have hidden the consequence of a gene deletion that was not entirely positive. Further examination of the deleted genes within Fragment 1 may

isolate a gene or DNA element, preservation of which may improve the performance of SynBac1, although this is not expected to have a major impact.

7.2.2 Different Recombination Systems can be Simultaneously Utilised for Baculovirus Engineering

To enable seamless engineering of the baculovirus vector, new strategies for inserting and excising DNAs were explored. A set of unique site-specific recombinases and their corresponding recognition sites, Dre/RoxP, VCre/VloxP, SCre/SloxP and Vika/Vox were investigated. These are akin, but strictly orthogonal, to each other as well as to the Cre/LoxP already implemented in SynBac1. The four recombinases were successfully expressed in *E. coli* and purified via a rapid one-step procedure exploiting polyhistidine-tag affinity chromatography. Donor plasmids containing four of the different site-specific recognition sites were constructed, barcoded by different antibiotic markers following the modular concept originally laid out for MultiBac.[150] Different reporter fluorescent proteins were inserted for validation purposes. The activity of the recombinases was tested by utilising the plasmids that contained the fragments destined for deletion in the original baculovirus genome minimisation plan confirming the activity of the recombinases. Further tests were performed by fusing donor and acceptor plasmids together in an *in vitro* recombination reaction. All four systems generated the expected DNA fusions and did not cross-react with the Cre/LoxP system. This approach was also demonstrated to work when deployed on bacmid representing the baculoviral genome, albeit at lower efficiency, likely due to the large size of the bacmid (>100 kb). In summary, the reactions are simple to perform following a standardised protocol and enable seam- and marker-less engineering of the baculoviral genome of choice.

Additionally, two of the recombination systems, Dre/RoxP and VCre/VloxP, were verified by fusing the fluorescent donor plasmids into the recognition sites engineered into the baculoviral genome (Chapter 6). Although the efficiency of the reaction was relatively low, successful fusion was accomplished and all expected fluorescent proteins were expressed, validating the functionality of the insertion sites. The presence of the recognition sites in the baculoviral vector offers a highly useful way to insert DNAs of choice, e.g. multiple target protein encoding genes or auxiliary DNA conferring desired functions, encoding for e.g. glycosylases or chaperones, in a distributed fashion across the genome, thus avoiding potentially detrimental accumulation of large swathes of foreign DNA as would be the case if only one insertion site, such as attTn7, was utilised. Protein expression cassettes may also be inserted into the bacmid via these additional

sites, which may decrease the likelihood of deletion upon serial passaging of the virus, a frequent feature for large-scale protein production. Further work should now address the stability of these integrations to conclusively establish the loci used as 'safe' sites. Also, the recombinase purification protocol could be improved by adding ion exchange and size exclusion steps.

7.2.3 Targeted Single Gene Deletions Identify Truly Dispensable DNA

The strategy to further minimise the SynBac1 genome was re-thought based on the results obtained when pursuing the original bold, but ultimately unsuccessful, approach. To gain more insight into the essential nature of the genes within the planned fragment deletions and additional genes found elsewhere in the baculoviral genome, single gene deletion viruses were pursued (Chapter 4). A total of 67 knockout viruses were generated within the framework of the SynBac1 genome using Red/ET, and a blasticidin cassette inserted in place of the deleted gene. Each single gene deletion virus was screened by colony PCR and examined in terms of fluorescent protein expression of EYFP and AmCyan compared to the control, SynBac1. Overall, the results highlighted that a considerable number of genes in the original fragment design could not be wholesale deleted, possibly due to overlapping genes and other DNA elements. Eight *bona fide* candidates for complete gene deletion were identified and validated, with five of them, ac20, ac32, ac87, ac110 and ac112/113 displaying higher or matching levels of EYFP and AmCyan fluorescence reporters as compared to the starting SynBac1 genome (subjective to a statistical analysis). The remaining three genes, ac7, ac45 and ac56 also showed comparable albeit consistently slightly lower levels of fluorescence reporter signal. In summary, these investigations identified truly dispensable DNA regions, which were subsequently used for expanding the scope and functionality of the SynBac1 genome.

The conclusions drawn from these results indicated a higher than anticipated complexity of the baculovirus genome. Specifically, an inactivation of the protein may result in a viable virus, however the encoding DNA may be required for driving expression of neighbouring genes, maintaining the topology or stabilisation of the genome during virus replication. In this study, the single gene deletions encompassed the entire open reading frame from start to end, whereas in numerous studies beforehand, the classification of gene essentiality has been mostly determined by inactivating the gene through inserted mutations or partial deletions. In those cases, it does not necessarily mean the genes are genuinely not essential, as parts of the genes are still present. Rather the proteins that would be under normal circumstances expressed from these genes are

dispensable. For the genome minimization goal of this project, complete gene deletion was pursued to determine the maximum DNA sequences in the genome that can be deleted. However, due to proximity or overlapping of genes on the plus and minus strand of the baculoviral genome, the deletion of whole gene sequences can evidently demonstrate negative impact on the virus.

Future work can include a second, refined round of testing that omits the now deleted genes and focuses on the deletions identified currently as deleterious. The refined single 'gene' deletion test should include deleting smaller fragments than the entire genes, preserving 30-50 bases on either side of the gene to avoid any effects on neighbouring genes and their promoter/terminator regions. Additional targets omitted in the original plan could be deleted and therefore should be investigated in further genome minimisations. The genes reported to be non-essential include ac12, a F-box like protein, ac55 and ac73, for which a BmNPV homolog knockout generated viable virus, ac124, involved in chitinase transcription and three more genes implicated in oral infection, ac96 (pif-4), ac108 (pif-9) and ac145.[41,44,285,337–339] Additionally, solely the terminator of ac68 was disturbed in the original Fragment 5 deletion, but may perhaps be completely deleted, as well as ac97, which lacks homologs and might therefore be non-functional.[42,286]

7.2.4 Tn7 Site Relocation and VloxP/RoxP Insertion Confirmed Results of the Single Gene Deletion Trial

The single gene deletion screening revealed a set of genes suitable for wholesale deletion. These dispensable genes were earmarked as potential new loci for relocating the Tn7 transposition site, to improve genome stability. Chapter 5 presented the work of generating SynBac1 variants containing four alternative Tn7 sites, as well as the insertion of the VloxP and RoxP recognition sites. The EYFP cassette under the control of the polh promoter was then inserted into the LoxP site of the newly generated SynBac1 Tn7 variants to enable later validation compared to the EMBacY system. The SynBac1 variants were examined in terms of EYFP and AmCyan expression compared to EMBacY, with all SynBac1 viruses matching the progenitor, evidencing full functionality of the relocations. Out of the six virus variants tested, SynBac1 ac32Tn7 ac20VloxP displayed the least variations between the fluorescence values obtained in triplicate studies, with both EYFP and AmCyan values consistently between 90-120% compared to the control over all experiments performed. Another two promising genomes were SynBac1 ac110Tn7 and ac112/113Tn7 with EYFP and AmCyan levels reaching the same values as EMBacY but at 48 hpi already. In conclusion, the relocation of the Tn7 sites and insertion of

the RoxP and VloxP sites validated the previously performed single gene deletion trial in Chapter 4. The chosen genes (ac20, ac32, ac87, ac110 and ac112/113) could indeed be deleted without affecting the virus replication and protein expression capabilities of the engineered baculovirus. It should be noted as due to a lack of a statistical analysis performed when comparing the SynBac variants to EMBacY, not all of the SynBac variants necessarily performed as well as the EMBacY virus, thus the claim that all the genes could be deleted without negatively impacting the virus is only suggestive.

7.2.5 SynBac1 Variants: Target Protein Expression and Genome Stability

SynBac1 variants containing the original and four newly located Tn7 sites were subject to genome stability and protein expression analyses benchmarked against EMBacY (Chapter 6). Viral infection, represented by the number of infected cells, was tracked over five generations utilising the EYFP located on the bacmid backbone. Moreover, mCherry and mTagBFP inserted into the Tn7 cassette were tracked as exemplars of recombinant target protein expression. Genome instability would be reported by loss of mCherry and mTagBFP fluorescence in cells which produce EYFP, correlating to losses incurred during virus replication. Of note, the EYFP fluorescence in the viral LoxP site has been shown to be stable in a very large number of recombinant protein production experiments, by the Berger lab and many others,[160,318] indicating that the site where EYFP had been inserted is a 'safe' site not prone to deletion during amplifications, and thus suitable as a robust infection control. The experiments performed here suggest that the EMBacY virus had the highest initial infection rate, but also exhibited higher genome instability resulting in the steepest decline of mCherry and mTagBFP among the genomes analysed. On the other hand, the best SynBac1 variants reached infection efficiencies of only about 60-70% of EMBacY, however the rate at which the mCherry and mTagBFP fluorescence diminished, was markedly reduced. This would suggest that SynBac1 variants were lower titer viruses compared to EMBacY. These lower titer SynBac1 viruses then exhibit higher stability of the genome.[76,77,213] However, it has to be noted that the exact viral titers were not determined in the experiments performed. Rather, a standard operating protocol (SOP), originally developed for MultiBac, was utilised for all viruses, which stringently safeguards a multiplicity of infection (MOI) of less than 1.[160] The enhanced stability of the SynBac1 variants observed is therefore suggestive but remains somewhat speculative at this moment. To conclusively establish SynBac variants as superior in stability to EMBacY, the experiments should

be repeated with precise virus titrations and potentially adjustment of virus volumes administered during infection.

The results suggested that SynBac1 produced lower titer virus, however the heterologous protein production capabilities of the SynBac variants somewhat surprisingly proved to be equal to EMBacY as evidenced by expressing an array of test target proteins including a multi-subunit protein complex. Across the different expression trials, the SynBac1 variants consistently displayed similar protein amounts produced as EMBacY. The only variation seen was a higher expression of a subunit, TAF6, of the protein complex tested with EMBacY, however the remaining two subunits were equally expressed with the SynBac variants, possibly indicating that the apparently higher levels of TAF6 in EMBacY in reality may be a superposition of TAF6 and a baculoviral protein absent in the SynBac repertoire. The fluorescence intensities determined by fluorimetry in the genome stability experiments similarly suggested that heterologous protein expression of the best SynBac1 variants matched EMBacY, compared to the cytometry data. Thus, the SynBac1 variants may exhibit increased protein productivity levels, even if they produce weaker viruses. This would be remarkable as it would mean that a lower titer SynBac virus produces at least as much target protein as a higher titer EMBacY virus. To validate this result, the viral titer needs to be determined and related back to the level of target protein produced.

Overall, further work needs to be undertaken to confirm and corroborate the apparent gain in genome stability of the SynBac1 variants. Nonetheless, it can be stated that the SynBac1 viruses produce, seemingly at lower virus titers, at least equivalent amounts of proteins as compared to the progenitor, EMBacY. In light of a lack of viral titer determination and precise statistical tests to verify the performance of the SynBac1 viruses to EMBacY, the claims can only be true when following the standard operating procedure (SOP) of virus generation in this thesis. However, utilising this SOP, the SynBac1 variants were able to express comparable levels of protein to EMBacY, thus confirming that the interventions performed including deletions, relocation of the Tn7 attachment site and outfitting with additional site-specific integration modalities, resulted in functional viruses without negatively impacting target protein production capabilities.

7.2.6 Alternative Strategies for Baculovirus Improvement

The genome minimisation method explored in this thesis has not been the sole technique utilised for baculovirus improvement. Numerous groups dedicate their research for the goal of improving the baculovirus expression system, with several strategies employed. The following section

exemplifies some of the various approaches that have been reported and could be applied in the future to upgrade our own systems.

7.2.6.1 Transfer Plasmid Modifications

One method involved the modification of the transfer and helper plasmids to reduce DNA carryover during transfection. During preparation of the recombinant bacmid, a high percentage of the DNA may come from the untransposed expression vectors and helper plasmid that are capable of replication in the standard DH10β strain. The cationic polymers or lipid-based reagents utilised for insect cell transfection are highly sensitive to overall amount of DNA present, thus varying amounts of transfer and helper plasmid in the bacmid preparation affects the reproducibility of transfections, causing variations in virus titers. To eliminate unwanted DNA carryover, reports of modifying the transfer vectors with temperature sensitive origins have been described.[340,341] Similarly to the use of the pRed/ET plasmid in this thesis, the plasmids are removed after transposition by culturing at higher temperatures. Another modification included the R6Kγ conditional origin of replication. The resulting transfer vectors can be maintained in *E. coli* strains carrying the *pir* gene, but replication is not supported upon transformation of *pir* negative strain containing the bacmid, thus eliminating the background plasmid levels in bacmid preparations.[342] Nevertheless, the R6Kγ origin is already a part of the original MultiBac system plasmid suite that will remain the basis of a future SynBac system. However, other conditional origins of replication orthogonal to R6Kγ exist, whether or not they should be implemented depends necessarily on the cost/benefit ratio.

7.2.6.2 Tn7 Transposition Improvement

At time of discovery the complete Tn7 transposase segment was placed in the helper plasmid that is required for targeting the Tn7 attachment site in the baculoviral genomes utilised and developed in this thesis. However, the function of all the five transposition proteins (TnsABCDE) were unclear at that time. It has since been determined that TnsA and TnsB form the active transposase, TnsC interacts with TnsAB and directs it to the target DNA via interaction with TnsD or TnsE, which are DNA binding proteins. Transposition into the attTn7 site is mediated by TnsD, allowing vertical transfer into daughter cells, whereas TnsE mediates horizontal transmission among plasmids that can move between other bacteria.[137,343] TnsE is thus dispensable for heterologous gene insertion and construction of a new helper plasmid that did not contain the TnsE gene was described. The natural attachment site of the Tn7 transposase in the *E. coli* genome is proximal to the stop codon of *glmS,* from which the attTn7 site in the

bacmid is derived. The natural site, if unoccupied, sequesters transfer plasmid by transposition of the insert into this locus instead, effectively halving the number of productive integration events in the baculoviral attTn7 site.[344] In common commercial kits (e.g. Bac-to-Bac™, MultiBac), the bacteria comprising the baculoviral genome as a bacmid had their glmS integration site inactivated exactly for this reason. In the work here presented, fresh competent *E. coli* cells with intact *glmS* were necessarily used to accept the newly generated genome. To maximise productive transposition efficiency, the *glmS* site in the *E. coli* chromosome also should be blocked e.g. by insertion of an antibiotic resistance marker.[342]

An alternative approach to the canonical Tn7 transposition system is the utilisation of a CRISPR-associated transposase. Two recent reports characterise RNA-guided DNA insertion with CRISPR-associated transposases from two different species, *Scytonema hofmanni* and *Vibrio cholerae*.[345,346] These systems consist of Tn7-like transposase subunits and CRISPR effectors with deficient nuclease activity. The RNA-guided DNA transposition occurs unidirectionally and inserts approximately 50 to 60 base pairs, dependent on the different system, downstream of the protospacer adjacent motif. The integration efficiencies of both of the systems were analysed by targeting multiple sites in the *E. coli* genome, with 80-95% on-target insertions without positive selection. These results suggest high specificity of the RNA-guided DNA transposition and could be employed as fully programmable integrases. Similarly to the TnsABCDE cassette found within the helper plasmid, the cascades for the CRISPR-associated transposes can be provided in trans. The gRNA could additionally be provided on a separate donor or integrated into the transfer vector containing the heterologous expression cassette flanked by the transposase recognition sequences. The flexibility of integrating the expression cassette anywhere in the genome, as long as a suitable PAM is present and not within an essential gene, could facilitate the discovery of the best integration site to minimise genome instability and increase heterologous gene expression.

7.2.6.3 Delaying Cell Lysis and Enhancing Protein Production

The lytic nature of baculovirus infection sets a limit to protein production. Numerous ways have been employed to delay cell death and increase protein expression, including the deletion of the p26, p10 and p74 genes, likewise implemented in Fragment 1 in SynBac1. Another approach includes inserting specific elements into the baculovirus genome or the expression vector, such as a vankyrin genes. Viral ankyrins are derived from the insect *Campoletis sonorensis ichnovirus* and were shown to delay the lysis of baculovirus infected cells when co-expressed from a dual vector or provided from a stably transformed insect cell line. The reduced cell lysis is a result of

apoptosis inhibition through modulation of the host cell immune response. Depending on the protein target, protein yields were increased by a factor of 4- to 15- fold.[347,348] As the new SynBac1 genomes contain new recognition sites, VloxP and RoxP, the vankyrin gene, P-vank-1, could be inserted into one of the sites and positive effects on protein production examined.

Further methods to delay cell lysis involved construction of expression cassettes consisting of transactivation factors, IE1 and IE0, under the control of the polh promoter, as well as the transcriptional enhancer, hr1. The presence of extra copies of IE1/IE0 in baculovirus infected cells revealed reduced cell lysis, thus overexpression of these factors may promote a longer time frame for protein expression due to the delayed cytopathic effects. The addition of the hr1 sequence increased the protein expression from the cassette but only when the gene of interest was expressed from the p10 promoter and not polh or p6.9.[167] The elements on their own were not sufficient to boost protein production, indicating their interconnectedness, however others reported increase in protein expression when utilising the hr1 sequence alone.[349]

Additionally, approaches to prevent the initiation of the apoptotic process by blocking insect cell caspases by RNA interference has been reported. Protein production has been previously improved by supressing expression of caspases in mammalian cells and stably transformed insect cell lines expressing short-hairpin RNAs (shRNAs), as well as long double stranded RNAs (dsRNAs).[350–352] More recently, expression cassettes containing a shRNA against the insect cell caspase (Sf-caspase-1) were constructed and inserted into the baculovirus vector. The suppression of Sf-caspase-1 resulted in improved accumulation and higher activity of the reporter proteins, however the overall protein yield did not differ, suggesting the suppression of cell lysis promoted better protein folding instead.[353]

Multiple baculovirus encoded genes contain anti-apoptotic activity to counteract the host induced apoptosis and shutoff of protein synthesis. Among the anti-apoptotic genes are p35, p49, iap and apsup, which all have been proven to perform a key role in the virus infection.[354–357] Similarly to the insertion of extra copies of the transcriptional enhancers, IE1, IE0 and hr1, additional copies of the anti-apoptotic genes in the baculovirus genome may aid in higher suppression of host apoptosis.

The above modifications provide numerous possible approaches to improve the efficiency and productivity of the baculovirus system. The field is abundant with many more approaches of enhancing the system, such as protocol optimisations and creation of a variety of expression vectors allowing multiple protein expression. Reports utilising the baculovirus expression system

or studying the virus itself are constantly published, thus our knowledge is continually expanding, providing ample impetus to corroborate, modify and update genome minimisation strategies. To explore 'minimal baculovirus genome space', several interesting and useful methods could also be employed, such as the earlier described SCRaMbLEing technique (Section 1.4.3.4). Utilising this approach would enable the discovery of more complex gene interactions, duplications of beneficial genes or relocation of genes within the baculovirus genome. In summary, there is a plethora of work to be done with the goal of improving the stability and protein expression yields of a minimal, increasingly synthetic baculovirus expression vector system. The work presented in this thesis provided important pieces to this puzzle, highlighting the complexity of the baculovirus genome, with the creation of new SynBac variants paving the way for future genome engineering.

References

1. Ackermann, H. W. & Smirnoff, W. A. A morphological investigation of 23 baculoviruses. *J. Invertebr. Pathol.* **41**, 269–280 (1983).

2. Jehle, J. A. *et al.* On the classification and nomenclature of baculoviruses: A proposal for revision. *Archives of Virology* vol. 151 1257–1266 (2006).

3. Rohrmann, G. F., Pearson, M. N., Bailey, T. J., Becker, R. R. & Beaudreau, G. S. N-terminal polyhedrin sequences and occluded Baculovirus evolution. *J. Mol. Evol.* **17**, 329–333 (1981).

4. Martignoni, M. E. & Iwai, P. J. A catalog of viral diseases of insects, mites, and ticks. *Fourth ed. USDA For. Serv. PNW-195* **195**, 897–911 (1986).

5. Harrison, R. L. *et al.* ICTV virus taxonomy profile: Baculoviridae. *J. Gen. Virol.* **99**, 1185–1186 (2018).

6. Thézé, J., Bézier, A., Periquet, G., Drezen, J. M. & Herniou, E. A. Paleozoic origin of insect large dsDNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15931–15935 (2011).

7. Wu, G. A., Jun, S. R., Sims, G. E. & Kim, S. H. Whole-proteome phylogeny of large dsDNA virus families by an alignment-free method. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12826–12831 (2009).

8. Jehle, J. A. *et al.* Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology* **346**, 180–193 (2006).

9. Zanotto, P. M., Kessing, B. D. & Maruniak, J. E. Phylogenetic Interrelationships among Baculoviruses: Evolutionary Rates and Host Associations. *J. Invertebr. Pathol.* **62**, 147–164 (1993).

10. Pearson, M. N. & Rohrmann, G. F. Transfer, Incorporation, and Substitution of Envelope Fusion Proteins among Members of the Baculoviridae, Orthomyxoviridae, and Metaviridae (Insect Retrovirus) Families. *J. Virol.* **76**, 5301–5304 (2002).

11. Ikeda, M., Hamajima, R. & Kobayashi, M. Baculoviruses: diversity, evolution and manipulation of insects. *Entomol. Sci.* **18**, 1–20 (2015).

12. Fuchs, L., Y., Woods, M., S. & Weaver, R., F. Viral transcription during Autographa californica Nuclear Polyhedrosis Virus Infection: a novel RNA polymerase induced in infected Spodoptera frugiperda cells. *J. Virol.* **48**, 641–646 (1983).

13. Guarino, L. A., Xu, B., Jin, J. & Dong, W. A virus-encoded RNA polymerase purified from baculovirus-infected cells. *J. Virol.* **72**, 7985–91 (1998).

14. Miele, S. A. B., Garavaglia, M. J., Belaich, M. N. & Ghiringhelli, P. D. Baculovirus: Molecular Insights on Their Diversity and Conservation. *Int. J. Evol. Biol.* **2011**, (2011).

15. Tweeten, K. A., Bulla, L. A. & Consigli, R. A. Characterization of an Extremely Basic Protein Derived from Granulosis Virus Nucleocapsids †. *J. Virol.* **33**, 866–876 (1980).

16. Wilson, M. E., Mainprize, T. H., Friesen, P. D. & Miller, L. K. Location, transcription, and sequence of a baculovirus gene encoding a small arginine-rich polypeptide. *J. Virol.* **61**, 661–666 (1987).

17. Wilson, M. E. & Price, K. H. Association of Autographa californica nuclear polyhedrosis virus (AcMNPV) with the nuclear matrix. *Virology* **167**, 233–241 (1988).

18. Oppenheimer, D. I. & Volkman, L. E. Proteolysis of p6.9 induced by cytochalasin D in Autographa californica M nuclear polyhedrosis virus-infected cells. *Virology* **207**, 1–11 (1995).

19. Katsuma, S. & Kokusho, R. A Conserved Glycine Residue Is Required for Proper Functioning of a Baculovirus VP39 Protein. *J. Virol.* **91**, (2017).

20. Russell, R. L. Q., Funk, C. J. & Rohrmann, G. F. Association of a baculovirus-encoded protein with the capsid basal region. *Virology* **227**, 142–152 (1997).

21. Ohkawa, T., Volkman, L. E. & Welch, M. D. Actin-based motility drives baculovirus transit to the nucleus and cell surface. *J. Cell Biol.* **190**, 187–195 (2010).

22. Braunagel, S. C., Russell, W. K., Rosas-Acosta, G., Russell, D. H. & Summers, M. D. Determination of the protein composition of the occlusion-derived virus of Autographa californica nucleopolyhedrovirus. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9797–802 (2003).

23. Marek, M., Merten, O.-W., Galibert, L., Vlak, J. M. & van Oers, M. M. Baculovirus VP80 Protein and the F-Actin Cytoskeleton Interact and Connect the Viral Replication Factory with the Nuclear Periphery. *J. Virol.* **85**, 5350–5362 (2011).

24. Vanarsdall, A. L., Pearson, M. N. & Rohrmann, G. F. Characterization of baculovirus constructs lacking either the Ac 101, Ac 142, or the Ac 144 open reading frame. *Virology* **367**, 187–195 (2007).

25. Braunagel, S. C., He, H., Ramamurthy, P. & Summers, M. D. Transcription, translation, and cellular localization of three Autographa californica nuclear polyhedrosis virus structural proteins: ODV-E18, ODV-E35, and ODV-EC27. *Virology* **222**, 100–114 (1996).

26. Beames, B. & Summers, M. D. Location and nucleotide sequence of the 25K protein missing from baculovirus few polyhedra (FP) mutants. *Virology* **168**, 344–53 (1989).

27. Harrison, R. L., Jarvis, D. L. & Summers, M. D. The Role of the AcMNPV25KGene, "FP25," in Baculoviruspolhandp10Expression. *Virology* **226**, 34–46 (1996).

28. Rosas-Acosta, G., Braunagel, S. C. & Summers, M. D. Effects of Deletion and Overexpression of the Autographa californica Nuclear Polyhedrosis Virus FP25KGene on Synthesis of Two Occlusion-Derived Virus Envelope Proteins and Their Transport into Virus-Induced Intranuclear Membranes. *J. Virol.* **75**, 10829–10842 (2001).

29. Wang, Y. *et al.* Autographa californica Multiple Nucleopolyhedrovirus Nucleocapsid Protein BV/ODV-C42 Mediates the Nuclear Entry of P78/83. *J. Virol.* **82**, 4554–4561 (2008).

30. Olszewski, J. & Miller, L. K. Identification and characterization of a baculovirus structural protein, VP1054, required for nucleocapsid formation. *J. Virol.* **71**, 5040–50 (1997).

31. Guan, Z. *et al.* The Autographa californica Multiple Nucleopolyhedrovirus ac54 Gene Is Crucial for Localization of the Major Capsid Protein VP39 at the Site of Nucleocapsid Assembly . *J. Virol.* **90**, 4115–4126 (2016).

32. Braunagel, S. & Summers, M. Molecular Biology of the Baculovirus Occlusion-Derived Virus Envelope. *Curr. Drug Targets* **8**, 1084–1095 (2007).

33. Whitford, M. & Faulkner, P. Nucleotide sequence and transcriptional analysis of a gene encoding gp41, a structural glycoprotein of the baculovirus Autographa californica nuclear polyhedrosis virus. *J. Virol.* **66**, 4763–8 (1992).

34. Olszewski, J. & Miller, L. K. A role for baculovirus GP41 in budded virus production. *Virology* **233**, 292–301 (1997).

35. Yuan, M. *et al.* Identification of Autographa californica Nucleopolyhedrovirus ac93 as a Core Gene and Its Requirement for Intranuclear Microvesicle Formation and Nuclear Egress of Nucleocapsids. *J. Virol.* **85**, 11664–11674 (2011).

36. Hong, T., Summers, M. D. & Braunagel, S. C. N-terminal sequences from Autographa californica nuclear polyhedrosis virus envelope proteins ODV-E66 and ODV-E25 are sufficient to direct reporter proteins to the nuclear envelope, intranuclear microvesicles and the envelope of occlusion derived virus. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4050–4055 (1997).

37. Chen, L. *et al.* Autographa californica multiple nucleopolyhedrovirus odv-e25 (Ac94) is required for budded virus infectivity and occlusion-derived virus formation. *Arch. Virol.* **157**, 617–625 (2012).

38. Wang, X. *et al.* Baculovirus Per Os Infectivity Factor Complex: Components and Assembly . *J. Virol.* **93**, (2019).

39. Ohkawa, T., Washburn, J. O., Sitapara, R., Sid, E. & Volkman, L. E. Specific binding of Autographa californica M nucleopolyhedrovirus occlusion-derived virus to midgut cells of Heliothis virescens larvae is mediated by products of pif genes Ac119 and Ac022 but not by Ac115. *J. Virol.* **79**, 15258–64 (2005).

40. Peng, K. *et al.* Characterization of novel components of the baculovirus per os infectivity factor complex.

J. Virol. **86**, 4981–8 (2012).

41. Fang, M., Nie, Y., Harris, S., Erlandson, M. A. & Theilmann, D. A. Autographa californica Multiple Nucleopolyhedrovirus Core Gene ac96 Encodes a Per Os Infectivity Factor (pif-4). *J. Virol.* **83**, 12569–12578 (2009).

42. Li, G., Wang, J., Deng, R. & Wang, X. Characterization of AcMNPV with a deletion of ac68 gene. *Virus Genes* **37**, 119–127 (2008).

43. Liu, J. *et al.* The Autographa californica multiple nucleopolyhedrovirus ac110 gene encodes a new per os infectivity factor. *Virus Res.* **221**, 30–37 (2016).

44. Boogaard, B., Evers, F., van Lent, J. W. M. & van Oers, M. M. The baculovirus Ac108 protein is a per os infectivity factor and a component of the ODV entry complex. *J. Gen. Virol.* **100**, 669–678 (2019).

45. Braunagel, S. C., Elton, D. M., Ma, H. & Summers, M. D. Identification and analysis of an Autographa californica nuclear polyhedrosis virus structural protein of the occlusion-derived virus envelope: ODV-E56. *Virology* **217**, 97–110 (1996).

46. Harrison, R. L., Sparks, W. O. & Bonning, B. C. Autographa californica multiple nucleopolyhedrovirus ODV-E56 envelope protein is required for oral infectivity and can be substituted functionally by Rachiplusia ou multiple nucleopolyhedrovirus ODV-E56. *J. Gen. Virol.* **91**, 1173–1182 (2010).

47. Carstens, E. B., Krebs, A. & Gallerneault, C. E. Identification of an amino acid essential to the normal assembly of Autographa californica nuclear polyhedrosis virus polyhedra. *J. Virol.* **58**, 684–688 (1986).

48. Ji, X. *et al.* How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *EMBO J.* **29**, 505–514 (2010).

49. Whitt, M. A. & Manning, J. S. A phosphorylated 34-kDa protein and a subpopulation of polyhedrin are thiol linked to the carbohydrate layer surrounding a baculovirus occlusion body. *Virology* **163**, 33–42 (1988).

50. Gross, C. H., Russell, R. L. Q. & Rohrmann, G. F. Orgyia pseudotsugata baculovirus p10 and polyhedron envelope protein genes: Analysis of their relative expression levels and role in polyhedron structure. *J. Gen. Virol.* **75**, 1115–1123 (1994).

51. Blissard, G. W. & Wenz, J. R. Baculovirus gp64 envelope glycoprotein is sufficient to mediate pH-dependent membrane fusion. *J. Virol.* **66**, 6829–6835 (1992).

52. Jiang, Y., Deng, F., Rayner, S., Wang, H. & Hu, Z. Evidence of a major role of GP64 in group I alphabaculovirus evolution. *Virus Res.* **142**, 85–91 (2009).

53. Beniya, H., Braunagel, S. C. & Summers, M. D. Autographa californica nuclear polyhedrosis virus: Subcellular localization and protein trafficking of BV/ODV-E26 to intranuclear membranes and viral envelopes. *Virology* **240**, 64–75 (1998).

54. Nie, Y. & Theilmann, D. A. Deletion of AcMNPV AC16 and AC17 results in delayed viral gene expression in budded virus infected cells but not transfected cells. *Virology* **404**, 168–179 (2010).

55. Wang, Q. *et al.* Budded baculovirus particle structure revisited. *J. Invertebr. Pathol.* **134**, 15–22 (2016).

56. Volkman, L. E. & Summers, M. D. Autographa californica nuclear polyhedrosis virus: Comparative infectivity of the occluded, alkali-liberated, and nonoccluded forms. *J. Invertebr. Pathol.* **30**, 102–103 (1977).

57. Funk, C. J., Braunagel, S. C. & Rohrmann, G. F. Baculovirus Structure. in *The Baculoviruses* 7–32 (Springer US, 1997). doi:10.1007/978-1-4899-1834-5_2.

58. Au, S., Wu, W. & Panté, N. Baculovirus Nuclear Import: Open, Nuclear Pore Complex (NPC) Sesame. *Viruses* **5**, 1885–1900 (2013).

59. Dow, J. A. pH gradients in lepidopteran midgut. *J. Exp. Biol.* **172**, 355–375 (1992).

60. Toprak, U. *et al.* Role of enhancin in mamestra configurata nucleopolyhedrovirus virulence: Selective degradation of host peritrophic matrix proteins. *J. Gen. Virol.* **93**, 744–753 (2012).

61. Kawasaki, Y., Matsumoto, S. & Nagamine, T. Analysis of baculovirus IE1 in living cells: Dynamics and spatial relationship to viral structural proteins. *J. Gen. Virol.* **85**, 3575–3583 (2004).

62. Nagamine, T., Kawasaki, Y. & Matsumoto, S. Induction of a subnuclear structure by the simultaneous expression of baculovirus proteins, IE1, LEF3, and P143 in the presence of hr. *Virology* **352**, 400–407 (2006).

63. Nagamine, T., Abe, A., Suzuki, T., Dohmae, N. & Matsumoto, S. Co-expression of four baculovirus proteins, IE1, LEF3, P143, and PP31, elicits a cellular chromatin-containing reticulate structure in the nuclei of uninfected cells. *Virology* **417**, 188–195 (2011).

64. Blissard, G. W. & Rohrmann, G. F. Baculovirus Diversity and Molecular Biology. *Annu. Rev. Entomol.* **35**, 127–155 (1990).

65. Passarelli, A. L. Barriers to success: How baculoviruses establish efficient systemic infections. *Virology* vol. 411 383–392 (2011).

66. Flipsen, J. T. M., van Lent, J. W. M., Goldbach, R. W. & Vlak, J. M. Expression of Polyhedrin and p10 in the Midgut of AcMNPV-Infected Spodoptera exigua Larvae: An Immunoelectron Microscopic Investigation. *J. Invertebr. Pathol.* **61**, 17–23 (1993).

67. Harrison, S. C. Viral membrane fusion. *Virology* vols 479–480 498–507 (2015).

68. Dong, S. *et al.* Autographa californica Multicapsid Nucleopolyhedrovirus Efficiently Infects Sf9 Cells and Transduces Mammalian Cells via Direct Fusion with the Plasma Membrane at Low pH. *J. Virol.* **84**, 5351–5359 (2010).

69. Long, G., Pan, X., Kormelink, R. & Vlak, J. M. Functional Entry of Baculovirus into Insect and Mammalian Cells Is Dependent on Clathrin-Mediated Endocytosis. *J. Virol.* **80**, 8830–8833 (2006).

70. Qin, F. *et al.* Dissecting the cell entry pathway of baculovirus by single particle tracking and quantitative electron microscopic analysis Downloaded from. *J. Virol* (2019) doi:10.1128/JVI.00033-19.

71. Shi, Y. *et al.* Three-dimensional visualization of the Autographa californica multiple nucleopolyhedrovirus occlusion-derived virion envelopment process gives new clues as to its mechanism. *Virology* **476**, 298–303 (2015).

72. Kool, M., Voeten, J. T. M., Goldbach, R. W., Tramper, J. & Vlak, J. M. Identification of seven putative origins of Autographa californica multiple nucleocapsid nuclear polyhedrosis virus DNA replication. *J. Gen. Virol.* **74**, 2661–2668 (1993).

73. Leisy, D. J. & Rohrmann, G. F. Characterization of the replication of plasmids containing hr sequences in baculovirus-infected Spodoptera frugiperda cells. *Virology* **196**, 722–730 (1993).

74. Pearson, M., Bjornson, R., Pearson, G. & Rohrmann, G. The Autographa californica baculovirus genome: Evidence for multiple replication origins. *Science (80-.).* **257**, 1382–1384 (1992).

75. Kool, M., Voncken, J. W., van Lier, F. L., Tramper, J. & Vlak, J. M. Detection and analysis of Autographa californica nuclear polyhedrosis virus mutants with defective interfering properties. *Virology* **183**, 739–46 (1991).

76. Kool, M., Van Den Berg, P. M. M. M., Tramper, J., Goldbach, R. W. & Vlak, J. M. Location of Two Putative Origins of DNA Replication of Autographa californica Nuclear Polyhedrosis Virus. *Virology* **192**, 94–101 (1993).

77. Lee, H. Y. & Krell, P. J. Generation and analysis of defective genomes of Autographa californica nuclear polyhedrosis virus. *J. Virol.* **66**, 4339–4347 (1992).

78. Lu, A., Krell, P. J., Vlak, J. M. & Rohrmann, G. F. Baculoviruses DNA replication. in *The Baculoviruses* 171–216 (Plenum Press, New York, 1997).

79. Luque, T., Finch, R., Crook, N., O'Reilly, D. R. & Winstanley, D. The complete sequence of the Cydia pomonella granulovirus genome. *J. Gen. Virol.* **82**, 2531–2547 (2001).

80. Nagamine, T., Kawasaki, Y., Iizuka, T. & Matsumoto, S. Focal Distribution of Baculovirus IE1 Triggered by Its Binding to the hr DNA Elements. *J. Virol.* **79**, 39–46 (2005).

81. Friesen, P. D. Regulation of Baculovirus Early Gene Expression. in *The Baculoviruses* 141–170 (Springer US, 1997). doi:10.1007/978-1-4899-1834-5_6.

82. Guarino, L. A., Gonzalez, M. A. & Summers, M. D. Complete Sequence and Enhancer Function of the Homologous DNA Regions of Autographa californica Nuclear Polyhedrosis Virus. *J. Virol.* **60**, 224–229 (1986).

83. Landais, I. *et al.* Functional analysis of evolutionary conserved clustering of bZIP binding sites in the baculovirus homologous regions (hrs) suggests a cooperativity between host and viral transcription factors. *Virology* **344**, 421–431 (2006).

84. Carstens, E. B. & Wu, Y. No single homologous repeat region is essential for DNA replication of the baculovirus Autographa californica multiple nucleopolyhedrovirus. *J. Gen. Virol.* **88**, 114–122 (2007).

85. Bossert, M. & Carstens, E. B. Sequential deletion of AcMNPV homologous regions leads to reductions in budded virus production and late protein expression. *Virus Res.* (2018) doi:10.1016/J.VIRUSRES.2018.08.010.

86. Kool, M., Goldbach, R. W. & Vlak, J. M. A putative non-hr origin of DNA replication in the HindIII-K fragment of Autographa californica multiple nucleocapsid nuclear polyhedrosis virus. *J. Gen. Virol.* **75**, 3345–3352 (1994).

87. Pearson, M. N., Bjornson, R. M., Ahrens, C. & Rohrmann, G. F. Identification and characterization of a putative origin of DNA replication in the genome of a baculovirus pathogenic for Orgyia pseudotsugata. *Virology* **197**, 715–725 (1993).

88. Ahrens, C. H., Pearson, M. N. & Rohrmann, G. F. Identification and characterization of a second putative origin of DNA replication in a baculovirus of Orgyia pseudotsugata. *Virology* **207**, 572–576 (1995).

89. Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W. & Vlak, J. M. Identification and functional analysis of a non-hr origin of DNA replication in the genome of Spodoptera exigua multicapsid nucleopolyhedrovirus. *J. Gen. Virol.* **78**, 1497–1506 (1997).

90. Pijlman, G. P. *et al.* Pivotal role of the non-hr origin of DNA replication in the genesis of defective interfering baculoviruses. *J. Virol.* **76**, 5605–11 (2002).

91. Pijlman, G. P., Vermeesch, A. M. G. & Vlak, J. M. Cell line-specific accumulation of the baculovirus non-hr origin of DNA replication in infected insect cells. *J. Invertebr. Pathol.* **84**, 214–9 (2003).

92. Huang, J. & Levin, D. B. Identification and functional analysis of a putative non-hr origin of DNA replication from the Spodoptera littoralis type B multinucleocapsid nucleopolyhedrovirus. *J. Gen. Virol.* **80**, 2263–2274 (1999).

93. DePamphilis, M. L. Eukaryotic DNA Replication: Anatomy of An Origin. *Annu. Rev. Biochem.* **62**, 29–63 (1993).

94. Lee, H. & Krell, P. J. Reiterated DNA fragments in defective genomes of Autographa californica nuclear polyhedrosis virus are competent for AcMNPV-dependent DNA replication. *Virology* **202**, 418–429 (1994).

95. Habib, S. & Hasnain, S. E. Differential Activity of Two Non-hrOrigins during Replication of the Baculovirus Autographa californica Nuclear Polyhedrosis Virus Genome. *J. Virol.* **74**, 5182–5189 (2000).

96. Huang, J. & Levin, D. B. Expression, purification and characterization of the Spodoptera littoralis nucleopolyhedrovirus (SpliNPV) DNA polymerase and interaction with the SpliNPV non-hr origin of DNA replication. *J. Gen. Virol.* **82**, 1767–1776 (2001).

97. Pijlman, G. P., van Schijndel, J. E. & Vlak, J. M. Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. *J. Gen. Virol.* **84**, 2669–2678 (2003).

98. Wu, Y.-L. *et al.* Identification of a High-Efficiency Baculovirus DNA Replication Origin That Functions in Insect and Mammalian Cells. *J. Virol.* **88**, 13073–13085 (2014).

99. Miele, S. A. B. *et al.* Identification of Multiple Replication Stages and Origins in the

Nucleopolyhedrovirus of Anticarsia gemmatalis. *Viruses* **11**, (2019).

100. Herniou, E. A., Olszewski, J. A., Cory, J. S. & O'Reilly, D. R. The genome sequence and evolution of baculoviruses. *Annu. Rev. Entomol.* **48**, 211–234 (2003).

101. Crouch, E. A. & Passarelli, A. L. Genetic Requirements for Homologous Recombination in Autographa californica Nucleopolyhedrovirus. *J. Virol.* **76**, 9323–9334 (2002).

102. Oppenheimer, D. I. & Volkman, L. E. Evidence for roiling circle replication of Autographa californica M nucleopolyhedrovirus genomic DNA. *Arch. Virol.* **142**, 2107–2113 (1997).

103. Mikhailov, V. S. Replication of the Baculovirus Genome. *Molecular Biology* vol. 37 250–259 (2003).

104. Wu, Y., Liu, G. & Carstens, E. B. Replication, Integration, and Packaging of Plasmid DNA following Cotransfection with Baculovirus Viral DNA. *J. Virol.* **73**, 5473–5480 (1999).

105. Kool, M., Ahrens, C. H., Vlak, J. M. & Rohrmann, G. F. Replication of baculovirus DNA. *Journal of General Virology* vol. 76 2103–2118 (1995).

106. Hajós, J. P. *et al.* High frequency recombination between homologous baculoviruses in cell culture. *Arch. Virol.* **145**, 159–164 (2000).

107. Lehman, I. R. & Boehmer, P. E. Replication of herpes simplex virus DNA. *Journal of Biological Chemistry* vol. 274 28059–28062 (1999).

108. Summers, M. D. & Smith, G. E. A manual of methods for baculovirus vectors and insect cell culture procedures. *Texas Agric. Exp. Stn. Bull.* **No. 1555**, (1987).

109. Jung, J. G. *et al.* Characterization and Application of Oviductal Epithelial Cells In Vitro in Gallus domesticus1. *Biol. Reprod.* **85**, 798–807 (2011).

110. Nobiron, I., O'Reilly, D. R. & Olszewski, J. A. Autographa californica nucleopolyhedrovirus infection of Spodoptera frugiperda cells: A global analysis of host gene regulation during infection, using a differential display approach. *J. Gen. Virol.* **84**, 3029–3039 (2003).

111. Yang, S. & Miller, L. K. Activation of baculovirus very late promoters by interaction with very late factor 1. *J. Virol.* **73**, 3404 (1999).

112. Merrington, C. L., Kitts, P. A., King, L. A. & Possee, R. D. An Autographa californica nucleopolyhedrovirus lef-2 mutant: Consequences for DNA replication and very late gene expression. *Virology* **217**, 338–348 (1996).

113. Mishra, G., Chadha, P., Chaudhury, I. & Das, R. H. Inhibition of Autographa californica nucleopolyhedrovirus (AcNPV) polyhedrin gene expression by DNAzyme knockout of its serine/threonine kinase (pk1) gene. *Virus Res.* **135**, 197–201 (2008).

114. Rosinski, M., Reid, S. & Nielsen, L. K. Kinetics of baculovirus replication and release using real-time quantitative polymerase chain reaction. *Biotechnol. Bioeng.* **77**, 476–480 (2002).

115. Kelly, B. J., King, L. A. & Possee, R. D. Introduction to baculovirus molecular biology. in *Methods in Molecular Biology* vol. 1350 25–50 (Humana Press Inc., 2016).

116. Assenberg, R., Wan, P. T., Geisse, S. & Mayr, L. M. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology* vol. 23 393–402 (2013).

117. Fernandes, F., Teixeira, A. P., Carinhas, N., Carrondo, M. J. & Alves, P. M. Insect cells as a production platform of complex virus-like particles. *Expert Rev. Vaccines* **12**, 225–236 (2013).

118. Van Oers, M. M., Pijlman, G. P. & Vlak, J. M. Thirty years of baculovirus-insect cell protein expression: From dark horse to mainstream technology. *Journal of General Virology* vol. 96 6–23 (2015).

119. Airenne, K. J. *et al.* Baculovirus: An insect-derived vector for diverse gene transfer applications. *Molecular Therapy* vol. 21 739–749 (2013).

120. Kost, T. A. & Condreay, J. P. Recombinant baculoviruses as mammalian cell gene-delivery vectors. *Trends in Biotechnology* vol. 20 173–180 (2002).

121. Kitts, P. A. Construction of baculovirus recombinants. *Cytotechnology* **20**, 111–123 (1996).

122. Agathos, S. N., Jeong, Y. -H & Venkat, K. Growth kinetics of free and immobilized insect cell cultures. *Ann. N. Y. Acad. Sci.* **589**, 372–398 (1990).

123. Smith, G. E., Fraser, M. J. & Summers, M. D. Molecular engineering of the Autographa californica nuclear polyhedrosis virus genome: Deletion mutations within the polyhedrin gene. *J. Virol.* **46**, 584–593 (1983).

124. Van Rijn, P. A., Van Gennip, H. G. P. & Moormann, R. J. M. An experimental marker vaccine and accompanying serological diagnostic test both based on envelope glycoprotein E2 of classical swine fever virus (CSFV). *Vaccine* **17**, 433–440 (1999).

125. Eustermann, S. *et al.* Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature* **556**, 386–390 (2018).

126. Smith, G. E., Summers, M. D. & Fraser, M. J. *Production of human beta interferon in insect cells infected with a baculovirus expression vector. Molecular and Cellular Biology* http://mcb.asm.org/ (1983).

127. Hartig, P. C. & Cardon, M. C. Rapid efficient production of baculovirus expression vectors. *J. Virol. Methods* **38**, 61–70 (1992).

128. Kitts, P. A., Ayres, M. D. & Possee, R. D. Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Res.* **18**, 5667–5672 (1990).

129. Possee, R. D. *et al.* Nucleotide sequence of the Autographa californica nuclear polyhedrosis 9.4 kbp EcoRI-I and -R (Polyhedrin gene) region. *Virology* **185**, 229–241 (1991).

130. Kitts, P. A. & Possee, R. D. A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* **14**, 810-812+814+816 (1993).

131. Luckow, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. *J. Virol.* **67**, 4566–4579 (1993).

132. Timmis, K., Cabello, F. & Cohen, S. N. Cloning, isolation, and characterization of replication regions of complex plasmid genomes. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 2242–2246 (1975).

133. Lovett, M. A. & Helinski, D. R. Method for the isolation of the replication region of a bacterial replicon: construction of a mini-F'kn plasmid. *J. Bacteriol.* **127**, 982 (1976).

134. Tsutsui, H. & Matsubara, K. Replication Control and Switch-Off Function as Observed with a Mini-F Factor Plasmid. *J. Bacteriol.* **147**, 509 (1981).

135. Lichtenstein, C. & Brenner, S. Site-specific properties of Tn 7 transposition into the E. coli Chromosome. *MGG Mol. Gen. Genet.* **183**, 380–387 (1981).

136. Gay, N. J., Tybulewicz, V. L. J. & Walker, J. E. Insertion of transposon Tn7 into the Escherichia coli glmS transcriptional terminator. *Biochem. J.* **234**, 111–117 (1986).

137. Peters, J. E. & Craig, N. L. Tn7: Smarter than we thought. *Nat. Rev. Mol. Cell Biol.* **2**, 806–814 (2001).

138. Mitra, R., McKenzie, G. J., Yi, L., Lee, C. A. & Craig, N. L. Characterization of the TnsD-attTn7 complex that promotes site-specific insertion of Tn7. *Mob. DNA* **1**, 18 (2010).

139. Jarvis, D. L. Baculovirus–Insect Cell Expression Systems. *Methods Enzymol.* **463**, 191–222 (2009).

140. Irons, S. L., Chambers, A. C., Lissina, O., King, L. A. & Possee, R. D. Protein production using the baculovirus expression system. *Curr. Protoc. Protein Sci.* **91**, 5.5.1-5.5.22 (2018).

141. Possee, R. D. *et al.* Generation of baculovirus vectors for the high-throughput production of proteins in insect cells. *Biotechnol. Bioeng.* **101**, 1115–1122 (2008).

142. Hitchman, R. B., Locanto, E., Possee, R. D. & King, L. A. Optimizing the baculovirus expression vector system. *Methods* vol. *55* 52–57 (2011).

143. Suzuki, T. *et al.* Efficient protein production using a Bombyx mori nuclear polyhedrosis virus lacking the

cysteine proteinase gene. *J. Gen. Virol.* **78**, 3073–3080 (1997).

144. Hitchman, R. B. *et al.* Improved expression of secreted and membrane-targeted proteins in insect cells. *Biotechnol. Appl. Biochem.* **56**, 85–93 (2010).

145. Vlak, J. M. *et al.* Functional studies on the p10 gene of Autographa californica nuclear polyhedrosis virus using a recombinant expressing a p10-beta-galactosidase fusion gene. *J. Gen. Virol.* **69 (Pt 4)**, 765–776 (1988).

146. Chaabihi, H. *et al.* Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. *J. Virol.* **67**, 2664–71 (1993).

147. Kuzio, J., Jaques, R. & Faulkner, P. Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. *Virology* **173**, 759–763 (1989).

148. Rodems, S. M. & Friesen, P. D. The hr5 transcriptional enhancer stimulates early expression from the Autographa californica nuclear polyhedrosis virus genome but is not required for virus replication. *J. Virol.* **67**, 5776–5785 (1993).

149. Hitchman, R. B., Possee, R. D. & King, L. A. High-throughput baculovirus expression in insect cells. *Methods Mol. Biol.* **824**, 609–627 (2012).

150. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22**, 1583–1587 (2004).

151. Pelosse, M. *et al.* MultiBac: from protein complex structures to synthetic viral nanosystems. *BMC Biol.* **15**, 99 (2017).

152. Gupta, K. *et al.* Multibac: Baculovirus-mediated multigene DNA cargo delivery in insect and mammalian cells. *Viruses* vol. 11 (2019).

153. Metcalf, W. W., Jiang, W. & Wanner, B. L. Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6Kλ origin plasmids at different copy numbers. *Gene* **138**, 1–7 (1994).

154. Imasaki, T. *et al.* Architecture of the Mediator head module. *Nature* **475**, 240–243 (2011).

155. Palmberger, D. *et al.* SweetBac: A new approach for the production of mammalianised glycoproteins in insect cells. *PLoS One* **7**, e34226 (2012).

156. Palmberger, D., Klausberger, M., Berger, I. & Grabherr, R. MultiBac turns sweet. *Bioengineered* **4**, 78–83 (2013).

157. Koehler, C. *et al.* Genetic code expansion for multiprotein complex engineering. *Nat. Methods* 1–8 (2016) doi:10.1038/nmeth.4032.

158. Berger, I. & Poterszman, A. Baculovirus expression: old dog, new tricks. *Bioengineered* **6**, 316–322 (2015).

159. Vijayachandran, L. S. *et al.* Gene gymnastics: Synthetic biology for baculovirus expression vector system engineering. *Bioengineered* **4**, 279–287 (2013).

160. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3**, 1021–1032 (2006).

161. Ikonomou, L., Schneider, Y.-J. & Agathos, · S N. Insect cell culture for industrial production of recombinant proteins. *Appl Microbiol Biotechnol* **62**, 1–20 (2003).

162. Lynn, D. E. & Harrison, R. L. Routine maintenance and storage of lepidopteran insect cell lines and baculoviruses. in *Methods in Molecular Biology* vol. 1350 197–221 (Humana Press Inc., 2016).

163. Chan, L. C. L. & Reid, S. Development of serum-free media for lepidopteran insect cell lines. in *Methods in Molecular Biology* vol. 1350 161–196 (Humana Press Inc., 2016).

164. Scholz, J. & Suppmann, S. A new single-step protocol for rapid baculovirus-driven protein production in insect cells. *BMC Biotechnol.* **17**, 83 (2017).

165. Monteiro, F., Carinhas, N., Carrondo, M. J. T., Bernal, V. & Alves, P. M. Toward system-level understanding of baculovirus-host cell interactions: From molecular fundamental studies to large-scale proteomics approaches. *Frontiers in Microbiology* vol. 3 391 (2012).

166. Ho, Y., Lo, H. R., Lee, T. C., Wu, C. P. Y. & Chao, Y. C. Enhancement of correct protein folding in vivo by a non-lytic baculovirus. *Biochem. J.* **382**, 695–702 (2004).

167. Gómez-Sebastián, S., López-Vidal, J. & Escribano, J. M. Significant Productivity Improvement of the Baculovirus Expression Vector System by Engineering a Novel Expression Cassette. *PLoS One* **9**, e96562 (2014).

168. Lin, G., Li, G., Granados, R. R. & Blissard, G. W. Stable cell lines expressing baculovirus P35: Resistance to apoptosis and nutrient stress, and increased glycoprotein secretion. *Vitr. Cell. Dev. Biol. - Anim.* **37**, 293–302 (2001).

169. Szewczyk, B., Hoyos-Carvajal, L., Paluszek, M., Skrzecz, I. & Lobo De Souza, M. Baculoviruses - Re-emerging biopesticides. *Biotechnology Advances* vol. 24 143–160 (2006).

170. Inceoglu, A. B., Kamita, S. G. & Hammock, B. D. Genetically Modified Baculoviruses: A Historical Overview and Future Outlook. *Advances in Virus Research* vol. 68 323–360 (2006).

171. Black, B. C., Brennan, L. A., Dierks, P. M. & Gard, I. E. Commercialization of Baculoviral Insecticides. in *The Baculoviruses* 341–387 (Springer US, 1997). doi:10.1007/978-1-4899-1834-5_13.

172. Stewart, L. M. D. *et al.* Construction of an improved baculovirus insecticide containing an insect-specific toxin gene. *Nature* **352**, 85–88 (1991).

173. Eldridge, R., O'Reilly, D. R., Hammock, B. D. & Miller, L. K. Insecticidal properties of genetically engineered baculoviruses expressing an insect juvenile hormone esterase gene. *Appl. Environ. Microbiol.* **58**, 1583 (1992).

174. O'Reilly, D. R. & Miller, L. K. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. *Science (80-.).* **245**, 1110–1112 (1989).

175. Inceoglu, A. B. *et al.* Recombinant baculoviruses for insect control. *Pest Manag. Sci.* **57**, 981–987 (2001).

176. Heldens, J. G. M. *et al.* Spodoptero exigua multicapsid nucleopolyhedrovirus deletion mutants generated in cell culture lack virulence in vivo. *J. Gen. Virol.* **77**, 3127–3134 (1996).

177. Sun, X. History and current status of development and use of viral insecticides in China. *Viruses* vol. 7 306–319 (2015).

178. Kriz, A. *et al.* A plasmid-based multigene expression system for mammalian cells. *Nat. Commun.* **1**, 1–6 (2010).

179. Haffke, M., Viola, C., Nie, Y. & Berger, I. Tandem recombineering by SLIC cloning and Cre-LoxP fusion to generate multigene expression constructs for protein complex research. *Methods Mol. Biol.* **1073**, 131–140 (2013).

180. Condreay, J. P., Witherspoon, S. M., Clay, W. C. & Kost, T. A. Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 127–132 (1999).

181. Mansouri, M. *et al.* Highly efficient baculovirus-mediated multigene delivery in primary cells. *Nat. Commun.* **7**, 1–13 (2016).

182. Mansouri, M., Ehsaei, Z., Taylor, V. & Berger, P. Baculovirus-based genome editing in primary cells. *Plasmid* **90**, 5–9 (2017).

183. Mansouri, M. & Berger, P. Baculovirus for gene delivery to mammalian cells: Past, present and future. *Plasmid* **98**, 1–7 (2018).

184. Makkonen, K. E., Airenne, K. & Ylä-Herttulala, S. Baculovirus-mediated gene delivery and RNAi applications. *Viruses* vol. 7 2099–2125 (2015).

185. Hindriksen, S. *et al.* Baculoviral delivery of CRISPR/Cas9 facilitates efficient genome editing in human cells. *PLoS One* **12**, e0179514 (2017).

186. Hu, L. *et al.* The major hurdle for effective baculovirus transduction into mammalian cells is passing early endosomes. *J. Virol.* JVI.00709-19 (2019) doi:10.1128/JVI.00709-19.

187. Meyer, C. *et al.* Pseudotyping exosomes for enhanced protein delivery in mammalian cells. *Int. J. Nanomedicine* **12**, 3153–3170 (2017).

188. Ono, C. *et al.* Baculovirus as a Tool for Gene Delivery and Gene Therapy. *Viruses* **10**, 510 (2018).

189. Tramper, J. & Vlak, J. M. Bioreactor design for growth of shear-sensitive mammalian and insect cells. *Adv Biotechnol Process.* **7**, 199–228 (1988).

190. Agathos, S. N. *Insect cell bioreactors. Cytotechnology* vol. 20 (1996).

191. Goosen, M. F. A. Large-scale insect cell culture: methods, applications and products. *Curr. Opin. Biotechnol.* **2**, 365–369 (1991).

192. Caron, A. W., Archambault, J. & Massie, B. High-level recombinant protein production in bioreactors using the baculovirus-insect cell expression system. *Biotechnol. Bioeng.* **36**, 1133–1140 (1990).

193. Tom, R. L., Caron, A. W., Massie, B. & Kamen, A. A. Scale-up of recombinant virus and protein production in stirred-tank reactors. *Methods in molecular biology (Clifton, N.J.)* vol. 39 203–224 (1995).

194. Klaus, D. M. Clinostats and bioreactors. *Gravit Sp. Biol Bull* **14**, 55–64 (2001).

195. Singh, V. Disposable Bioreactor for Cell Culture Using Wave-Induced Agitation. *Cytotechnology* **30**, 149–158 (1999).

196. Chu, L. & Robinson, D. K. Industrial choices for protein production by large-scale cell culture. *Current Opinion in Biotechnology* vol. 12 180–187 (2001).

197. Kompier, R., Tramper, J. & Vlak, J. M. A continuous process for the production of baculovirus using insect-cell cultures. *Biotechnol. Lett.* **10**, 849–854 (1988).

198. van Lier, F. L. J., van den End, E. J., de Gooijer, C. D., Vlak, J. M. & Tramper, J. Continuous production of baculovirus in a cascade of insect-cell reactors. *Appl. Microbiol. Biotechnol.* **33**, 43–47 (1990).

199. Krell, P. J. Passage effect of virus infection in insect cells. in 125–137 (Springer, Dordrecht, 1996). doi:10.1007/0-306-46850-6_10.

200. De Gooijer, C. D. *et al.* A structured dynamic model for the baculovirus infection process in insect-cell reactor configurations. *Biotechnol. Bioeng.* **40**, 537–548 (1992).

201. Scholz, J. & Suppmann, S. A re-usable wave bioreactor for protein production in insect cells. *MethodsX* **3**, 497–501 (2016).

202. Zitzmann, J., Sprick, G., Weidner, T., Schreiber, C. & Czermak, P. Process Optimization for Recombinant Protein Expression in Insect Cells. in *New Insights into Cell Culture Technology* (InTech, 2017). doi:10.5772/67849.

203. Stern, A. & Andino, R. Viral Evolution: It is All About Mutations. in *Viral Pathogenesis: From Basics to Systems Biology: Third Edition* 233–240 (Elsevier Inc., 2016). doi:10.1016/B978-0-12-800964-2.00017-3.

204. Chateigner, A. *et al.* Ultra deep sequencing of a baculovirus population reveals widespread genomic variations. *Viruses* **7**, 3625–3646 (2015).

205. Chen, X. *et al.* The sequence of the Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* **82**, 241–257 (2001).

206. Friesen, P. D. & Nissen, M. S. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol. Cell. Biol.* **10**, 3067–3077 (1990).

207. Muñoz, D., Murillo, R., Krell, P. J., Vlak, J. M. & Caballero, P. Four genotypic variants of a Spodoptera exigua Nucleopolyhedrovirus (Se-SP2) are distinguishable by a hypervariable genomic region. *Virus Res.* **59**, 61–74 (1999).

208. Jehle, J. A. The expansion of a hypervariable, non-hr ori-like region in the genome of Cryptophlebia leucotreta granulovirus provides in vivo evidence for the utilization of baculovirus non-hr oris during

replication. *J. Gen. Virol.* **83**, 2025–2034 (2002).

209. Dai, X. *et al.* Isolation of Spodoptera exigua baculovirus recombinant with a 10·6 kbp genome deletion that retains biological activity. *J. Gen. Virol.* **81**, 2545–2554 (2000).

210. Kumar, S. & Miller, L. K. Effects of serial passage of Autographa Californica nuclear polyhedrosis virus in cell culture. *Virus Res.* **7**, 335–349 (1987).

211. Giri, L., Feiss, M. G., Bonning, B. C. & Murhammer, D. W. Production of baculovirus defective interfering particles during serial passage is delayed by removing transposon target sites in fp25k. *J. Gen. Virol.* **93**, 389 (2012).

212. Carstens, E. B. Identification and nucleotide sequence of the regions of Autographa californica nuclear polyhedrosis virus genome carrying insertion elements derived from Spodoptera frugiperda. *Virology* **161**, 8–17 (1987).

213. Kool, M., Voncken, J. W., van Lier, F. L., Tramper, J. & Vlak, J. M. Detection and analysis of Autographa californica nuclear polyhedrosis virus mutants with defective interfering properties. *Virology* **183**, 739–46 (1991).

214. von Magnus, P. Incomplete Forms of Influenza Virus. *Adv. Virus Res.* **2**, 59–79 (1954).

215. Roux, L., Simon, A. E. & Holland, J. J. Effects of defective interfering viruses on virus replication and pathogenesis in vitro and in vivo. *Adv. Virus Res.* **40**, 181–211 (1991).

216. Bangham, C. M. & Kirkwood, T. B. L. Defective interfering particles: Effects in modulating virus growth and persistence. *Virology* **179**, 821–826 (1990).

217. Wickham, T. J. *et al.* Baculovirus defective interfering particles are responsible for variations in recombinant protein production as a function of multiplicity of infection. *Biotechnol. Lett.* **13**, 483–488 (1991).

218. van Lier, F. L. J. *et al.* Continuous β-galactosidase production with a recombinant baculovirus insect-cell system in bioreactors. *J. Biotechnol.* **22**, 291–298 (1992).

219. Pijlman, G. P. *Molecular mechanisms In baculovlrus defective Interference : keys for Improved recombinant protein expression.* (2003).

220. Pijlman, G. P., van den Born, E., Martens, D. E. & Vlak, J. M. Autographa californica Baculoviruses with Large Genomic Deletions Are Rapidly Generated in Infected Insect Cells. *Virology* **283**, 132–138 (2001).

221. Arends, H. M. & Jehle, J. A. Homologous recombination between the inverted terminal repeats of defective transposon TCp3.2 causes an inversion in the genome of Cydia pomonella granulovirus. *J. Gen. Virol.* **83**, 1573–1578 (2002).

222. Pijlman, G. P., Van Schijndel, J. E. & Vlak, J. M. Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. (2003) doi:10.1099/vir.0.19438-0.

223. Smith, G. A. & Enquist, L. W. Construction and Transposon Mutagenesis inEscherichia coli of a Full-Length Infectious Clone of Pseudorabies Virus, an Alphaherpesvirus. *J. Virol.* **73**, 6405–6414 (1999).

224. Smith, G. A. & Enquist, L. W. A self-recombining bacterial artificial chromosome and its application for analysis of herpesvirus pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4873–4878 (2000).

225. Muyrers, J. P. P., Zhang, Y., Testa, G. & Stewart, A. F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27**, (1999).

226. Noll, S., Hampp, G., Bausbacher, H., Pellegata, N. S. & Kranz, H. Site-directed mutagenesis of multi-copy-number plasmids: Red/ET recombination and unique restriction site elimination. *Biotechniques* **46**, 527–533 (2009).

227. Maier, S. *et al.* Functional characterization of different ORFs including luciferase-like monooxygenase genes from the mensacarcin gene cluster. *ChemBioChem* **16**, 1175–1182 (2015).

228. Ata, E. B. *et al.* Development and characterization of ORF68 negative equine herpes virus type–1, Ab4p

strain. *J. Virol. Methods* **261**, 121–131 (2018).

229. Fujimoto, R., Osakabe, T., Saito, M., Kurosawa, N. & Isobe, M. Minimum length of homology arms required for effective Red/ET recombination. *Biosci. Biotechnol. Biochem.* **73**, 2783–2786 (2009).

230. Sawitzke, J. A. *et al.* Recombineering: In Vivo Genetic Engineering in E. coli, S. enterica, and Beyond. *Methods Enzymol.* **421**, 171–199 (2007).

231. Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. & Copeland, N. G. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36 (2005).

232. Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. Recombineering: A homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206–223 (2009).

233. Sternberg, N. & Hamilton, D. Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *J. Mol. Biol.* **150**, 467–486 (1981).

234. Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5166–5170 (1988).

235. Long, D. P., Tan, B., Zhao, A. C., Xu, L. X. & Xiang, Z. H. Progress in Cre/lox site-specific recombination system in higher eukaryotes. *Yi chuan = Hereditas / Zhongguo yi chuan xue hui bian ji* vol. 34 177–189 (2012).

236. Bradley, A., Su, H. & Wang, X. Nested chromosomal deletions induced with retroviral vectors in mice. *Nat. Genet.* **24**, 92–95 (2000).

237. Wu, S., Ying, G., Wu, Q. & Capecchi, M. R. Toward simpler and faster genome-wide mutagenesis in mice. *Nat. Genet.* **39**, 922–30 (2007).

238. Togawa, Y., Nunoshiba, T. & Hiratsu, K. Cre/lox-based multiple markerless gene disruption in the genome of the extreme thermophile Thermus thermophilus. *Mol. Genet. Genomics* **293**, 277–291 (2018).

239. Kratochwil, C. F. & Rijli, F. M. The cre/lox system to assess the development of the mouse brain. *Methods Mol. Biol.* **1082**, 295–313 (2014).

240. Chandrasegaran, S. & Carroll, D. Origins of Programmable Nucleases for Genome Engineering. *Journal of Molecular Biology* vol. 428 963–989 (2016).

241. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **346**, 1258096 (2014).

242. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* vol. 32 347–350 (2014).

243. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).

244. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology* vol. 20 490–507 (2019).

245. Suenaga, T., Kohyama, M., Hirayasu, K. & Arase, H. Engineering large viral DNA genomes using the CRISPR-Cas9 system. *Microbiol. Immunol.* **58**, 513–522 (2014).

246. Gowripalan, A., Smith, S., Stefanovic, T. & Tscharke, D. C. Rapid poxvirus engineering using CRISPR/Cas9 as a selection tool. *Commun. Biol.* **3**, 1–13 (2020).

247. Pazmiño-Ibarra, V., Mengual-Martí, A., Targovnik, A. M. & Herrero, S. Improvement of baculovirus as protein expression vector and as biopesticide by CRISPR/Cas9 editing. *Biotechnol. Bioeng.* (2019) doi:10.1002/bit.27139.

248. O'Reilly, D. R. & Miller, L. K. Regulation of expression of a baculovirus ecdysteroid UDPglucosyltransferase gene. *J. Virol.* **64**, 1321–8 (1990).

249. Cello, J., Paul, A. V. & Wimmer, E. Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template. *Science (80-.).* **297**, 1016–1018 (2002).

250. Smith, H. O., Hutchison, C. A., Pfannkoch, C. & Venter, J. C. Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15440–15445 (2003).

251. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science (80-.).* **319**, 1215–1220 (2008).

252. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science (80-.).* **329**, 52–56 (2010).

253. Hutchison, C. A. *et al.* Design and synthesis of a minimal bacterial genome. *Science (80-.).* **351**, aad6253–aad6253 (2016).

254. Rees-Garbutt, J. *et al.* Designing minimal genomes using whole-cell models. *Nat. Commun.* **11**, 836 (2020).

255. Annaluru, N. *et al.* Total synthesis of a functional designer eukaryotic chromosome. *Science (80-.).* **344**, 55–58 (2014).

256. Mitchell, L. A. *et al.* Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond. *Science (80-.).* **355**, (2017).

257. Richardson, S. M. *et al.* Design of a synthetic yeast genome. *Science (80-.).* **355**, 1040–1044 (2017).

258. Shen, Y. *et al.* Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome. *Science (80-.).* **355**, (2017).

259. Wu, Y. *et al.* Bug mapping and fitness testing of chemically synthesized chromosome X. *Science (80-.).* **355**, (2017).

260. Xie, Z. X. *et al.* 'Perfect' designer chromosome v and behavior of a ring derivative. *Science (80-.).* **355**, (2017).

261. Shen, Y. *et al.* SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. *Genome Res.* **26**, 36–49 (2016).

262. Liu, W. *et al.* Rapid pathway prototyping and engineering using in vitro and in vivo synthetic genome SCRaMbLE-in methods. *Nat. Commun.* **9**, 1–12 (2018).

263. Neumann, H. *et al.* A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. *Mol. Cell* **36**, 153–163 (2009).

264. Virdee, S. *et al.* Traceless and site-specific ubiquitination of recombinant proteins. *J. Am. Chem. Soc.* **133**, 10708–10711 (2011).

265. Tharp, J. M., Wang, Y. S., Lee, Y. J., Yang, Y. & Liu, W. R. Genetic incorporation of seven ortho-substituted phenylalanine derivatives. *ACS Chem. Biol.* **9**, 884–890 (2014).

266. Coleman, J. R. *et al.* Virus attenuation by genome-scale changes in codon pair bias. *Science (80-.).* **320**, 1784–1787 (2008).

267. Martínez, M. A., Jordan-Paiz, A., Franco, S. & Nevot, M. Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. *Trends in Microbiology* vol. 24 134–147 (2016).

268. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).

269. Isaacs, F. J. *et al.* Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science (80-.).* **333**, 348–353 (2011).

270. Lajoie, M. J. *et al.* Probing the limits of genetic recoding in essential genes. *Science (80-.).* **342**, 361–363 (2013).

271. Napolitano, M. G. *et al.* Emergent rules for codon choice elucidated by editing rare arginine codons in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E5588–E5597 (2016).

272. Lau, Y. H. *et al.* Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA.

Nucleic Acids Res. **45**, 6971–6980 (2017).

273. Chuang, K., Nguyen, E., Sergeev, Y. & Badea, T. C. Novel Heterotypic Rox Sites for Combinatorial Dre Recombination Strategies. *G3* **6**, 559–571 (2016).

274. Anastassiadis, K. *et al.* Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. *Dis. Model. Mech.* **2**, 508–515 (2009).

275. Suzuki, E. & Nakayama, M. VCre/VloxP and SCre/SloxP: new site-specific recombination systems for genome engineering. *Nucleic Acids Res.* **39**, e49–e49 (2011).

276. Karimova, M. *et al.* Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system. *Nucleic Acids Res.* **41**, e37–e37 (2013).

277. McInerney, P., Adams, P. & Hadi, M. Z. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. *Mol. Biol. Int.* **2014**, 287430 (2014).

278. Marx, V. PCR: the price of infidelity. *Nat. Methods* **13**, 475–479 (2016).

279. Bird, A. W. *et al.* high-efficiency counterselection recombineering for site-directed mutagenesis in bacterial artificial chromosomes. *Artic. Nat. methods* | **9**, 103 (2012).

280. Karstentischer, B., von Einem, J., Kaufer, B. & Osterrieder, N. Two-step Red-mediated recombination for versatile high-efficiency markerless DNA manipulation in *Escherichia coli*. *Biotechniques* **40**, 191–197 (2006).

281. Rivero-Mü ller, A., Laji, S. & Huhtaniemi, I. Assisted large fragment insertion by Red/ET-recombination (ALFIRE)—an alternative and enhanced method for large fragment recombineering. *Nucleic Acids Res.* **35**, e78 (2007).

282. Guarino, L. A., Mistretta, T.-A. & Dong, W. DNA binding activity of the baculovirus late expression factor PP31. *Virus Res.* **90**, 187–195 (2002).

283. Gomi, S., Zhou, C. E., Yih, W., Majima, K. & Maeda, S. Deletion Analysis of Four of Eighteen Late Gene Expression Factor Gene Homologues of the Baculovirus, BmNPV. *Virology* **230**, 35–47 (1997).

284. Yamagishi, J., Burnett, E. D., Harwood, S. H. & Blissard, G. W. The AcMNPV pp31 gene is not essential for productive AcMNPV replication or late gene transcription but appears to increase levels of most viral transcripts. *Virology* **365**, 34–47 (2007).

285. Ono, C. *et al.* Phenotypic grouping of 141 BmNPVs lacking viral gene sequences. *Virus Res.* **165**, 197–206 (2012).

286. Harrison, R. L. & Bonning, B. C. Comparative analysis of the genomes of Rachiplusia ou and Autographa californica multiple nucleopolyhedroviruses. *J. Gen. Virol.* **84**, 1827–1842 (2003).

287. Zeng, X. *et al.* Functional analysis of the Autographa californica nucleopolyhedrovirus IAP1 and IAP2. *Sci. China Ser. C Life Sci.* **52**, 761–770 (2009).

288. Passarelli, A. L. & Miller, L. K. In vivo and in vitro analyses of recombinant baculoviruses lacking a functional cg30 gene. *J. Virol.* **68**, 1186–90 (1994).

289. Katsuma, S., Kawaoka, S., Mita, K. & Shimada, T. Genome-wide survey for baculoviral host homologs using the Bombyx genome sequence. *Insect Biochem. Mol. Biol.* **38**, 1080–1086 (2008).

290. Lin, G. & Blissard, G. W. Analysis of an Autographa californica nucleopolyhedrovirus lef-11 knockout: LEF-11 is essential for viral DNA replication. *J. Virol.* **76**, 2770–9 (2002).

291. Guarino, L. A., Mistretta, T.-A. & Dong, W. Baculovirus lef-12 is not required for viral replication. *J. Virol.* **76**, 12032–43 (2002).

292. CHEN, C.-J. & THIEM, S. M. Differential Infectivity of TwoAutographa californicaNucleopolyhedrovirus Mutants on Three Permissive Cell Lines Is the Result oflef-7Deletion. *Virology* **227**, 88–95 (1997).

293. Mitchell, J. K., Byers, N. M. & Friesen, P. D. Baculovirus F-box protein LEF-7 modifies the host DNA damage response to enhance virus multiplication. *J. Virol.* **87**, 12592–9 (2013).

294. REILLY, L. M. & GUARINO, L. A. The Viral Ubiquitin Gene ofAutographa californicaNuclear Polyhedrosis Virus Is Not Essential for Viral Replication. *Virology* **218**, 243–247 (1996).

295. Ge, J. *et al.* AcMNPV ORF38 protein has the activity of ADP-ribose pyrophosphatase and is important for virus replication. *Virology* **361**, 204–211 (2007).

296. Raynes, D. A., Hartshorne, D. J. & Guerriero, V. Sequence and expression of a baculovirus protein with antigenic similarity to telokin. *J. Gen. Virol.* **75**, 1807–1809 (1994).

297. Iwanaga, M., Kurihara, M., Kobayashi, M. & Kang, W. Characterization of Bombyx mori Nucleopolyhedrovirus orf68 Gene That Encodes a Novel Structural Protein of Budded Virus. *Virology* **297**, 39–47 (2002).

298. Krappa, R. & Knebel-Mörsdorf, D. Identification of the very early transcribed baculovirus gene PE-38. *J. Virol.* **65**, 805–12 (1991).

299. Wu, X., Stewart, S. & Theilmann, D. A. Alternative transcriptional initiation as a novel mechanism for regulating expression of a baculovirus trans activator. *J. Virol.* **67**, 5833–42 (1993).

300. Milks, M. L., Washburn, J. O., Willis, L. G., Volkman, L. E. & Theilmann, D. A. Deletion of pe38 attenuates AcMNPV genome replication, budded virus production, and virulence in heliothis virescens. *Virology* **310**, 224–234 (2003).

301. Lu, A. & Miller, L. K. Differential requirements for baculovirus late expression factor genes in two cell lines. *J. Virol.* **69**, 6265–72 (1995).

302. Prikhod'ko, E. A., Lu, A., Wilson, J. A. & Miller, L. K. In vivo and in vitro analysis of baculovirus ie-2 mutants. *J. Virol.* **73**, 2460–8 (1999).

303. Claussen, M., Koch, R., Jin, Z.-Y. & Suter, B. Functional characterization of Drosophila Translin and Trax. *Genetics* **174**, 1337–47 (2006).

304. Crawford, A. M. & Miller, L. K. Characterization of an early gene accelerating expression of late genes of the baculovirus Autographa californica nuclear polyhedrosis virus. *J. Virol.* **62**, 2773–81 (1988).

305. Iwahori, S., Ikeda, M. & Kobayashi, M. Association of Sf9 cell proliferating cell nuclear antigen with the DNA replication site of Autographa californica multicapsid nucleopolyhedrovirus. *J. Gen. Virol.* **85**, 2857–2862 (2004).

306. Detvisitsakun, C., Berretta, M. F., Lehiy, C. & Passarelli, A. L. Stimulation of cell motility by a viral fibroblast growth factor homolog: Proposal for a role in viral pathogenesis. *Virology* **336**, 308–317 (2005).

307. Detvisitsakun, C., Cain, E. L. & Passarelli, A. L. The Autographa californica M nucleopolyhedrovirus fibroblast growth factor accelerates host mortality. *Virology* **365**, 70–78 (2007).

308. Detvisitsakun, C., Hutfless, E. L., Berretta, M. F. & Passarelli, A. L. Analysis of a baculovirus lacking a functional viral fibroblast growth factor homolog. *Virology* **346**, 258–265 (2006).

309. Lu, M. & Iatrou, K. Characterization of a domain of the genome of BmNPV containing a functional gene for a small capsid protein and harboring deletions eliminating three open reading frames that are present in AcNPV. *Gene* **185**, 69–75 (1997).

310. Javed, M. A. *et al.* Autographa californica Multiple Nucleopolyhedrovirus AC83 is a Per Os Infectivity Factor (PIF) Protein Required for Occlusion-Derived Virus (ODV) and Budded Virus Nucleocapsid Assembly as well as Assembly of the PIF Complex in ODV Envelopes. *J. Virol.* **91**, (2017).

311. Boogaard, B., van Oers, M. & van Lent, J. An Advanced View on Baculovirus per Os Infectivity Factors. *Insects* **9**, 84 (2018).

312. Roncarati, R. & Knebel-Mörsdorf, D. Identification of the early actin-rearrangement-inducing factor gene, arif-1, from Autographa californica multicapsid nuclear polyhedrosis virus. *J. Virol.* **71**, 7933–41 (1997).

313. Gearing, K. L. & Possee, R. D. *Functional analysis of a 603 nucleotide open reading frame upstream of the polyhedrin gene of Autographa californica nuclear polyhedrosis virus. Journal of General Virology* vol. 71 www.microbiologyresearch.org (1990).

314. Pijlman, G. P., van Schinjndel, J. E. & Vlak, J. M. Spontaneous excision of BAC vector sequences from

bacmid-derived baculovirus expression vectors upon passage in insect cells. *J. Gen. Virol.* **84**, 2669–2678 (2003).

315. Aulicino, F., Capin, J. & Berger, I. Synthetic virus-derived nanosystems (Svns) for delivery and precision docking of large multifunctional dna circuitry in mammalian cells. *Pharmaceutics* vol. 12 1–20 (2020).

316. Pijlman, G. P. *et al.* Relocation of the attTn7 Transgene Insertion Site in Bacmid DNA Enhances Baculovirus Genome Stability and Recombinant Protein Expression in Insect Cells. *Viruses* **12**, (2020).

317. Bieniossek, C., Imasaki, T., Takagi, Y. & Berger, I. MultiBac: Expanding the research toolbox for multiprotein complexes. *Trends in Biochemical Sciences* vol. 37 49–57 (2012).

318. Stolt-Bergner, P. *et al.* Baculovirus-driven protein expression in insect cells: A benchmarking study. *J. Struct. Biol.* (2018) doi:10.1016/j.jsb.2018.03.004.

319. Guarino, L. A. & Summers, M. D. Nucleotide Sequence and Temporal Expression of a Baculovirus Regulatory Gene. *J. Virol.* **61**, 2091–2099 (1987).

320. Hooft van Iddekinge, B. J. L., Smith, G. E. & Summers, M. D. Nucleotide sequence of the polyhedrin Gene of Autographa californica nuclear polyhedrosis virus. *Virology* **131**, 561–565 (1983).

321. Roelvink, P. W. *et al.* Dissimilar expression of Autographa californica multiple nucleocapsid nuclear polyhedrosis virus polyhedrin and p10 genes. *J. Gen. Virol.* **73**, 1481–1489 (1992).

322. Chaabihi, H. *et al.* Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. *J. Virol.* **67**, 2664–2671 (1993).

323. Hitchman, R. B. *et al.* Genetic modification of a baculovirus vector for increased expression in insect cells. *Recomb. protein expression. Chitinase. Cathepsin Cell Biol Toxicol* **26**, 57–68 (2010).

324. Antonova, S. V. *et al.* Chaperonin CCT checkpoint function in basal transcription factor TFIID assembly. *Nat. Struct. Mol. Biol.* **25**, 1119–1127 (2018).

325. D, S., Shyam Mohan, A. H. & Rao, S. N. Expression and purification of codon-optimized cre recombinase in E. coli. *Protein Expr. Purif.* **167**, 105546 (2020).

326. Cantor, E. J. & Chong, S. Intein-mediated rapid purification of Cre recombinase. *Protein Expr. Purif.* **22**, 135–140 (2001).

327. Wickham, T. J. *et al.* Baculovirus defective interfering particles are responsible for variations in recombinant protein production as a function of multiplicity of infection. *Biotechnol. Lett.* **13**, 483–488 (1991).

328. Pijlman, G. P., Van Den Born, E., Martens, D. E. & Vlak, J. M. Autographa californica baculoviruses with large genomic deletions are rapidly generated in infected insect cells. *Virology* **283**, 132–138 (2001).

329. Berger, I., Fitzgerald, D. J. & Richmond, T. J. Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* **22**, 1583–1587 (2004).

330. Aulicino, F. *et al.* Baculovirus-vectored precision delivery of large DNA cargoes in human genomes. *bioRxiv* 2020.08.17.253898 (2020) doi:10.1101/2020.08.17.253898.

331. Pijlman, G. P. *et al.* Pivotal Role of the Non-hr Origin of DNA Replication in the Genesis of Defective Interfering Baculoviruses. *J. Virol.* **76**, 5605–5611 (2002).

332. Hartman, J. L., Garvik, B. & Hartwell, L. Principles for the Buffering of Genetic Variation. *Princ. Buffering Genet. Var.* **291**, 1001–1004 (2001).

333. Costanzo, M. *et al.* A global genetic interaction network maps a wiring diagram of cellular function. *Science (80-.).* **353**, (2016).

334. Celaj, A. *et al.* Highly Combinatorial Genetic Interaction Analysis Reveals a Multi-Drug Transporter Influence Network. *Cell Syst.* **10**, 25-38.e10 (2020).

335. Taka, H., Ono, C., Sato, M., Asano, S. & Bando, H. Complex genetic interactions among non-essential genes of BmNPV revealed by multiple gene knockout analysis. *J. Insect Biotechnol. Sericology* **82**, 1_025-1_032 (2013).

336. Hitchman, R. B. *et al.* Genetic modification of a baculovirus vector for increased expression in insect cells. *Cell Biol. Toxicol.* **26**, 57–68 (2010).

337. Costa Navarro, G. S. *et al.* The autographa californica multiple nucleopolyhedrovirus Ac12: A non-essential F box-like protein that interacts with cellular SKP1 component of the E3 ubiquitin ligase complex. *Virus Res.* **260**, 67–77 (2019).

338. Fang, Z., Que, Y., Li, J. & Zhang, Z. The deletion of the AcMNPV ac124 gene resulted in a decrease in chitinase transcription. *Virus Res.* **263**, 151–158 (2019).

339. Beperet, I., Simón, O., Williams, T., López-Ferber, M. & Caballero, P. The '11K' gene family members sf68, sf95 and sf138 modulate transmissibility and insecticidal properties of Spodoptera frugiperda multiple nucleopolyhedrovirus. *J. Invertebr. Pathol.* **127**, 101–109 (2015).

340. Leusch, M. S., Lee, S. C. & Olins, P. O. A novel host-vector system for direct selection of recombinant baculoviruses (bacmids) in Escherichia coli. *Gene* **160**, 191–194 (1995).

341. Huang, Z. *et al.* Introduction of temperature-sensitive helper and donor plasmids into Bac-to-Bac baculovirus expression systems. *Virol. Sin.* **30**, 379–385 (2015).

342. Mehalko, J. L. & Esposito, D. Engineering the transposition-based baculovirus expression vector system for higher efficiency protein production from insect cells. *J. Biotechnol.* **238**, 1–8 (2016).

343. Skelding, Z., Queen-Baker, J. & Craig, N. L. Alternative interactions between the Tn7 transposase and the Tn7 target DNA binding protein regulate target immunity and transposition. *EMBO J.* **22**, 5904–5917 (2003).

344. Oppon, J. C., Sarnovsky, R. J., Craig, N. L. & Rawlings, D. E. A Tn7-like transposon is present in the glmUS region of the obligately chemoautolithotrophic bacterium Thiobacillus ferrooxidans. *J. Bacteriol.* **180**, 3007–3012 (1998).

345. Strecker, J. *et al.* RNA-guided DNA insertion with CRISPR-associated transposases. *Science (80-.).* **364**, 48–53 (2019).

346. Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S. & Sternberg, S. H. Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219–225 (2019).

347. Fath-Goodin, A., Kroemer, J., Martin, S., Reeves, K. & Webb, B. A. Polydnavirus Genes that Enhance the Baculovirus Expression Vector System. *Adv. Virus Res.* **68**, 75–90 (2006).

348. Steele, K. H. *et al.* Improving the Baculovirus Expression Vector System with Vankyrin-Enhanced Technology. *Biotechnol. Prog.* (2017) doi:10.1002/btpr.2516.

349. Venkaiah, B., Viswanathan, P., Habib, S. & Hasnain, S. E. An additional copy of the homologous region (hr1) sequence in the Autographa californica multinucleocapsid polyhedrosis virus genome promotes hyperexpression of foreign genes. *Biochemistry* **43**, 8143–8151 (2004).

350. Sung, Y. H., Lee, J. S., Park, S. H., Koo, J. & Lee, G. M. Influence of co-down-regulation of caspase-3 and caspase-7 by siRNAs on sodium butyrate-induced apoptotic cell death of Chinese hamster ovary cells producing thrombopoietin. *Metab. Eng.* **9**, 452–464 (2007).

351. Wang, Q., Zhou, Y., Chen, K. & Ju, X. Suppression of Bm-Caspase-1 Expression in BmN Cells Enhances Recombinant Protein Production in a Baculovirus Expression Vector System. *Mol. Biotechnol.* **58**, 319–327 (2016).

352. Lai, Y. K. *et al.* Enhanced recombinant protein production and differential expression of molecular chaperones in sf-caspase-1-repressed stable cells after baculovirus infection. *BMC Biotechnol.* **12**, 83 (2012).

353. Zhang, X., Xu, K., Ou, Y., Xu, X. & Chen, H. Development of a baculovirus vector carrying a small hairpin RNA for suppression of sf-caspase-1 expression and improvement of recombinant protein production. *BMC Biotechnol.* **18**, 24 (2018).

354. Hay, B. A., Wolff, T., Rubin, G. M. Expression of baculovirus P35 prevents cell death in Drosophila. *Development* **120**, 2121–9 (1994).

355. Seshagiri, S. & Miller, L. K. Baculovirus inhibitors of apoptosis (IAPs) block activation of Sf-caspase-1. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13606–13611 (1997).

356. Lannan, E., Vandergaast, R. & Friesen, P. D. Baculovirus Caspase Inhibitors P49 and P35 Block Virus-Induced Apoptosis Downstream of Effector Caspase DrICE Activation in Drosophila melanogaster Cells. *J. Virol.* **81**, 9319–9330 (2007).

357. Yamada, H., Kitaguchi, K., Hamajima, R., Kobayashi, M. & Ikeda, M. Novel Apoptosis Suppressor Apsup from the Baculovirus Lymantria dispar Multiple Nucleopolyhedrovirus Precludes Apoptosis by Preventing Proteolytic Processing of Initiator Caspase Dronc. *J. Virol.* **87**, 12925–12934 (2013).

358. Rohrmann, G. F. The AcMNPV genome: Gene content, conservation, and function - Baculovirus Molecular Biology - NCBI Bookshelf. in *Baculovirus Molecular Biology* (2019).

359. Li, S., Li, L., Zhao, H. & Liu, W. Disruption of Autographa Californica Multiple Nucleopolyhedrovirus ac111 Results in Reduced per os Infectivity in a Host-Dependent Manner. *Viruses* **10**, (2018).

360. Cohen, D. P. A., Marek, M., Davies, B. G., Vlak, J. M. & Van Oers, M. M. Encyclopedia of Autographa californica Nucleopolyhedrovirus Genes *. *Virol. Sin.* **24**, 359–414 (2009).

361. Gong, M., Jin, J. & Guarino, L. A. Mapping of ORF121, a factor that activates baculovirus early gene expression. *Virology* **244**, 495–503 (1998).

362. Li, Y. & Miller, L. K. Expression and functional analysis of a baculovirus gene encoding a truncated protein kinase homolog. *Virology* **206**, 314–323 (1995).

363. Chen, C. J. & Thiem, S. M. Differential infectivity of two Autographa californica nucleopolyhedorvirus mutants on three permissive cell lines is the result of lef-7 deletion. *Virology* **227**, 88–95 (1997).

364. D'Amico, V. *et al.* Deletion of v-chiA from a baculovirus reduces horizontal transmission in the field. *Appl. Environ. Microbiol.* **79**, 4056–64 (2013).

365. Hawtin, R. E. *et al.* Liquefaction of Autographa californica Nucleopolyhedrovirus-Infected Insects Is Dependent on the Integrity of Virus-Encoded Chitinase and Cathepsin Genes. *Virology* **238**, 243–253 (1997).

366. Hodgson, J. J., Arif, B. M. & Krell, P. J. Interaction of Autographa californica multiple nucleopolyhedrovirus cathepsin protease progenitor (proV-CATH) with insect baculovirus chitinase as a mechanism for proV-CATH cellular retention. *J. Virol.* **85**, 3918–29 (2011).

367. Oomens, A. G. P. & Blissard, G. W. Requirement for GP64 to drive efficient budding of Autographa californica multicapsid nucleopolyhedrovirus. *Virology* **254**, 297–314 (1999).

368. Wolgamot, G. M., Gross, C. H., Russell, R. L. Q. & Rohrmann, G. F. Immunocytochemical characterization of p24, a baculovirus capsid-associated protein. *J. Gen. Virol.* **74**, 103–107 (1993).

369. Gross, C. H., Wolgamot, G. M., Russell, R. L. Q., Pearson, M. N. & Rohrmann, G. F. A baculovirus encoded 16-kDa glycoprotein localizes near the nuclear membrane of infected cells. *Virology* **192**, 386–390 (1993).

370. Zuidema, D., Klinge-Roode, E. C., Van Lent, J. W. M. & Vlak, J. M. Construction and analysis of an Autographa californica nuclear polyhedrosis virus mutant lacking the polyhedral envelope. *Virology* **173**, 98–108 (1989).

371. Peng, Y., Liang, A. & Fu, Y. Function and Application Analysis of Ac132 Protein in Autographa californica Multiple Nucleopolyhedrovirus. *Biotechnol. Bioprocess Eng.* **23**, 655–661 (2018).

372. Li, L. & Rohrmann, G. F. Characterization of a Baculovirus Alkaline Nuclease. *J. Virol.* **74**, 6401–6407 (2000).

373. Gomi, S., Majima, K. & Maeda, S. Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus. *J. Gen. Virol.* **80**, 1323–1337 (1999).

374. Lee, J.-C., Chen, H.-H. & Chao, Y.-C. Persistent Baculovirus Infection Results from Deletion of the Apoptotic Suppressor Gene p35. *J. Virol.* **72**, 9157–9165 (1998).

375. Clem, R. J., Fechheimer, M. & Miller, L. K. Prevention of apoptosis by a baculovirus gene during

infection of insect cells. *Science (80-.).* **254**, 1388–1390 (1991).

376. Clem, R. Baculoviruses and Apoptosis: A Diversity of Genes and Responses. *Curr. Drug Targets* **8**, 1069–1074 (2007).

377. Simón, O., Williams, T., Caballero, P. & Possee, R. D. Effects of Acp26 on in vitro and in vivo productivity, pathogenesis and virulence of Autographa californica multiple nucleopolyhedrovirus. *Virus Res.* **136**, 202–205 (2008).

378. Van Oers, M. M. & Vlak, J. M. The Baculovirus 10-kDa Protein. *Journal of Invertebrate Pathology* vol. 70 1–17 (1997).

379. Faulkner, P., Kuzio, J., Williams, G. V. & Wilson, J. A. Analysis of p74, a PDV envelope protein of Autographa californica nucleopolyhedrovirus required for occlusion body infectivity in vivo. *J. Gen. Virol.* **78**, 3091–3100 (1997).

380. Xi, Q., Wang, J., Deng, R. & Wang, X. Characterization of AcMNPV with a deletion of me53 gene. *Virus Genes* **34**, 223–232 (2007).

381. Schultz, K. L. W., Wetter, J. A., Fiore, D. C. & Friesen, P. D. Transactivator IE1 Is Required for Baculovirus Early Replication Events That Trigger Apoptosis in Permissive and Nonpermissive Cells. *J. Virol.* **83**, 262–272 (2009).

382. Stewart, T. M., Huijskens, I., Willis, L. G. & Theilmann, D. A. The Autographa californica Multiple Nucleopolyhedrovirus ie0-ie1 Gene Complex Is Essential for Wild-Type Virus Replication, but either IE0 or IE1 Can Support Virus Growth. *J. Virol.* **79**, 4619–4629 (2005).

383. Harrison, R. L. & Bonning, B. C. Comparative analysis of the genomes of Rachiplusia ou and Autographa californica multiple nucleopolyhedroviruses. *J. Gen. Virol.* **84**, 1827–1842 (2003).

384. Slack, J. M., Lawrence, S. D., Krell, P. J. & Arif, B. M. Trypsin cleavage of the baculovirus occlusion-derived virus attachment protein P74 is prerequisite in per os infection. *J. Gen. Virol.* **89**, 2388–2397 (2008).

385. Zhang, J.-H., Ohkawa, T., Washburn, J. O. & Volkman, L. E. Effects of Ac150 on virulence and pathogenesis of Autographa californica multiple nucleopolyhedrovirus in noctuid hosts. *J. Gen. Virol.* **86**, 1619–1627 (2005).

386. Ohkawa, T., Rowe, A. R. & Volkman, L. E. Identification of Six Autographa californica Multicapsid Nucleopolyhedrovirus Early Genes That Mediate Nuclear Localization of G-Actin. *J. Virol.* **76**, 12281–12289 (2002).

387. Jiang, S. S. *et al.* Temporal transcription program of recombinant Autographa californica multiple nucleopolyhedrosis virus. *J. Virol.* **80**, 8989–99 (2006).

388. Bai, H., Liang, A. & Fu, Y. Ac154 carried out anti-apoptotic role during AcMNPV infection process in the host insect cells. *Mol. Cell. Biochem.* **463**, 79–90 (2020).

389. Takagi, T., Taylor, G. S., Kusakabe, T., Charbonneau, H. & Buratowski, S. A protein tyrosine phosphatase-like protein from baculovirus has RNA 5′-triphosphatase and diphosphatase activities. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9808–9812 (1998).

390. Bideshi, D. K., Renault, S., Stasiak, K., Federici, B. A. & Bigot, Y. Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. *J. Gen. Virol.* **84**, 2531–2544 (2003).

391. Eldridge, R., Li, Y. & Miller, L. K. Characterization of a baculovirus gene encoding a small conotoxinlike polypeptide. *J. Virol.* **66**, 6563–71 (1992).

392. Kool, M., Ahrens, C. H., Goldbach, R. W., Rohrmann, G. F. & Vlak, J. M. Identification of genes involved in DNA replication of the Autographa californica baculovirus. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11212–11216 (1994).

393. O'Reilly, D. R. & Miller, L. K. Improvement of a Baculovirus Pesticide by Deletion of the EGT Gene. *Bio/Technology* **9**, 1086–1089 (1991).

394. O'Reilly, D. R. & Miller, L. K. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. *Science* **245**, 1110–2 (1989).

395. Burks, J. K., Summers, M. D. & Braunagel, S. C. BV/ODV-E26: A palmitoylated, multifunctional structural protein of Autographa californica nucleopolyhedrovirus. *Virology* **361**, 194–203 (2007).

396. Wang, Y. *et al.* ac18 is not essential for the propagation of Autographa californica multiple nucleopolyhedrovirus. *Virology* **367**, 71–81 (2007).

397. Pearson, M. N., Russell, R. L. Q. & Rohrmann, G. F. Characterization of a Baculovirus-Encoded Protein That Is Associated with Infected-Cell Membranes and Budded Virions. *Virology* **291**, 22–31 (2001).

398. Lung, O. Y., Cruz-Alvarez, M. & Blissard, G. W. Ac23, an envelope fusion protein homolog in the baculovirus Autographa californica multicapsid nucleopolyhedrovirus, is a viral pathogenicity factor. *J. Virol.* **77**, 328–39 (2003).

399. McLachlin, J. R., Yang, S. & Miller, L. K. A baculovirus mutant defective in PKIP, a protein which interacts with a virus-encoded protein kinase. *Virology* **246**, 379–391 (1998).

400. Lauzon, H. A. M., Jamieson, P. B., Krell, P. J. & Arif, B. M. Gene organization and sequencing of the Choristoneura fumiferana defective nucleopolyhedrovirus genome. *J. Gen. Virol.* **86**, 945–961 (2005).

401. Zhang, B., Liang, A. & Fu, Y. Ac34 protein of AcMNPV promoted progeny virus production and induced the apoptosis in host Sf9 cells. *Biotechnol. Lett.* **41**, 147–158 (2019).

402. Cai, Y. *et al.* An ac34 Deletion Mutant of Autographa californica Nucleopolyhedrovirus Exhibits Delayed Late Gene Expression and a Lack of Virulence In Vivo. *J. Virol.* **86**, 10432–10443 (2012).

403. Reilly, L. M. & Guarino, L. A. The viral ubiquitin gene of Autographa californica nuclear polyhedrosis virus is not essential for viral replication. *Virology* **218**, 243–247 (1996).

404. Lin, G. & Blissard, G. W. Analysis of an Autographa californica Nucleopolyhedrovirus lef-11 Knockout: LEF-11 Is Essential for Viral DNA Replication. *J. Virol.* **76**, 2770–2779 (2002).

405. Yu, M. & Carstens, E. B. Characterization of an Autographa californica multiple nucleopolyhedrovirus mutant lacking the ac39(p43) gene. *Virus Res.* **155**, 300–306 (2011).

406. Katsuma, S., Fujii, T., Kawaoka, S. & Shimada, T. Bombyx mori nucleopolyhedrovirus SNF2 global transactivator homologue (Bm33) enhances viral pathogenicity in B. mori larvae. *J. Gen. Virol.* **89**, 3039–3046 (2008).

407. Katsuma, S. & Shimada, T. Bombyx mori nucleopolyhedrovirus ORF34 is required for efficient transcription of late and very late genes. *Virology* **392**, 230–237 (2009).

408. Tao, X. Y. *et al.* Functional Characterization of Autographa californica Multiple Nucleopolyhedrovirus ORF43 and Phenotypic Changes of ORF43-Knockout Mutant. *J. Microbiol.* **51**, 515–521 (2013).

409. Rohrmann, G. F. The AcMNPV genome: Gene content, conservation, and function. (2013).

410. Hong, T., Braunagel, S. C. & Summers, M. D. Transcription, Translation, and Cellular Localization of PDV-E66: A Structural Protein of the PDV Envelope of Autographa californica Nuclear Polyhedrosis Virus. *Virology* **204**, 210–222 (1994).

411. Xiang, X. *et al.* Autographa californica multiple nucleopolyhedrovirus odv-e66 is an essential gene required for oral infectivity. *Virus Res.* **158**, 72–78 (2011).

412. Chen, Y. *et al.* Helicoverpa armaigera nucleopolyhedrovirus ORF50 is an early gene not essential for virus propagation in vitro and in vivo. *Virus Genes* **45**, 149–160 (2012).

413. Guo, Z.-J., Wang, D.-X., Yao, Q., Chen, K.-P. & Zhang, C.-X. Identification of a novel functional nuclear localization signal in the protein encoded by open reading frame 47 of Bombyx mori nucleopolyhedrovirus. *Arch. Virol.* **155**, 1943–1950 (2010).

414. Wang, R. *et al.* Proteomics of the Autographa californica nucleopolyhedrovirus budded virions. *J. Virol.* **84**, 7233–42 (2010).

415. Lu, A. & Miller, L. K. Identification of three late expression factor genes within the 33.8- to 43.4-map-unit

region of Autographa californica nuclear polyhedrosis virus. *J. Virol.* **68**, 6710–8 (1994).

416. Crouch, E. A., Cox, L. T., Morales, K. G. & Passarelli, A. L. Inter-subunit interactions of the Autographa californica M nucleopolyhedrovirus RNA polymerase. *Virology* **367**, 265–274 (2007).

417. Tian, C.-H. *et al.* Bombyx mori nucleopolyhedrovirus ORF51 encodes a budded virus envelope associated protein. *Virus Genes* **38**, 171–177 (2009).

418. Wu, J. & Miller, L. K. Sequence, Transcription and Translation of a Late Gene of the Autographa californica Nuclear Polyhedrosis Virus Encoding a 34.8K Polypeptide. *J. Gen. Virol.* **70**, 2449–2459 (1989).

419. Li, Z. *et al.* Characterization of a chitin-binding protein GP37 of Spodoptera litura multicapsid nucleopolyhedrovirus. *Virus Res.* **96**, 113–122 (2003).

420. Liu, X. *et al.* Synergistic effects of Cydia pomonella granulovirus GP37 on the infectivity of nucleopolyhedroviruses and the lethality of Bacillus thuringiensis. *Arch. Virol.* **156**, 1707–1715 (2011).

421. Vialard, J. E., Yuen, L. & Richardson, C. D. Identification and characterization of a baculovirus occlusion body glycoprotein which resembles spheroidin, an entomopoxvirus protein. *J. Virol.* **64**, 5804–11 (1990).

422. Krell, P. J., Cheng, X.-W. & Arif, B. M. P34.8 (GP37) is not essential for baculovirus replication. *J. Gen. Virol.* **82**, 299–305 (2001).

423. Vanarsdall, A. L., Okano, K. & Rohrmann, G. F. Characterization of the replication of a baculovirus mutant lacking the DNA polymerase gene. *Virology* **331**, 175–180 (2005).

424. Ito, E., Sahri, D., Knippers, R. & Carstens, E. B. Baculovirus proteins IE-1, LEF-3, and P143 interact with DNA in vivo: A formaldehyde cross-linking study. *Virology* **329**, 337–347 (2004).

425. Li, L., Harwood, S. H. & Rohrmann, G. F. Identification of Additional Genes That Influence Baculovirus Late Gene Expression. *Virology* **255**, 9–19 (1999).

426. Wu, X. & Guarino, L. A. Autographa californica nucleopolyhedrovirus orf69 encodes an RNA cap (nucleoside-2'-O)-methyltransferase. *J. Virol.* **77**, 3430–40 (2003).

427. Lu, A. & Miller, L. K. Species-specific effects of the hcf-1 gene on baculovirus virulence. *J. Virol.* **70**, 5123–30 (1996).

428. King, L. A. *et al.* In vitro host range of Autographa californica nucleopolyhedrovirus recombinants lacking functional p35, iap1 or iap2. *J. Gen. Virol.* **80**, 1055–1066 (1999).

429. Zhu, S., Wang, W., Wang, Y., Yuan, M. & Yang, K. The Baculovirus Core Gene ac83 Is Required for Nucleocapsid Assembly and Per Os Infectivity of Autographa californica Nucleopolyhedrovirus. *J. Virol.* **87**, 10573–10586 (2013).

430. Martins, A. & Shuman, S. Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5'-kinase, and polynucleotide 3'-phosphatase activities. *J. Biol. Chem.* **279**, 18220–31 (2004).

431. Croizier, G. *et al.* The pnk/pnl gene (ORF 86) of Autographa californica nucleopolyhedrovirus is a non-essential, immediate early gene. *J. Gen. Virol.* **79**, 629–637 (1998).

432. Thiem, S. M. & Miller, L. K. A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. *J. Virol.* **63**, 4489–97 (1989).

433. Zhang, M.-J. *et al.* Disruption of Bombyx mori nucleopolyhedrovirus ORF71 (Bm71) results in inefficient budded virus production and decreased virulence in host larvae. *Virus Genes* **45**, 161–168 (2012).

434. Thiem, S. M. & Miller, L. K. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus Autographa californica nuclear polyhedrosis virus. *J. Virol.* **63**, 2008–2018 (1989).

435. Wu, W. *et al.* Autographa californica Multiple Nucleopolyhedrovirus 38K Is a Novel Nucleocapsid Protein That Interacts with VP1054, VP39, VP80, and Itself. *J. Virol.* **82**, 12356–12364 (2008).

436. Rohrmann, G. F. Polyhedrin structure. *The Journal of general virology* vol. 67 1499–1513 (1986).

437. Pham, D. Q. D., Hice, R. H., Sivasubramanian, N. & Federici, B. A. The 1629-bp open reading frame of the Autographa californica multinucleocapsid nuclear polyhedrosis virus encodes a virion structural protein. *Gene* **137**, 275–280 (1993).

438. Mishra, G., Chadha, P. & Das, R. H. Serine/threonine kinase (pk-1) is a component of Autographa californica multiple nucleopolyhedrovirus (AcMNPV) very late gene transcription complex and it phosphorylates a 102 kDa polypeptide of the complex. *Virus Res.* **137**, 147–149 (2008).

439. Tao, X. Y. *et al.* Autographa californica Multiple Nucleopolyhedrovirus ORF11 Is Essential for Budded-Virus Production and Occlusion-Derived-Virus Envelopment. *J. Virol.* **89**, 373–383 (2015).

440. Nagamine, T., Inaba, T. & Sako, Y. A nuclear envelop-associated baculovirus protein promotes intranuclear lipid accumulation during infection. *Virology* **532**, 108–117 (2019).

441. Okano, K., Mikhailov, V. S. & Maeda, S. Colocalization of Baculovirus IE-1 and Two DNA-Binding Proteins, DBP and LEF-3, to Viral Replication Factories †. *J. Virol.* **73**, 110–119 (1999).

442. Guo, Y., Liang, A. & Fu, Y. Ac25 in Autographa californica multiple nucleopolyhedrovirus was crucial for progeny budded virion production. *Biotechnol. Lett.* (2019) doi:10.1007/s10529-019-02725-y.

443. Shen, H. *et al.* Characterization of Bombyx mori nucleopolyhedrovirus with a knockout of Bm17. *Cytotechnology* **64**, 711–718 (2012).

444. McLachlin, J. R., Escobar, J. C., Harrelson, J. A., Clem, R. J. & Miller, L. K. Deletions in the Ac-iap1 gene of the baculovirus AcMNPV occur spontaneously during serial passage and confer a cell line-specific replication advantage. *Virus Res.* **81**, 77–91 (2001).

445. Ikeda, M., Yamada, H., Ito, H. & Kobayashi, M. Baculovirus IAP1 induces caspase-dependent apoptosis in insect cells. *J. Gen. Virol.* **92**, 2654–2663 (2011).

446. Passarelli, A. L. & Miller, L. K. Identification and transcriptional regulation of the baculovirus lef-6 gene. *J. Virol.* **68**, 4458–67 (1994).

447. Huang, J. *et al.* Open reading frame Bm21 of Bombyx mori nucleopolyhedrovirus is not essential for virus replication in vitro, but its deletion extends the median survival time of infected larvae. *J. Gen. Virol.* **89**, 922–930 (2008).

448. Shen, Y., Feng, M. & Wu, X. Bombyx mori nucleopolyhedrovirus ORF40 is essential for budded virus production and occlusion-derived virus envelopment. *J. Gen. Virol.* **99**, 837–850 (2018).

449. Tian, C. H. *et al.* Involvement of Bombyx mori nucleopolyhedrovirus ORF41 (Bm41) in BV production and ODV envelopment. *Virology* **387**, 184–192 (2009).

450. Liu, C. *et al.* Autographa californica multiple nucleopolyhedrovirus ac53 plays a role in nucleocapsid assembly. *Virology* **382**, 59–68 (2008).

451. Xu, X. *et al.* Aggregation of AcMNPV LEF-10 and Its Impact on Viral Late Gene Expression. *PLoS One* **11**, e0154835 (2016).

452. Guan, Z. *et al.* The Autographa californica Multiple Nucleopolyhedrovirus ac54 Gene Is Crucial for Localization of the Major Capsid Protein VP39 at the Site of Nucleocapsid Assembly. *J. Virol.* **90**, 4115 (2016).

453. Biswas, S., Willis, L. G., Fang, M., Nie, Y. & Theilmann, D. A. Autographa californica Nucleopolyhedrovirus AC141 (Exon0), a Potential E3 Ubiquitin Ligase, Interacts with Viral Ubiquitin and AC66 To Facilitate Nucleocapsid Egress. *J. Virol.* **92**, (2017).

454. Feng, G. & Krell, P. J. Autographa californica Multiple Nucleopolyhedrovirus DNA Polymerase C Terminus Is Required for Nuclear Localization and Viral DNA Replication. *J. Virol.* **88**, 10918–10933 (2014).

455. Ke, J., Wang, J., Deng, R. & Wang, X. Autographa californica multiple nucleopolyhedrovirus ac66 is required for the efficient egress of nucleocapsids from the nucleus, general synthesis of preoccluded virions and occlusion body formation. *Virology* **374**, 421–431 (2008).

456. Hu, X. *et al.* Bm59 is an early gene, but is unessential for the propagation and assembly of Bombyx mori

nucleopolyhedrovirus. *Mol. Genet. Genomics* **291**, 145–154 (2016).

457. Du, M. F., Yin, X. M., Guo, Z. J. & Zhu, L. J. Characterization of a late gene, ORF60 from Bombyx mori nucleopolyhedrovirus. *J. Biochem. Mol. Biol.* **39**, 737–742 (2006).

458. Guo, Z. J. *et al.* Open reading frame 60 of the Bombyx mori nucleopolyhedrovirus plays a role in budded virus production. *Virus Res.* **151**, 185–191 (2010).

459. Shen, H., Chen, K., Yao, Q. & Zhou, Y. Characterization of the Bm61 of the bombyx mori nucleopolyhedrovirus. *Curr. Microbiol.* **59**, 65–70 (2009).

460. Shi, A. *et al.* Autographa californica Nucleopolyhedrovirus *ac75* is Required for the Nuclear Egress of Nucleocapsids and Intranuclear Microvesicle Formation. *J. Virol.* **92**, JVI.01509-17 (2017).

461. Hu, Z. *et al.* Autographa californica Multiple Nucleopolyhedrovirus ac76 Is Involved in Intranuclear Microvesicle Formation. *J. Virol.* **84**, 7437–7447 (2010).

462. Guo, Y.-J., Fu, S.-H. & Li, L.-L. Autographa californica multiple nucleopolyhedrovirus ac75 is required for egress of nucleocapsids from the nucleus and formation of de novo intranuclear membrane microvesicles. *PLoS One* **12**, e0185630 (2017).

463. McLachlin, J. R. & Miller, L. K. Identification and characterization of vlf-1, a baculovirus gene involved in very late gene expression. *J. Virol.* **68**, 7746–56 (1994).

464. Vanarsdall, A. L., Okano, K. & Rohrmann, G. F. Characterization of the Role of Very Late Expression Factor 1 in Baculovirus Capsid Structure and DNA Processing. *J. Virol.* **80**, 1724–1733 (2006).

465. Garavaglia, M. J., Miele, S. A. B., Iserte, J. A., Belaich, M. N. & Ghiringhelli, P. D. The ac53, ac78, ac101, and ac103 Genes Are Newly Discovered Core Genes in the Family Baculoviridae. *J. Virol.* **86**, 12069–12079 (2012).

466. Chen, L. *et al.* Bombyx mori nucleopolyhedrovirus (BmNPV) Bm64 is required for BV production and per os infection. *Virol. J.* **12**, 173 (2015).

467. Wu, W. & Passarelli, A. L. The Autographa californica M Nucleopolyhedrovirus ac79 Gene Encodes an Early Gene Product with Structural Similarities to UvrC and Intron-Encoded Endonucleases That Is Required for Efficient Budded Virus Production. *J. Virol.* **86**, 5614–5625 (2012).

468. Tang, Q. *et al.* Bm65 is essential for the propagation of bombyx mori nucleopolyhedrovirus. *Curr. Microbiol.* **66**, 22–29 (2013).

469. Chen, H. qing, Chen, K. ping, Yao, Q., Guo, Z. jian & Wang, L. ling. Characterization of a late gene, ORF67 from Bombyx mori nucleopolyhedrovirus. *FEBS Lett.* **581**, 5836–5842 (2007).

470. Gauthier, D., Thirunavukkarasu, K., Faris, B. L., Russell, D. L. & Weaver, R. F. Characterization of an Autographa californica multiple nucleopolyhedrovirus dual mutant: ORF82 is required for budded virus production, and a point mutation in LEF-8 alters late and abolishes very late transcription. *J. Gen. Virol.* **93**, 364–373 (2012).

471. Knebel-Morsdorf, D., Quadt, I., Li, Y., Montier, L. & Guarino, L. A. Expression of Baculovirus Late and Very Late Genes Depends on LEF-4, a Component of the Viral RNA Polymerase Whose Guanyltransferase Function Is Essential. *J. Virol.* **80**, 4168–4173 (2006).

472. Jin, J., Dong, W. & Guarino, L. A. The LEF-4 subunit of baculovirus RNA polymerase has RNA 5'-triphosphatase and ATPase activities. *J. Virol.* **72**, 10011–9 (1998).

473. Wang, D. & Zhang, C.-X. Helicoverpa armigera nucleopolyhedrovirus ORF80 encodes a late, nonstructural protein. *J. Biochem. Mol. Biol.* **40**, 65–71 (2007).

474. Shen, H. *et al.* Characterization of Bombyx mori nucleopolyhedrovirus orf74, a novel gene involved in virulence of virus. *Virus Genes* **38**, 487–494 (2009).

475. Long, C. M., Rohrmann, G. F. & Merrill, G. F. The conserved baculovirus protein p33 (Ac92) is a flavin adenine dinucleotide-linked sulfhydryl oxidase. *Virology* **388**, 231–235 (2009).

476. Wu, W. & Passarelli, A. L. Autographa californica Multiple Nucleopolyhedrovirus Ac92 (ORF92, P33) Is Required for Budded Virus Production and Multiply Enveloped Occlusion-Derived Virus Formation. *J.*

Virol. **84**, 12351–12361 (2010).

477. Nie, Y., Fang, M. & Theilmann, D. A. Autographa californica multiple nucleopolyhedrovirus core gene ac92 (p33) is required for efficient budded virus production. *Virology* **409**, 38–45 (2011).

478. Russell, R. L. Q. & Rohrmann, G. F. A 25-kda protein is associated with the envelopes of occluded baculovirus virions. *Virology* **195**, 532–540 (1993).

479. Luo, X.-C. *et al.* Effects of Early or Overexpression of the Autographa californica Multiple Nucleopolyhedrovirus orf94 (ODV-e25) on Virus Replication. *PLoS One* **8**, e65635 (2013).

480. Lu, A. & Carstens, E. B. Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus Autographa californica nuclear polyhedrosis virus. *Virology* **181**, 336–47 (1991).

481. Wu, Y. & Carstens, E. B. A baculovirus single-stranded DNA binding protein, LEF-3, mediates the nuclear localization of the putative helicase P143. *Virology* **247**, 32–40 (1998).

482. Xu, H. J., Yang, Z. N., Wang, F. & Zhang, C. X. Bombyx mori nucleopolyhedrovirus ORF79 encodes a 28-kDa structural protein of the ODV envelope. *Arch. Virol.* **151**, 681–695 (2006).

483. Dong, Z. Q. *et al.* Bombyx mori nucleopolyhedrovirus ORF79 is a per os infectivity factor associated with the PIF complex. *Virus Res.* **184**, 62–70 (2014).

484. Wu, W. *et al.* Autographa californica Multiple Nucleopolyhedrovirus Nucleocapsid Assembly Is Interrupted upon Deletion of the 38K Gene. *J. Virol.* **80**, 11475–11485 (2006).

485. Harwood, S. H., Li, L., Ho, P. S., Preston, Å. K. & Rohrmann, G. F. AcMNPV late expression factor-5 interacts with itself and contains a Zinc ribbon domain that is required for maximal late transcription activity and is homologous to elongation factor TFIIS. *Virology* **250**, 118–134 (1998).

486. Guarino, L. A., Dong, W. & Jin, J. In Vitro Activity of the Baculovirus Late Expression Factor LEF-5. *J. Virol.* **76**, 12663–12675 (2002).

487. Su, J., Lung, O. & Blissard, G. W. The Autographa californica multiple nucleopolyhedrovirus lef-5 gene is required for productive infection. *Virology* **416**, 54–64 (2011).

488. Peng, Y. *et al.* The protamine-like DNA-binding protein p6.9 epigenetically up-regulates autographa californica multiple nucleopolyhedrovirus gene transcription in the late infection phase. *Virol. Sin.* **27**, 57–68 (2012).

489. Wang, M. *et al.* Specificity of baculovirus P6.9 basic DNA-binding proteins and critical role of the C terminus in virion formation. *J. Virol.* **84**, 8821–8 (2010).

490. Braunagel, S. C., Guidry, P. A., Rosas-Acosta, G., Engelking, L. & Summers, M. D. Identification of BV/ODV-C42, an Autographa californica Nucleopolyhedrovirus orf101-Encoded Structural Protein Detected in Infected-Cell Complexes with ODV-EC27 and p78/83. *J. Virol.* **75**, 12331–12338 (2001).

491. Gandhi, K. M., Ohkawa, T., Welch, M. D. & Volkman, L. E. Nuclear localization of actin requires AC102 in Autographa californica multiple nucleopolyhedrovirus-infected cells. *J. Gen. Virol.* **93**, 1795–803 (2012).

492. Lu, A., Craig, A., Casselman, R. & Carstens, E. B. Nucleotide sequence, insertional mutagenesis, and transcriptional mapping of a conserved region of the baculovirus Autographa californica nuclear polyhedrosis virus (map unit 64.8-66.9). *Can. J. Microbiol.* **42**, 1267–1273 (1996).

493. Yuan, M. *et al.* A highly conserved baculovirus gene p48 (ac103) is essential for BV production and ODV envelopment. *Virology* **379**, 87–96 (2008).

494. Marek, M., van Oers, M. M., Devaraj, F. F., Vlak, J. M. & Merten, O.-W. Engineering of baculovirus vectors for the manufacture of virion-free biopharmaceuticals. *Biotechnol. Bioeng.* **108**, 1056–1067 (2011).

495. Becker, D. & Knebel-Mörsdorf, D. Sequence and temporal appearance of the early transcribed baculovirus gene HE65. *J. Virol.* **67**, 5867–72 (1993).

496. Tang, Q. *et al.* Bm91 is an envelope component of ODV but is dispensable for the propagation of Bombyx mori nucleopolyhedrovirus. *J. Invertebr. Pathol.* **113**, 70–77 (2013).

497. Fang, M., Nie, Y. & Theilmann, D. A. Deletion of the AcMNPV core gene ac109 results in budded virions that are non-infectious. *Virology* **389**, 66–74 (2009).

498. Lin, L. *et al.* ac109 is required for the nucleocapsid assembly of Autographa californica multiple nucleopolyhedrovirus. *Virus Res.* **144**, 130–135 (2009).

499. Alfonso, V. *et al.* AcMNPV Core Gene ac109 Is Required for Budded Virion Transport to the Nucleus and for Occlusion of Viral Progeny. *PLoS One* **7**, (2012).

500. Javed, M. A. *et al.* Autographa californica Multiple Nucleopolyhedrovirus AC83 is a Per Os Infectivity Factor (PIF) Protein Required for Occlusion-Derived Virus (ODV) and Budded Virus Nucleocapsid Assembly as well as Assembly of the PIF Complex in ODV Envelopes. *J. Virol.* **91**, (2017).

501. Yamada, H., Shibuya, M., Kobayashi, M. & Ikeda, M. Identification of a novel apoptosis suppressor gene from the baculovirus Lymantria dispar multicapsid nucleopolyhedrovirus. *J. Virol.* **85**, 5237–42 (2011).

502. Fang, M., Dai, X. & Theilmann, D. A. Autographa californica Multiple Nucleopolyhedrovirus EXON0 (ORF141) Is Required for Efficient Egress of Nucleocapsids from the Nucleus. *J. Virol.* **81**, 9859–9869 (2007).

503. Dai, X., Stewart, T. M., Pathakamuri, J. A., Li, Q. & Theilmann, D. A. Autographa californica Multiple Nucleopolyhedrovirus exon0 (orf141), Which Encodes a RING Finger Protein, Is Required for Efficient Production of Budded Virus. *J. Virol.* **78**, 9633–9644 (2004).

504. McCarthy, C. B., Dai, X., Donly, C. & Theilmann, D. A. Autographa californica multiple nucleopolyhedrovirus ac142, a core gene that is essential for BV production and ODV envelopment. *Virology* **372**, 325–339 (2008).

505. Yang, Z. N., Xu, H. J., Park, E. Y. & Zhang, C. X. Characterization of Bombyx mori nucleopolyhedrovirus with a deletion of bm118. *Virus Res.* **135**, 220–229 (2008).

506. McCarthy, C. B. & Theilmann, D. A. AcMNPV ac143 (odv-e18) is essential for mediating budded virus production and is the 30th baculovirus core gene. *Virology* **375**, 277–291 (2008).

507. Lapointe, R. *et al.* Characterization of Two Autographa californica Nucleopolyhedrovirus Proteins, Ac145 and Ac150, Which Affect Oral Infectivity in a Host-Dependent Manner. *J. Virol.* **78**, 6439–6448 (2004).

508. Dickison, V. L., Willis, L. G., Sokal, N. R. & Theilmann, D. A. Deletion of AcMNPV ac146 eliminates the production of budded virus. *Virology* **431**, 29–39 (2012).

509. Ge, J. Q., Zhao, J. F., Shao, Y. M., Tian, C. H. & Zhang, C. X. Characterization of an early gene orf122 from Bombyx mori nucleopolyhedrovirus. *Mol. Biol. Rep.* **36**, 543–548 (2009).

Chapter 8: Appendix

8.1 List of plasmids

Table 8.1 –List of used plasmids in this thesis summarizing important details, antibiotic resistance, and source.

Name	Details	Resistance	Source
pDS	Donor, R6Kγ origin	Spec	Berger Group
pACEBac1	Acceptor, Tn7 transposition sequences, polh promoter, LoxP site	Gent	Berger Group
pUCDM	Donor, R6Kγ origin, polh and p10 promoter	Chlor	Berger Group
pUCDM_YFP	Polh driven EYFP	Chlor	Berger Group
pFB_GDAP288	Polh driven GDAP288	Amp	Berger Group
pRedET	Expression of proteins required for HR, temperature sensitive	Tet	F.Stewart
pDS-Frag4A	Cassette for Fragment 4 native DNA deletion by HR	Spec	D.Balaji
pDS-Frag4	Cassette for insertion of rewired Fragment 4 by HR	Spec	D.Balaji
pDS-Frag4-RoxAmpRox	Cassette for insertion of rewired Fragment 4 by HR, RoxP sites	Spec	This study
pDS-Frag3	Cassette for Fragment 3 native DNA deletion by HR	Spec	D.Balaji
pDS-Frag3	Cassette for insertion of rewired Fragment 3 by HR	Spec	D.Balaji
pDS-Frag3VoxAmpVox	Cassette for insertion of rewired Fragment 3 by HR, Vox sites	Spec	GenScript
pUADM_RoxP	Donor, RoxP site	Amp	This study
pUADM_RoxP_AmCyan	Donor, RoxP site, polh driven AmCyan	Amp	This study
pUBDM_SLoxP	Donor, SLoxP site	Blas	H.Crocker
pUBDM_SLoxP_mTagBFP	Donor, SLoxP site, polh driven mTagBFP	Blas	This study
pUPDM_VLoxP	Donor, VloxP site	Puro	Z.Guo
pUPDM_VLoxP_EGFP	Donor, VloxP site, polh driven EGFP	Puro	Z.Guo
pUZDM_Vox	Donor, Vox site	Zeo	Z.Guo
pUZDM_Vox_mCherry	Donor, Vox site, polh driven mCherry	Zeo	This study
pACEBac1_RoxP	Acceptor, RoxP site	Gent	GenScript
pACEBac1_VloxP	Acceptor, VloxP site	Gent	GenScript

pACEBac1_SLoxP	Acceptor, SLoxP site	Gent	GenScript
pACEBac1_Vox	Acceptor, Vox site	Gent	GenScript
pACEBac1_RoxP x pUADM_RoxP_AmCyan	Acceptor-donor fusion	Gent, Amp	This study
pACEBac1_VloxP x pUPDM_VloxP_EGFP	Acceptor-donor fusion	Gent, Puro	This study
pACEBac1_SLoxP x pUBDM_SLoxP_mTagBFP	Acceptor-donor fusion	Gent, Blas	This study
pACEBac1_Vox x pUZDM_Vox_mCherry	Acceptor-donor fusion	Gent, Zeo	This study
pET28a-Dre-6xHis	Dre expression in BL21 start	Kan	GenScript
pET28a-VCre-6xHis	VCre expression in BL21 start	Kan	GenScript
pET28a-SCre-6xHis	SCre expression in BL21 start	Kan	GenScript
pET28a-Vika-6xHis	Vika expression in BL21 start	Kan	GenScript
pKD46	Expression of proteins required for HR, temperature sensitive	Amp	AstraZeneca
PL-17-0352	BSD cassette	Gent	AstraZeneca
pCPSS2185	Polh driven AmCyan	Gent	AstraZeneca
pCPSS2190	Polh driven PPAR	Gent	AstraZeneca
pCPSS2191	Polh driven Keap1	Gent	AstraZeneca
pCPSS2187	Polh driven Lck	Gent	AstraZeneca
pCPSS1728	Polh driven hPI3Kg	Gent	AstraZeneca
pDS_ac20VloxP	Cassette for VloxP insertion at ac20	Spec	GenScript
pDS_ac87RoxP	Cassette for RoxP insertion at ac87	Spec	GenScript
pDS_ac32RoxP_Tn7	Cassette for Tn7 insertion at ac32	Spec	GenScript
pDS_ac87RoxP_Tn7	Cassette for Tn7 insertion at ac87	Spec	GenScript
pDS_ac110RoxP_Tn7	Cassette for Tn7 insertion at ac110	Spec	GenScript
pDS_ac112/113RoxP_Tn7	Cassette for Tn7 insertion at ac112/113	Spec	GenScript
pUCDM_YFP_polh	Polh driven EYFP, p10 promoter removed	Chlor	This study
pUPDM_VloxP_mCherry	Donor, VloxP site, polh driven mCherry	Puro	This study
3xColour (named Dual colour plasmid in this thesis)	pKL_NLSmTagBFP x pUCDM_SNAPTag_GP64 x pSPL_mCherry	Kan, Chlor, Spec	H.Crocker
pACEBac1 polH TAF5 p10 TAF6 polH TAF9 p10 mCherry	TAF569	Gent	F.Aulicino

8.2 Sequences of interest

8.2.1 Chapter 3

8.2.1.1 pDS

ataacttcgtataatgtatgctatacgaagttatggtacgtactaagctctcatgtttcacgtactaagctctcatgtttaacgtactaagctctcatgtttaacgaactaaaccctcatggctaacgtactaagc
tctcatggctaacgtactaagctctcatgtttcacgtactaagctctcatgtttgaacaataaaattaatataaatcagcaacttaaatagcctctaaggttttaagttttataagaaaaaaaagaatatataag
gcttttaaagcttttaaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagcccttagagcctctcaaagcaattttgagtgacacaggaacacttaacggctgacataattcagctt
cacgctgccgcaagcactcagggcgcaagggctgctaaaggaagcggaacacgtagaaagccagtccgcagaaacggtgctgaccccggatgaatgtcagctgggaggcagaataaatgatcatat
cgtcaattattacctccacggggagagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaagtagcgtatgcgctcacgcaactggtccag
aaccttgaccgaacgcagcggtggtaacggcgcagtggcggttttcatggcttgttatgactgtttttttggggtacagtctatgcctcgggcatccaagcagcaagcgcgttacgccgtgggtcgatgtttga
tgttatggagcagcaacgatgttacgcagcagggcagtcgccctaaaacaaagttaaacatcatgagggaagcggtgatcgccgaagtatcgactcaactatcagaggtagttggcgtcatcgagcgcc
atctcgaaccgacgttgctggccgtacatttgtacggctccgcagtggatggcggcctgaagccacacagtgatattgatttgctggttacggtgaccgtaaggcttgatgaaacaacgcggcgagctttga
tcaacgaccttttggaaacttcggcttcccctggagagagcgagattctccgcgctgtagaagtcaccattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttggag
aatggcagcgcaatgacattcttgcaggtatcttcgagccagccacgatcgacattgatctggctatcttgctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcggaggaactcttt
gatccggttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccgcccgactgggctggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtacagcgcagta
accggcaaaatcgcgccgaaggatgtcgctgccgactgggcaatggagcgcctgccggcccagtatcagcccgtcatacttgaagctagacaggcttatcttggacaagaagaagatcgcttggcctc
gcgcgcagatcagttggaagaatttgtccactacgtgaaaggcgagatcaccaaggtagtcggcaaataatgtctaacaattcgttcaagccgacggatctatgtcgggtgcggagaaagaggtaatgaa
atggcacctaggtatcgatggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgaatttctagaaataattttgtttaa
ctttaagaaggagatatacatatgaggcctcggatcctgtaaaacgacggccagtgaattccccgggaagcttcgccagggttttcccagtcgagctcgatatcggtaccagcggataacaatttcacatc
cggatcgcgaacgcgtctcgagactagttccgtttaaacccatgtgcctggcag

8.2.1.2 pACEBac1 (Change LoxP site for RoxP, VloxP, SloxP or Vox variants)

ataacttcgtatagcatacattatacgaagttattatctgtaactataacggtcctaaggtagcgagtttaaacactagtatcgattcgcgacctactccggaatattaatagatcatggagataattaaaatg
ataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatcccggtccgaagcgcgcggaatt
caaaggcctacgtcgacgagctcacttgtcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctt
taaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcattt
ttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccac
cgctgagcaataactatcataacccctagggtatacccatctaattggaaccagataagtgaaatctagttccaaactattttgtcatttttaattttcgtattagcttacgacgctacacccagttcccatctat
tttgtcactcttccctaaataatccttaaaaactccatttccacccctcccagttcccaactattttgtccgcccacaaccggttgacttgggtcaactgtcagaccaagtttactcatatatactttagattgatt
taaaacttcatttttaatttaaaaggatctaggtgaagatcctttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgag
atcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatacc
aaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggt
tggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaga
aagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttc
gccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgtt
atcccctgattgacttgggtcgctcttcctgtggatgcgcagatgccctgcgtaagcgggtgtgggcggacaataaagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactagac
agaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctgaagtgcaaattgcccgtcgtattaaagaggggcgtggccaaggg
catgtaaagactatattcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgt
atgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcct
ccggtgctcgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacttgggcagaacgtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatg
aatgtcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttggtcagggc
cgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacg
gcgtaacgcgcttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgt
gagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgc

8.2.1.3 pUADM_RoxP_AmCyan

taactttaaataatgccaattatttaaagttaggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttggggggttggggtggggaaa
aggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttatt
gccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtcttcgt
cgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgcgacc
tactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtccca
ccatcgggcgcggatcccggtccgaagcgcgcggaattcaaaggggatccatggcactgagcaacaaattcattggagacgatatgaaaatgacttaccacatggacggctgcgttaacggtcactact
tcactgtcaagggagagggcagtggaaaaccttacgaaggtacacagacgtcgaccttcaaggtgacaatggctaacggtggccccttggccttctctttcgacatcctgtcaaccgtcttcatgtacggc
aacaggtgcttcacggcttaccctacctctatgcccgactacttcaagcaagccttccctgatggaatgtcatacgagagaactttcacatacgaagatggtggtgtggctactgcctcctgggagatcagc
ctgaagggtaactgtttcgaacacaaaagtaccttccacggagttaacttcccagcagacggtcctgtgatggcgaagaaaaccactggctgggaccctagcttcgagaagatgactgtctgcgacggta
ttctcaaaggcgatgttacagctttcctgatgctccagggcggtggtaactaccgctgtcaattccacacttcctacaagacaaagaaaccggtcacgatgcctcccaaccacgtggtcgagcatcgcat
cgcacgtactgacctcgataaaggcggaaactccgtgcagttgacggaacacgcagtcgcgcatattaccagcgttgtgccattcgacccggctttcctgtacaaggtcgttaggttcgaacaccatcac
catcaccattaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctg
aaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactc
atcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactatcataacccctagggg

cctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggatctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaacaccatcatacactaaatcagt
aagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgataccgggaagccctgggccaacttttggcgaaaat
gagacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaaaatgagtattcaacatttccgtgtc
gcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaag
atccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaa
tgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggag
gaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatgg
caacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggttta
ttgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaat
agacagatcgctgagataggtgcctcactgattaagcattggtaataattttttttaaggcagttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcggatgaatggcagaaattc
gaaagcaaattcgacccggtcgtcggttcagggcagggtcgttaaatagccgcttatgtctattgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctgaggccagtt
tgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctcccagctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagcagcccttg
cgccctgagtgcttgcggcagcgtgaagctaattctgtcagccgttaagtgttcctgtgtcactcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccacaaccgtta
aaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaaccttagaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatgagagcttag
tacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacatgagagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaacaggttgaact
gctgatcaacagatcctctacgcggccgcggtaccgatatc

8.2.1.4 pUBDM_SloxP_mTagBFP

ctcgtgtccgataactgtaattatcggacatgatctggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttggggggttggggtggg
gaaaaggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctt
tttattgccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtc
ttcgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgc
gacctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgt
cccaccatcgggcgcggatcccggtccgaagcgcgcggaattcaaaggggatccatgagcgagctgattaaggagaacatgcacatgaagctgtacatggagggcaccgtggacaaccatcacttca
agtgcacatccgagggcgaaggcaagccctacgagggcacccagaccatgagaatcaaggtggtcgagggcggccctctccccttcgccttcgacatcctggctactagcttcctctacggcagcaag
accttcatcaaccacacccagggcatccccgacttcttcaagcagtccttccctgagggcttcacatgggagagagtcaccacatacgaagacgggggcgtgctgaccgctacccaggacaccagcc
tccaggacggctgcctcatctacaacgtcaagatcagaggggtgaacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacgctgtaccccgctgacgg
cggcctggaaggcagaaacgacatggccctgaagctcgtgggcgggagccatctgatcgcaaacatcaagaccacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactat
gtggactacagactggaaagaatcaaggaggccaacaacgagacctacgtcgagcagcacgaggtggcagtggccagatactgcgacctccctagcaaactggggcacaagcttaatggatcctagt
ctagacctacgtcgacgagctcactagtcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgcttt
aaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttt
tttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccacc
gctgagcaataactatcataacccctaggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggatctgaacaggagggacagctgatagaaacagaagccactggagca
cctcaaaaacaccatcatacactaaatcagtaagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgatacc
gggaagccctgggccaacttttggcgaaaatgagacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaagg
aagctaaaatggccaagcctttgtctcaagaagaatccaccctcattgaaagagcaacggctacaatcaacagcatccccatctctgaagactacagcgtcgccagcgcagctctctctagcgacggc
cgcatcttcactggtgtcaatgtatatcattttactgggggaccttgtgcagaactcgtggtgctgggcactgctgctgctgcggcagctggcaacctgacttgtatcgtcgcgatcggaaatgagaacaggg
gcatcttgagcccctgcggacggtgtcgacaggtgcttctcgatctgcatcctgggatcaaagcgatagtgaaggacagtgatggacagccgacggcagttgggattcgtgaattgctgccctctggttatg
tgtgggagggctaattttttttaaggcagttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcggatgaatggcagaaattcgaaagcaaattcgacccggtcgtcggttcaggg
cagggtcgttaaatagccgcttatgtctattgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctgaggccagtttgctcaggctctccccgtggaggtaataattgac
gatatgatcatttattctgcctcccagctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagcagcccttgcgccctgagtgcttgcggcagcgtgaagctaatt
ctgtcagccgttaagtgttcctgtgtcactcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccacaaccgttaaaccttaaaagctttaaaagccttatatattcttttt
tttcttataaaacttaaaaccttagaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatgagagcttagtacgttagccatgagagcttagtacgttagccatg
agggtttagttcgttaaacatgagagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggta
ccgatatc

8.2.1.5 pUPDM_VloxP_EGFP

tcaatttctgagaactgtcattctcggaaattgactggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttggggggttggggtgggg
aaaaggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttt
tattgccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtctt
cgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgcg
acctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtc
ccaccatcgggcgcggatccatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgaggg
cgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccac
atgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctgg
tgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaacggcatca
aggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacc
cagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtccggactcagatctcgatagtc
tagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaaacctcccacacctccccctgaacctgaaacataaaatga
atgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttat
catgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactatcataacccctaggggcctctaaacgggtc
ttgaggggttttttgctgaaaggaggaactatatccggatctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaacaccatcatacactaaatcagtaagttggcagcatc
acccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgataccgggaagccctgggccaacttttggcgaaaatgagacgttgatcgg
cacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaaaatgaccgagtacaagcccacggtgcgcctcgccac
ccgcgacgacgtccccagggccgtacgcaccctcgccgccgcgttcgccgactaccccgccacgcgccacaccgtcgatccggaccgccacatcgagcgggtcaccgagctgcaagaactcttcc

tcacgcgcgtcgggctcgacatcggcaaggtgtgggtcgcggacgacggcgccgcggtggcggtctggaccacgccggagagcgtcgaagcgggggcggtgttcgccgagatcggcccgcgcatgg
ccgagttgagcggttcccggctggccgcgcagcaacagatggaaggcctcctggcgccgcaccggcccaaggagcccgcgtggttcctggccaccgtcggcgtctcgcccgaccaccagggcaagg
gtctgggcagcgccgtcgtgctccccggagtggaggcggccgagcgcgccggggtgcccgccttcctggagacctccgcgccccgcaacctccccttctacgagcggctcggcttcaccgtcaccgc
cgacgtcgagtgcccgaaggaccgcgcgacctggtgcatgacccgcaagcccggtgcctgattttttaaggcagttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcgga
tgaatggcagaaattcgaaagcaaattcgacccggtcgtcggttcagggcagggtcgttaaatagccgcttatgtctattgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaa
atgcctgaggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctcccagctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgctt
cctttagcagcccttgcgccctgagtgcttgcggcagcgtgaagctaattcccatgtcagccgttaagtgttcctgtgtcactcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctgg
cttgttgtccacaaccgttaaaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaaccttagaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtac
gtgaaacatgagagcttagtacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacatgagagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacg
tactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggtaccgatatc

8.2.1.6 pUZDM_Vox_mCherry

aataggtctgagaacgcccattctcagacgtattctggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttggggggttggggtggg
gaaaaggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctt
tttattgccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtc
ttcgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgc
gacctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgt
cccaccatcgggcgcggatccatggtgagcaagggcgaggaggataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccacgagttcgagatcgagg
gcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggtggcccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggcctacgt
gaagcaccccgccgacatccccgactacttgaagctgtccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccgtgacccaggactcctccctgcaggacg
gcgagttcatctacaaggtgaagctgcgcggcaccaacttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctcctccgagcggatgtaccccgaggacggcgccctgaa
gggcgagatcaagcagaggctgaagctgaaggacggcggccactacgacgctgaggtcaagaccacctacaaggccaagaagcccgtgcagctgcccggcgcctacaacgtcaacatcaagttgg
acatcacctcccacaacgaggactacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgagctgtacaagttctagagcctgcagtctcgacaagcttgtcg
agaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcag
cttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagccta
gaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactatcataacccctaggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaacta
tatccggatctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaacaccatcatacactaaatcagtaagttggcagcatcacccgacgcactttgcgccgaataaatac
ctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgataccgggaagccctgggccaacttttggcgaaaatgagacgttgatcggcacgtaagaggttccaactttcaccataat
gaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaaaatggccaagttgaccagtgccgttccggtgctcaccgcgcgcgacgtcgccggagcggtcgagtt
ctggaccgaccggctcgggttctcccgggacttcgtggaggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggaccaggtggtgccggacaacaccctggcctg
ggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaacttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgcc
ctgcgcgacccggccggcaactgcgtgcacttcgtggccgaggagcaggactgattttttaaggcagttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcggatgaatggc
agaaattcgaaagcaaattcgacccggtcgtcggttcagggcagggtcgttaaatagccgcttatgtctattgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctga
ggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctcccagctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagc
agcccttgcgccctgagtgcttgcggcagcgtgaagctaattctgtcagccgttaagtgttcctgtgtcactcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccac
aaccgttaaaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaaccttagaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatga
gagcttagtacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacatgagagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaaca
ggttgaactgctgatcaacagatcctctacgcggccgcggtaccgatatc

8.2.1.7 Dre recombinase coding sequence

atgggcagcagccaccatcaccaccaccatcatcaccatcatagtagtgaaaacctgtattttcaggggagtagtgaattaattatctctggcagttccggtggtttctgcgcaacattggcaaagagtac
caggaagccgcagaaaactttatgcggttcatgaatgatcagggtgcttatgcgccgaatactctgcgcgacctccggttggtgtttcattcctgggcgcgctggtgtcacgctcgccagcttgcctggtttc
cgatctcaccagaaatggcccgcgagtattttcttcagttgcatgatgccgatctggcttcgaccaccattgataagcactacgccatgcttaacatgctgctttctcattgtggccttccgccactttcggatg
ataaaagtgtttctctggctatgcggcgtatccggcgcgaagcggcaacggaaaaaggcgaacgaacaggccaggctataccgctgcgatgggacgacctgaaactgctcgacgtcctgttgtccaggt
cagaacggctggtggacctgcgcaaccgggcttttctctttgttgcatacaatacgctgatgcgtatgtcggaaatctcgcgtattcgtgtaggagatctggaccaaacaggtgacactgtcacgctacatat
ttcacacacgaagacaataacgaccgccgccgggcttgataaggtgctttcccgtcgcactactgctgtgctgaatgactggctggatgtttccgggcttcgtgaacatcctgacgcggtgctgttcccgcc
gatacaccgtagcaataaggccaggattacgacaacgccccttactgcacctgcaatggagaaaatattcagcgacgcctgggtgttgctgaataaaagggatgccacgccaaacaaagggagatac
cggacgtggaccgggcatagtgctcgtgtcggggccgctatcgatatggctgaaaagcaggtgtctatggtggagattatgcaggaaggaacctggaagaaaccggaaacacttatgcggtatctgcgc
cgtggcggtgtgtctgtgggggctaatagccgcttgatggattcataa

8.2.1.8 VCre recombinase coding sequence

atgggatccatgagtcaccatcaccaccaccatcatcaccatcatagtagtgaaaacctgtattttcaggggagtattgaaaaccagttaagtttgttgggagacttcagtggcgttcgcccggatgatgtca
agactgccattcaagcagcacagaaaaaagggattaatgtagccgaaaacgagcagttcaaagccgcttttgaacacctgttaaacgagttcaagaagcgcgaggagcgttacagtcccaatacctta
cgtcgtctggagtccgcttggacatgcttcgttgactggtgtcttgcgaatcatcgccattcgttaccggcaactcccgacacggtcgaagcatttttcatcgaacgtgcggaagaattacatcgtaatacctt
gagtgtttatcgttgggcaatttcgcgcgtgcatcgtgtcgctgggtgtccagatccctgtctggacatctatgttgaggaccgtcttaaggcgatcgcgcgtaaaaaggtacgtgaaggagaagctgtaaag
caggcttcaccgtttaatgagcaacacctgttaaaacttacaagtttgtggtaccgctcggacaaattacttcttcgccgcaacctggcgttgttagcagtagcgtatgagagtatgttacgcgcaagcgagt
tggcgaatattcgtgtttctgacatggagttggcaggtgacggtaccgccattttgacgattccgattacaaagaccaaccatagcggggagcctgatacgtgcatcttgagccaggacgttgtgtctcttctg
atggactatacggaagcgggcaagttggacatgtcttccgatgggtttttgtttgtaggggtgtcgaagcataatacgtgcattaagcccaaaaaagacaagcagaccggcgaggtacttcacaagccaat
cacaacaaagactgttgaaggggtcttctattccgcgtgggagacacttgaccttggccgtcagggcgtgaaaccgttcaccgcacacagcgctcgtgtcggcgccgcacaagacttgttaaagaaggg
ttataatacattacagattcaacagtccggtcgttggagttctggagcaatggttgcccgctacggccgcgcaatcttagcgcgcgatggcgctatggctcactctcgcgttaaaacacgttctgctccgatg
caatggggaaaggacgaaaaagattaa

8.2.1.9 SCre recombinase coding sequence

atgggatccatgagtcaccatcaccaccaccatcatcaccatcatagtagtgaaaacctgtattttcaggggagtatggtgccaaaaaaaaaacgcaaagactataaagacgacgacgataaaagtctt
ttaacaacaaataatcatagtgtagctctgagttatggggaaccgccatctacgctgaacgacagtttgaaagattcatatcagcgttctaccgacgagttgcaagcactgttgtcaaagccattagcccag
cttaccgatgctgataagctgcgtatccgcgaaatcacccaagcaaaactgaaacattttctggacaacggtcatcgcactcgccgcgcaaatacgtggcgtgcccttatgagccgctgggcaaaatttg
aatcatggtgtttaaccaacaacttaacaccgcttccggctacacctgaggtggtggctaccttcatcgagtactatcaagcgtccagctacactaccttgtcccaatacgcttgggcgatcaactcatttca
tgtcgaatgcggtttgctgagtccagtgtcctccaagaccgtgcaggataaacagaacgaaatccgtatcgtaaaacttgagagcggcgggcttgcccaggaacaagcgactccatttcgccttcaccac
ttgcagatgttaatcgagtcttacggggaaagcgaacgtttattagataagcgcaacttggcccttttgaacatcgcgtacgagtcactgctgcgcgagagcgaattgcttcgcattaaagttggccatctga
aatccaccttcgaaggtgactacgttttaagtgtcccatatactaagacaaacgatagtggggaggaggaggttgtcaatatcacccctctgggcttcaaacttatccagcgctatattcaaggtgctggctt
aacgaaagaggactatttatttcaacctattggacgctcgaacaaggtctccgtacaggcgaaacccatgtcgactcgtacagtagatcgcgtcttcttgtgggcgtttgagtcccttggcatcgatcgccat
tcggcttggtctggtcattctgcacgtattggcgccgcgcaggaccttttagcagccggttactccattgctcagattcaagaaaatggccgctggaagtcccccatgatggtgttacgttatgggaaggaca
ttaaggccaaagagagcgcaatggctaagatgttagccgagcgccgctaa

8.2.1.10 Vika recombinase coding sequence

Atgggatccatgagtcaccatcaccaccaccatcatcaccatcatagtagtgaaaacctgtattttcaggggagtaccgatttaacgccttttccgcccttagaacacctggagccggatgaatttgctgat
ttggtccgtaaggcgattaaacgcgatccgcaagcaggcgctcacccggctattcagtccgccatttctcactttcaggacgaattcgtacgccgtcagggtgaatggcagccagccaccttacaacgtct
tcgcaatgcatggaatgtatttgtccgctggtgtacccaccagggaatccctgctttacctgctcgccatcaagacgttgagcgctatctgattgaacgtcgtaacgaacttcaccgcaacacgttgaaggt
gcacctttgggccatcggaaagactcacgttatttcgggtcttcctaatccttgcgcccatcgttatgtcaaggctcagatggcccaaattacacatcaaaaagtgcgtgagcgcgaacgtattgagcaagc
gcccgctttccgtgagtctgacttagatcgcttaacggaactttggtctgcaacgcgctcagtcactcaacagcgtgatcttatgatcgtctctcttgcctatgaaaccttactgcgtaagaataatttagaaca
aatgaaggtaggagacattgaattctgtcaagatggctctgctctgatcacgatccctttctcaaagacaaatcattcgggtcgcgatgacgtacgctggattagcccccaggtcgcgaatcaagttcacgc
ttatcttcaactgcccaacattgatgctgaccctcaatgtttttattacaacgtgttaagcgttctgggaaggctttaaatcccgaatctcacaacacacttaatggacaccatcctgtgtctgagaaacttat
cagccgtgtgtttgagcgtgcgtggcgcgcgttaaaccacgaaactggaccgcgttatacaggtcattccgcccgtgtcggagccgcccaagatctgctgcaagagggttattctacactgcaagttatgc
aagccggtgggtggtcctcagagaaaatggtattacgttatggacgccatctgcacgcccatacatcggcgatggcgcagaaacgccgtcagcgctaa

8.2.1.11 HA-Gent-HA fragment 4 (removal of native fragment 4)

atgaaaaaagtagcgcttggaaaaattatcgaaaacacagtagaaagcaaatataaaagcaacagtgtgtcgtcgtcattgtcaacgggcgccagtgcaaaattgagtttaagcgaatattacaaaacttttgaagcaaataaagt
gggccagcacactacgtacgacgtggtcggcaagcgagattacacgaaatttgacaaattggtgaaaaaatattgaccaagggcatgtaaagactatattcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagc
cgatctcggcttgaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagc
gcgttggcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacttgggcagaacgtaagccgcgagagcgccaa
caaccgcttcttggtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcat
ggatttgacttggtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctccataacatcaaacatcgaccca
cggcgtaacgcgcttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgcta
cttgcattacagtttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgctttttttcattttatacaaacaaaatttatacgtattgttagcac
attgagtgtcattttcgctgtctgaactatcacaatcatcgtcatcatcatcatcattgtcatcgtcgtcgtcacgtttgcgtttgacactgcattttttttggttaattttcactaacactggttcttttcgatcgtacaattgattctgcatgt

8.2.1.12 HA-Frag4-HA fragment with AmpR

atgaaaaaagtagcgcttggaaaaattatcgaaaacacagtagaaagcaaatataaaagcaacagtgtgtcgtcgtcattgtcaacgggcgccagtgcaaaattgagtttaagcgaatattacaaaactt
ttgaagcaaataaagtgggccagcacactacgtacgacgtggtcggcaagcgagattacacgaaatttgacaaattggtgaaaaaatattgaccagcctgctgtcgtgaataccgtatatctcaacgctgt
ctgtgagattgtcgcattctagcctttttagtttttcgctcatcgacttgatattgtccgacacattttcgtcgatttgcgttttgatcaacgacttgagcagagacacgttaatcaactgttcaaattgatccatatt
aactatatcaacccgatgcgtatatggtgcgtaaaatatatttttttaaccctcttatactttgcactctgcgttaatacgcgttcgtgtacagacgtaatcatgttttctttttggataaaactcctactgagtttga
cctcatattagaccctcacaagttgcaaaacgtggcatttttttaccaatgaagaatttaaagttattttaaaaaatttcatcacagatttaaagaagaaccaaaaattaaattatttcaacagtttaatcgacca
attaatcaacgtgtacacagacgcgtcggtgaaaaacacgcagcccgacgtgttggctaaaattatcaaatcaacttgtgttatagtcacagatttgccgtccaacgtgtttctcaaaaagttgaagaccaa
caagtttacagacactattaattatttaattttgccccactttattttgtgggatcacaattttgttatatttttaaacaaagctttcaattctaaacatgaaaacgatctggttgacatttcgggcgctctgcagaaa
atcaaacttacacacggtgtcatcaaagatcagttgcagagcaaaaacgggtacgcggtccaatacttgtacgcgacgtttctcaacacggcctcgttctacgccaacgtgcaatgtttaaatggtgtcaa
cgaaattatgccgccgcggagcagcgtaaagcgctattatggacgtgatgtggacaacgtgcgtgcatggaccacgcgtcatcccaacattagccagctgagtacgcaagtctcggacgtccacattaa
cgagtcatctaccgactggaatgtaaaagtgggtctgggaatatttcccggcgctaacacagactgcgacggtgacaaaaaaattattacatttttacccaaacctaattccctaatcgactcggaatgcct
tttgtacggcgaccctcggtttaatttcatttgctttgacaaaaaccgtttgtcgtttgtgtcacaacaaatttattatttgtacaaaaatattgacgcaatggaggcgttgtttaaatctacaccattggtttacgcg
ctgtggcaaaaacataaacatgagcagtttgcacagaggctagagatgttgttgcgtgatttttgcttaattgccagttcaaacgctagttatttactttttaaacagcttacacagctcatagctaacgaagaa
atggtgtgcggagatgaagaaatattcaatttaggcggccaatttgtagacatgattaaaagcggtgctaaaggcagtcaaaatctgattaaaagcacgcaacaataccgacagactttaaatacagatat
tgaaactgtgtcttcacgagccaccaccagtttaaatagttacatatcttctcacaataaggtaaaagtgtgtggcgccgacatatatcataacacggttgtgttacagagcgtgtttattaaaaataactatgt
ttgttacaaaaacgacgaacgtacaatcatgaatatttgcgctttgccctctgagtttctgtttccagaacatttgctcgacatgttcattgaatgataatataaatagagcgcatttgattgcatgcaatcagtgt
tttattaattttagagcaactatgattatacttatagaacactaatcggaataataacctaagcctgctagcagcatactaactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagc
cagggatgtaacgcactgagaagcccttagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcag
gcaccgtgtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgg
gttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagag
caactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacac
tgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgac
gagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggacca
cttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacga
cggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgcgtcgcggtgcatggagccgggccacctcgacctctaa
ctttaaataatgccaattatttaaagttactgttcgcggccgctcaagtgctgtcgtaatcttggtcaaaaacgttatgttgaaacataataacaccttacagtaacatacaataaaacaacatagtatcgtat
ataattataaactttatttttcattttatacaaacaaaatttatacgtattgttagcacattgagtgtcattttcgctgtctgaactatcacaatcatcgtcatcatcatcatcattgtcatcgtcgtcgtcacgtttgc
gtttgacactgcattttttttggttaattttcactaacactggttcttttcgatcgtacaattgattctgcatgt

8.2.1.13 HA-Gent-HA fragment 3 (removal of native fragment 3)

gctttacgagtagaattctacgtgtaaaacataatcaagagatgatgtcatttgtttttcaaaactgaactcaagaaatgatgtcatttgtttttcaaaactgaactggctttacgagtagaattctacttgtaacg
catgatcaagggatgatgtcatttgtttttcaaaaccgaactcgctttacgagtagaattctacttgtaaaacataatcgaaagatgatgtcatttgtttttaaaattgaactggctttacgagtagaattctactt
gtaaaacacaatcgagagatgatgtcatattttgcacacggctctaattaaactcgctttacgagtaaaattctacttgtaacgcatgatcaagggatgatgtattggatgagtcatttgtttttcaaaactaaa
ctcgctttacgagtagaattctacttgtaacgcacgcccaagggatgatgtcatttatttgtgcaaagctgatgtcatcttttgcacacgattataaacacaatcaaataatgactcatttgtttttcaaaactga
actcgctttacgagtagaattctacttgtaaaacacaatcaagcgatgatgtcattttaaaaatgatgtcatttgtttttcaaaactaaactcgctttacgagtagaattctacgtgtaaaacccaagggcatgt
aaagactatattcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggcttgaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgtatgc
ccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccgg
tgctcgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacttgggcagaacgtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatgt
cttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccgaactcacgaccgaaaagatcaagagcagcccgcatggatttgacttggtcagggccgag
cctacatgtgcgaatgatgcccatacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgta
acgcgcttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgagcg
catacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgctcaatttttcattatcg
tattgttgcacatcaacaggctggacaccacgttgaactcgccgcagttttgcggcaagttggacccgccgcgcatccaatgcaaactttccgacattctgttgcctacgaacgattgattctttgtccattga
tcgaagcgagtgccttcgactttttcgtgtccagtgtggcttgttttaataaattctttgaaaatattgtcgggtgtattattaaatagcatgtatggtatgttgaagatgggataacgcttggcgtgcgggtcgtcat
gatttccaccgcgcaccacatatttgcgctcaattttatcaaaattggactggcgagacaaaaacgagacgggcgacaggcatatttgggcgtgcgtaccatcttcggccaagg

8.2.1.14 HA-Frag3-HA fragment with AmpR

gctttacgagtagaattctacgtgtaaaacataatcaagagatgatgtcatttgtttttcaaaactgaactcaagaaatgatgtcatttgtttttcaaaactgaactggctttacgagtagaattctacttgtaacg
catgatcaagggatgatgtcatttgtttttcaaaaccgaactcgctttacgagtagaattctacttgtaaaacataatcgaaagatgatgtcatttgtttttaaaattgaactggctttacgagtagaattctactt
gtaaaacacaatcgagagatgatgtcatattttgcacacggctctaattaaactcgctttacgagtaaaattctacttgtaacgcatgatcaagggatgatgtattggatgagtcatttgtttttcaaaactaaa
ctcgctttacgagtagaattctacttgtaacgcacgcccaagggatgatgtcatttatttgtgcaaagctgatgtcatcttttgcacacgattataaacacaatcaaataatgactcatttgtttttcaaaactga
actcgctttacgagtagaattctacttgtaaaacacaatcaagcgatgatgtcattttaaaaatgatgtcatttgtttttcaaaactaaactcgctttacgagtagaattctacgtgtaaaacggtgaatgcatct
tactcaaagtccatcaattcgtacgcgtttgtgtccaggtcgggcgttgaaaaattgtagcttgccattagatcggatagcgattcaaattttgtaagcgtttgtagcgcacgtttggcatcttgtttaaaattaca
cgacgacagacagtaaaaatattcctcgataagcatgactacacccatatcactgtttaagtgctcgacgtagttgttgcatgttatgtcgcgtgtgccgcgatacgcgtgatttcggtgaaaatcacaccac
aaccagtcggcgtgcgtgtaacaaagtcgacagcgaaacaatttatcgttttccaaaaaatttaaatactcgacagttttgcagcttagattccgcgtttgattcaccttaaaatcgtcgtcagcctctataat
ctcgggcaacagcttgccttgttgccccatcgtatcgatcacctcccccaagtggcccggtgttatattaagtcgtttaaaatcatttattgcttcctgcacgtcggcctggtaattttttgaccacgggcgtgga
aatcaattgccgttgaagggaaataattcgtggtgtgggtatcggccgcctgttgcacaattccaccagcggtggaggcaagggcgcattcacagcaaccgttgtcatttataagtaatagtgtaaaaatgc
aaatattcatcaaaacattgacgggcaaaaccattaccgccgaaacggaacccgcagagacggtggccgatcttaagcaaaaaattgccgataaagaaggtgtgcccgtagatcaacaaagacttat
ctttgcgggcaaacaactggaagattccaaaactatggccgattacaatattcagaaggaatctactcttcacatggtgttacgattacgaggagggtattaataataacaataataaaaaccattaaatat
acataaaagtttttttatttaatctgacatatttgtatcttgtgtattatcgctaaccattaaaagtgctggagccacagtgttgcggcgagtcttatagaagatcgttgtttggctggaactgagcttttccttttcct
gctgccgctaatgggagtgggcacgtactctgtagtagacggtgcaacgggcaacttgagcgctaccgtcttaaatttggccatacttttagtgatgaaatcgcgcgttaacacttcgtcgtaaatgttactta
gcagaggcgcaacattgtgattaaatgtctcgtttaacaagctgtaaaactccgaataaagcttatcgcgcatttcgcagctctccttcaattctgccaaatttgcgttggtaagcaccacagtctgtcttttttt
gctcgctggaattgctgcgttctcgcttgaagacgacgatgtcgatcggtcggccattttttttgcccagcttttcagtgtgatcaaaaatgaacacaaaatctgccaattcgggcttgtttttcaccaaatccca
catggccgggctactaggccactcgggctgcttgatcttagtgtaccaactgttaaacaaaatgtatttattgttgtaatcactttcttcttgcgtttggacattttgcgttcgtcttgcatgacaggcaccacgtt
aaggatatagttaatgttctttctttccaagaaatttacaataacggccagctggtccatgttggatttgttgtaagagctcgattccagtttattcaacagcttttcatttttgcacacggccgcagtctccggag
attgttgctccggcacgtttaccatgtttgcttcttgtaaacctttgaaacaacccgtttgtattcttgatgatatatttttttaatgcccaacaacctggcaattcgtttgtgatgaagacacaccttacgcttcgaa
catttgtcggtgattactgtgaaatggcctaaattagctcttatatattcttttatacgctcaaacgacacgatgtccaacatgtgcgcgcagacgttttctgtgttcatcgtgtgcttgagcgtgttgatggcttcc
ctgaacagcgcttgtatttcgctgcgagtcaagcagtccgaatcacacccgcctaagtgcgtgcaatttttgggggggcatcgttgtctatctttttcagagtggcgtagaaaaagtcctgcaattgcctattatc
aaaacgcgccttgacgctgcgcacaaaatcaaaaaattcaatgtaattgctgtaatcgtacgtgatcagttgtttgtcgttcatataattaaagtatttgttgagcggcacgatggccaggctgcgcgctattt
cgcaattgaagcgtcgcggttttaacattatacggtagtcattgccaaacgtgcccggcaacaacttcacggtgtacgtgttgggtttggcgttcacgttaatcaagttgccgcgcacgacgcctacgtatat
caaatacttgtaggtgacgccgtcatctttccattgtaacgtaaatggcaacttgtagatgaacgcgctgtcaaaaaaaccggccagtttcttccacaaactcgcgcacggctgtctcgtaaactttgcgtcg
caacaatcgcgatgacctcgtggtatggaaatttttttctaaaaaagtgtcgttcatgtcggcggcgggcgcgttcgcgctccggtacgcgcgacgggcacacagcaggacagccttgtccggctcgattat
cataaacaatcctgcagcgtttcgcattttacatatttgacacttaaaaaattgcgcacacgagcaccatcgtttgatacctaattgcaactatttacaatttatcagtttacgttgaacccgttttaatttttttaga
tccgtccttgttcagttgcaagttgactaaatgacaaaatttttcggttctgcaaaaccgcccttgtctgttccacccgttgtatttgaaaaaactttttttcacgcggcgacaactgcttgtataatattgcccaa
tgtaaacatgcaaaattgtttttgcgtcggtagtattttggcaggcaataatcaaaaaaaatccgtaagcaattctctgcatctattaatattcgttgcgtacgaatcgagtttttcaaaaattactttgtttgtatga
aaataacgtttgggcttctcacaataataatcttcgttgtagaacagaaacggtttgcgagaattggcacgtttgtccatgattggctcagtgtaacgattgattcaaatcaaaattgacaacacgtttgccgta
atgtgcaccggttcgcacacgtttgccgcgtatgtaatccatgtttatttcgctgtcgcaattgattacacgattgtgttgggcggcgcgttttattgaatttaggcgacgcgtcgacaactccaaaggattgtaa
agcgcagatttttccagagtaaacgagtttaagtggccaccgttgaaccattccagagccacgattgtgtacagcaaaaagaatatttcttgtcgacgttttcaaacgcaaacttgtttttttaggcaatagta
gtaaaattttaacgaattgtataaataaaacataaaattgccatttttaaagtaaaattctacatccgtgacgaacaaaaggtttactattttgttctccaacaagtgtgccaattttcttaagtacaccattgaat
ttttgtcgtcgtccatctcgatcaacaacacgtacggcgttttggaatttaaaattattctaaaattttcctgttgcaacgattccacagcgtccgaccaatatgacgctgccacctctagacagatgtatttctt
ggaaaacacgtgtcgtttgataacctcgctgatggacgtgatcgattgtaaatacttttcaaacgtcgcgtcttcccaaccacgcaccgaaacgggcgctgtcgtgtcgggctgatgtttgaaatccaaacc
actctgaattaacttggttgtgattcgtatgctcaactgttgacccaacgtgtagtgatcttcgtaggcgcgctcccacatcacgttacacacaaatttgacgagatcatcaacgtctttctgttgcaaaattcg
ccgcaaacgcgccacatcgcccttgtaccaccgatctcggcacacaagctgtagcatttttaaatcgtgatcgctcaagctattaattctggttagatttatatagtcgtcaatatcctcgggcgtggtttgcgt
catgtctgtaaaacgtgcaaaatcaaacatttttatgttgtagtcgaatctaacaaatccatcggcgttcacttgcacttcgcgctttacaaaacgaggtagcgtgtaatcgaacccgtttaaatagattgcgt
acaaaaaccagcacttcatcttccagtttgcacgcttgcggcaaaaattgtgtggtgtgctccaaccgggtgacaaacatgactatggaaaataacgcggaattcaacagacgactagagtacgtgggca
cgatcgccacaatgatgaaacgaacattgaacgttttacgacagcagggctattgcacgcaacaggatgcggattctttgtgcgtgtcagacgacacggcggcctggttatgcggccgtttgccgacctgc
aattttgtatcgttccgcgtgcacatcgaccagtttgagcatccaaatccggcgttggaatattttaaatttgaagaaagtctggcgcaacgccaacacgtgggcccgcgttacacgtacatgaattacacg
ctttttaaaaacgtcgtggccctcaaattggtcgtgtacacgcgcacgctacaagctaacatgtacgcggacgggttgccgtattttgtgcaaaatttttcagaaacaagctacaaacatgttcgtgtgtatgt
tagaaaacttggtgcgatacaagtagcgacattatcagtttacgaacaaattattgaagatacaataaatgaactcgtcgtcaatcacgttgattagataatgtccgaataggtctgagaacgcccattctca
gacgtattggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagcccttagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggt
agtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaa
agatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggc
gcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaatt
atgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttggg

aaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaata
gactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagat
ggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgcaata
ggtctgagaacgcccattctcagacgtattgtcgcggtgcatggagccgggccacctcgacctctaactttaaataatgccaattatttaaagttacaataaccatagtttaatcgaagagaatagccgtcg
ccacattatatttagggagaatgtgatattcaaaagactgactgttaacacaaaagactgatattgttgttgttacaaaatagataataaaacaaaaaataaattaaatacctggcactttttaatgtgaatatt
aaaaggttgttgcatatatatcgttaaacgtttatgtttactttcacgttagctcgtttcattgatgtaaacatttagttttataacagcgtcggtaattttatttttttaaagtaaacagaccaaaatcaaaggtgtctt
cgacaggtacgattattttcccattgacactgttttcgtgcacagatataattttatcaccgtttattattttgcccaaacacacgtactcgtttcttctcaagccaactatttctaaacaattcacttttctattatcg
tgtacgcaattaaaagtaaacgaagcgctacaattgtcgtattctattacaattctgcggcatttataaaatttattaatgttgacgcaaattccatgcagcgcatccatttcgtactgcaaatgcggcgcaatt
aaaaaatttcctcgtcgttgttaacaatcttgggcgctaaaaagcacgccaacacgcccacgtctttaatgcaatattccaatttgaacggcagttcctcggacatgtatattgtcacggtgggcgccaaag
gagcggctttagcaaaatgacacaagtaatcgcccgcaaaagtggcagatccttaaaagtttccaccagcctcttcaaaactgcgccggttttaaattccgcttcgaacatttttagcagtgattctaattgc
agctgctctttgatacaactaattttacgacgacgatgcgagcttttattcaaccgagcgtgcatgtttgcaatcgtgcaagcgttatcaatttttcattatcgtattgttgcacatcaacaggctggacaccacg
ttgaactcgccgcagttttgcggcaagttggacccgccgcgcatccaatgcaaactttccgacattctgttgcctacgaacgattgattctttgtccattgatcgaagcgagtgccttcgactttttcgtgtcca
gtgtggcttgttttaataaattctttgaaaatattgtcgggtgtattattaaatagcatgtatggtatgttgaagatgggataacgcttggcgtgcgggtcgtcatgatttccaccgcgcaccacatatttgcgctca
attttatcaaaattggactggcgagacaaaaacgagacgggcgacaggcatatttgggcgtgcgtaccatcttcggcca

8.2.2 Chapter 4

8.2.2.1 BSD cassette

Taatacgactcactatagggtagggataacagggtaatattctctagacatcattaattcctaatttttgttgacactctatcattgatagagttattttaccactccctatcagtgatagagaaaagtgaaatg
gcaactgttaaccagctggtgcgcaaaccgcgtgcgcgtaaggttgcaaaaagcaacgttcctgctctggaagcttgcccgcagaagcgtggtgtatgtactcgcgtatataccaccacccctaaaaaa
ccaaactctgcactgcgtaaagtttgtcgtgtgcgtctgactaacggcttcgaagttacctcttacatcggtggcgaaggccacaacctgcaagaacactccgtaattctgatccgtggtggccgtgtaaaa
gacctcccaggtgtgcgttaccacactgttcgcggtgcactggactgttccggtgttaaagaccgtaaacaagcgcgttctaagtacggtgtgaagcgtccaaaagcttaaaggaggacaatcatggcca
agcctttgtctcaagaagaatccaccctcattgaaagagcaacggctacaatcaacagcatccccatctctgaagactacagcgtcgccagcgcagctctctctagcgacggccgcatcttcactggtgt
caatgtatatcattttactgggggaccttgtgcagaactcgtggtgctgggcactgctgctgctgcggcagctggcaacctgacttgtatcgtcgcgatcggaaatgagaacaggggcatcttgagcccctg
cggacggtgtcgacaggtgcttctcgatctgcatcctgggatcaaagcgatagtgaaggacagtgatggacagccgacggcagttgggattcgtgaattgctgccctctggttatgtgtgggagggctaata
gggataacagggtaatcctcgactgtgccttctatg

8.2.3 Chapter 5

8.2.3.1 ac20VloxP fragment

actaaattaatactttatataatcaaaaataatactttatataatacttttcaatttctgagaactgtcattctcggaaattgaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccc
ttagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttcc
gtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcg
gtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattct
cagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgat
cggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagc
aatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggct
ggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaac
gaaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctcaatttctgagaactgtcattctcggaaattgatgtgcacgcgtccggtgaggtacacaaaactttc
ttgggcacgctgtaca

8.2.3.2 ac87RoxP fragment

tatataaacgcaatgtaacagtttttgctagtaccatcgcatacaacatttaactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccct
tagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcgg
taagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctc
agaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatc
ggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagca
atggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctg
gtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacg
aaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctaactttaaataatgccaattatttaaagttattaatctacatttattgtaacatttgtggtaatagtggcgt
tggttatac

8.2.3.3 ac32RoxPTn7 fragment

ataaaataattatttaaataaaaatgtttttattgtaaaatacacattgataactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccctt
agagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttccgt
gtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggt
aagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcg
gaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaa
tggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctgg

tttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctaactttaaataatgccaattatttaaagttaccgcctctccccgcgcgttggccgattcattaatgcag
ctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagc
ggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatccc
ccgggctgcaggaattcacataacaggaagaaaaatgccccgcttacgcagggcatccatttattactcaaccgtaaccgattttgccaggttacgcggctgtcgacctcgagggggggcccggtaccc
aattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaat
agcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcttcgaagacgctcgac
agcgacacacttgcatcggatgcagcccggttaacgtgccggcacggcctgggtaaccaggtattttgtccacataaccgtgcgcaaaatgttgtggataagcaggacacagcagcaatccacagcag
gcatacaaccgcacaccgaggttactccgttctacaggttacgacgacatgtcaatacttgcccttgacaggcattgatggaatcgtagtctcacgctgatagtctgatcgacaatacaagtgggaccgat
gaactttaagtccccacagcaaactggcgcttttatataaaaatttg

8.2.3.4 ac87RoxPTn7 fragment

tatataaacgcaatgtaacagtttttgctagtaccatcgcatacaacatttaactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccct
tagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttccg
tgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcgg
taagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctc
agaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatc
ggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagca
atggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctg
gtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacg
aaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctaactttaaataatgccaattatttaaagttaccgcctctccccgcgcgttggccgattcattaatgca
gctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgag
cggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatcc
cccgggctgcaggaattcacataacaggaagaaaaatgccccgcttacgcagggcatccatttattactcaaccgtaaccgattttgccaggttacgcggctgtcgacctcgagggggggcccggtacc
caattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaa
tagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcttcgaagacgctcgac
agcgacacacttgcatcggatgcagcccggttaacgtgccggcacggcctgggtaaccaggtattttgtccacataaccgtgcgcaaaatgttgtggataagcaggacacagcagcaatccacagcag
gcatacaaccgcacaccgaggttactccgttctacaggttacgacgacatgtcaatacttgcccttgacaggcattgatggaatcgtagtctcacgctgatagtctgatcgacaatacaagtgggacctta
atctacatttattgtaacatttgtggtaatagtggcgttggttatac

8.2.3.5 ac110RoxPTn7 fragment

aaacgggcactccatcgtgatgtatatgtcttacttcattagactttagataactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccctt
agagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttccgt
gtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggt
aagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcg
gaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaa
tggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctgg
tttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacga
aatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctaactttaaataatgccaattatttaaagttaccgcctctccccgcgcgttggccgattcattaatgcag
ctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagc
ggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatccc
ccgggctgcaggaattcacataacaggaagaaaaatgccccgcttacgcagggcatccatttattactcaaccgtaaccgattttgccaggttacgcggctgtcgacctcgagggggggcccggtaccc
aattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaat
agcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcttcgaagacgctcgac
agcgacacacttgcatcggatgcagcccggttaacgtgccggcacggcctgggtaaccaggtattttgtccacataaccgtgcgcaaaatgttgtggataagcaggacacagcagcaatccacagcag
gcatacaaccgcacaccgaggttactccgttctacaggttacgacgacatgtcaatacttgcccttgacaggcattgatggaatcgtagtctcacgctgatagtctgatcgacaatacaagtgggaccctttt
aaatcttatgctagttgaataaaatacatagtacttttatacgttt

8.2.3.6 ac112/113RoxPTn7 fragment

tataaaagtgtttgaattgcgagaccgtcaacataacgtttatcaacgcgtaactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagccc
ttagagcctctcaaagcaattttcagtgacaattctcatgtttgacagcttatcatcgataagctttaatgcggtagtttatcacagttaaattgctaacgcagtcaggcaccgtgtatgagtattcaacatttcc
gtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcg
gtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattct
cagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgat
cggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagc
aatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggct
ggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaac
gaaatagacagatcgctgagataggtgcctcactgattaagcattggtaagcctccccgcgttgctaactttaaataatgccaattatttaaagttaccgcctctccccgcgcgttggccgattcattaatgc
agctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtga
gcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatc
ccccgggctgcaggaattcacataacaggaagaaaaatgccccgcttacgcagggcatccatttattactcaaccgtaaccgattttgccaggttacgcggctgtcgacctcgagggggggcccggtac
ccaattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgta
atagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcttcgaagacgctcga

cagcgacacacttgcatcggatgcagcccggttaacgtgccggcacggcctgggtaaccaggtattttgtccacataaccgtgcgcaaaatgttgtggataagcaggacacagcagcaatccacagca
ggcatacaaccgcacaccgaggttactccgttctacaggttacgacgacatgtcaatacttgcccttgacaggcattgatggaatcgtagtctcacgctgatagtctgatcgacaatacaagtgggaccaa
caaattaatttaaacaagcgtttttattgacaatactcacatttgata

8.2.3.7 pUCDM_YFP_polh

ataacttcgtatagcatacattatacgaagttatctggagtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttgggggttggggtggggaaaa
ggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttattg
ccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagtatcgattcgcgacctactccggaatattaatagatcat
ggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatcccggtc
cgaagcgcgcggaattcaaagggctaccggtcgccaccatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgt
gtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccttcggctacggcctgcagtgcttcg
cccgctaccccgaccacatgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagtt
cgagggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaa
gcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgac
aaccactacctgagctaccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtccg
gactcagatctcgagctcaagcttcgaattctgcagtcgacggtaccgcgggcccgggatccaccggatctagataactgatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacc
tcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgc
attctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagc
aataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggatctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaa
aacaccatcatacactaaatcagtaagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgataccgggaagc
cctgggccaacttttggcgaaaatgagacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaa
aatggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggcctttt
taaagaccgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgcccgcctgatgaatgctcatccggaattccgtatggcaatgaaagacggtgagctggtgatatgggatagtgttca
cccttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacgacgatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctattt
ccctaaagggtttattgagaatatgtttttcgtctcagccaatccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacgcaa
ggcgacaaggtgctgatgccgctggcgattcaggttcatcatgccgtctgtgatggcttccatgtcggcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaattttttaag
gcagttattggtgcccttaaacgcctggtgctacgcctgaataagtgataataagcggatgaatggcagaaattcgaaagcaaattcgacccggtcgtcggttcagggcagggtcgttaaatagccgcttat
gtctattgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctgaggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctccc
agagcctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagcagcccttgcgccctgagtgcttgcggcagcgtgaagctaattctgtcagccgttaagtgttcc
tgtgtcactgaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccacaaccgttaaaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaacc
ttagaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatgagagcttagtacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacat
gagagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggtacc

8.2.4 Chapter 6

8.2.4.1 pUPDM_VloxP_mCherry

gaattagcttcacgctgccgcaagcactcagggcgcaagggctgctaaaggaagcggaacacgtagaaagccagtccgcagaaacggtgctgaccccggatgaatgtcagctgggaggcagaataa
atgatcatatcgtcaattattacctccacggggagagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataagcg
gctatttaacgaccctgtcctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgcttattatcacttattcaggcgtagcaaccaggcgtttaagggcaccaataactgccttaaaaa
aatcaggcaccgggcttgcgggtcatgcaccaggtcgcgcggtccttcgggcactcgacgtcggcggtgacggtgaagccgagccgctcgtagaaggggaggttgcggggcgcggaggtctccagga
aggcgggcaccccggcgcgctcggccgcctccactccggggagcacgacggcgctgcccagacccttgccctggtggtcgggcgagacgccgacggtggccaggaaccacgcgggctccttgggc
cggtgcggcgccaggaggccttccatctgttgctgcgcggccagccgggaaccgctcaactcggccatgcgcgggccgatctcggcgaacaccgcccccgcttcgacgctctccggcgtggtccaga
ccgccaccgcggcgccgtcgtccgcgacccacaccttgccgatgtcgagcccgacgcgcgtgaggaagagttcttgcagctcggtgacccgctcgatgtggcggtccggatcgacggtgtggcgcgtg
gcggggtagtcggcgaacgcggcggcgagggtgcgtacggccctggggacgtcgtcgcgggtggcgaggcgcaccgtgggcttgtactcggtcattttagcttccttagctcctgaaaatctcgataactc
aaaaaatacgcccggtagtgatcttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttattt
attctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcct
gttcagatccggatatagttcctcctttcagcaaaaaacccctcaagacccgtttagaggccccctaggggttatgatagttattgctcagcggtggcagcagccaactcagcttcctttcgggctttgttagca
gccggatcttctaggctcaagcagtgatcagatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttattt
gtaaccattataagctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggaggttttttaaagcaagtaaaacctctacaaatgtggtatggctgattatgat
cctctagtacttctcgacaagcttgtcgagactgcaggctctagattacttgtacagctcgtccatgccgccggtggagtggcggccctcggcgcgttcgtactgttccacgatggtgtagtcctcgttgtggg
aggtgatgtccaacttgatgttgacgttgtaggcgccgggcagctgcacgggcttcttggccttgtaggtggtcttgacctcagcgtcgtagtggccgccgtccttcagcttcagcctctgcttgatctcgccct
tcagggcgccgtcctcggggtacatccgctcggaggaggcctcccagcccatggtcttcttctgcattacggggccgtcggaggggaagttggtgccgcgcagcttcaccttgtagatgaactcgccgtcc
tgcagggaggagtcctgggtcacggtcaccacgccgccgtcctcgaagttcatcacgcgctcccacttgaagccctcggggaaggacagcttcaagtagtcggggatgtcggcggggtgcttcacgtag
gccttggagccgtacatgaactgaggggacaggatgtcccaggcgaagggcagggggccacccttggtcaccttcagcttggcggtctgggtgccctcgtaggggcggccctcgccctcgccctcgatc
tcgaactcgtggccgttcacggagccctccatgtgcaccttgaagcgcatgaactccttgatgatggccatgttatcctcctcgcccttgctcaccatggatccgcgcccgatggtgggacggtatgaataa
tccggaatatttataggtttttttattacaaaactgttacgaaaacagtaaaatacttatttatttgcgagatggttatcattttaattatctccatgatctattaatattccggagtaggtcgcgaatcgatactagt
atacggacctttaattcaacccaacacaatatattatagttaaataagaattattatcaaatcatttgtatattaattaaaatactatactgtaaattacattttatttacaatcactcgacgaagacttgatcacc
cgggatctcgagccatggtgctagcagctgatgcatagcatgcggtaccgggagatgggggaggctaactgaaacacggaaggagacaataccggaaggaacccgcgctatgacggcaataaaaag
acagaataaaacgcacgggtgttgggtcgtttgttcataaacgcggggttcggtcccagggctggcactctgtcgataccccaccgagaccccattgggaccaatacgcccgcgtttcttccttttcccca
ccccaaccccccaagttcgggtgaaggcccagggctcgcagccaacgtcggggcggcaagccctgccatagccactacgggtacgtttaaaccagtcaatttccgagaatgacagttctcagaaatgat
atcggtaccgcggccgcgtagaggatctgttgatcagcagttcaacctgttgatagtacgtactaagctctcatgtttcacgtactaagctctcatgtttaacgtactaagctctcatgtttaacgaactaaac
cctcatggctaacgtactaagctctcatggctaacgtactaagctctcatgtttcacgtactaagctctcatgtttgaacaataaaattaatataaatcagcaacttaaatagcctctaaggttttaagttttata
agaaaaaaaagaatatataaggcttttaaagcttttaaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagcccttagagcctctcaaagcaattttgagtgacacaggaacactt
aacggctgaca

8.2.4.2 pKL_NLSmTagBFP x pUCDM_SNAPTag_GP64 x pSPL_mCherry

ataacttcgtatagcatacattatacgaagttatctggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttgggggttggggtgggg
aaaaggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttt
tattgccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtctt
cgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgcg
acctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtc
ccaccatcgggcgcgatgagcgagctgattaaggagaacatgcacatgaagctgtacatggagggcaccgtggacaaccatcacttcaagtgcacatccgagggcgaaggcaagccctacgagggc
acccagaccatgagaatcaaggtggtcgagggcggccctctcccattcgccttcgacatcctggctactagcttcctctacggcagcaagaccttcatcaaccacacccagggcatccccgacttcttc
aagcagtccttccctgagggcttcacatgggagagagtcaccacatacgaagacgggggcgtgctgaccgctacccaggacaccagcctccaggacggctgcctcatctacaacgtcaagatcagag
gggtgaacttcacatccaacggccctgtgatgcagaagaaaacactcggctgggaggccttcaccgagacgctgtaccccgctgacggcggcctggaaggcagaaacgacatggccctgaagctcgt
gggcgggagccatctgatcgcaaacatcaagaccacatatagatccaagaaacccgctaagaacctcaagatgcctggcgtctactatgtggactacagactggaaagaatcaaggaggccaacaa
cgagacctacgtcgagcagcacgaggtggcagtggccagatactgcgacctccctagcaaactggggcacaagcttaatggatcctccggactcagatctcgagctgatccaaaaaagaagagaaag
gtagatccaaaaaagaagagaaaggtagatccaaaaaagaagagaaaggtaggatccaccggatctagataagatcccggtccgaagcgcgcggaattcaaaggcctacgtcgacgagctcactag
tcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctccccctga
acctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtcca
aactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactatcataacccct
aggagatccgaaccagataagtgaaatctagttccaaactattttgtcatttttaattttcgtattagcttacgacgctacacccagttcccatctattttgtcactcttccctaaataatccttaaaaactccattt
ccacccctcccagttcccaactattttgtccgcccacagcggggcatttttcttcctgttatgtttttaatcaaacatcctgccaactccatgtgacaaaccgtcatcttcggctactttttctctgtcacagaat
gaaaatttttctgtcatctcttcgttattaatgtttgtaattgactgaatatcaacgcttatttgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgca
gcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccaaacgtactgcggtaaagctcatcagcgtggtcgtgaagcaattcttagaaaaact
catcgagcatcaaatgaaactgcaatttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcg
gtctgcgattccgactcgtccaacatcaatacaacctattaatttcccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagtttatgcatttct
ttccagacttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattaca
aacaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacctggaatgctgttttcccggggatcgcagtggtgagtaaccatgcat
catcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaactctggcgc
atcgggcttcccatacaagcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctagagcaagacgtttcccgttga
atatggctcataacaccccttgtattactgtttatgtaagcagacagttttattgttcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttg
agatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagata
ccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccg
ggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatga
gaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggt
ttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgc
gttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtatttt
ctccttacgcatctgtgcggtatttcacaccgcatagaccagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgccctgcgtaagcgggtgtgggcggacaataaagtcttaaactg
aacaaaatagatctaaactatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttttgttatggctaaagcaaactcttcattttctg
aagtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatggtaaagactatattcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggcttgaacgaattgt
taggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttg
gcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacttgggcagaacgtaagccgcga
gagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccgaactcacgac
cgaaaagatcaagagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttttagggcgactgccctgctgcgtaacatcgttgctgc
tgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatgaaaaccgccactgc
gccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtttccacggtgtgc
gtcacccggcaaccttgggcagcagcgaagtcgaggcatttctgtcctggctggcgaacgagcgcaaggtttcggtctccacgcatcgtcaggcattggcggccttgctgttcttctacggcaaggtgctgt
gcacggatctgcccttgcttcaggagatcggtagacctcggccgtcgcggcgcttgccggtggtgctgaccccggatgaagtggttcgcatcctcggttttctggaaggcgagcatcgtttgttcgcccagg
actctagctatagttctagtggttggctacagctttgtttgtactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggtaccataacttcgtatagcatacattatacgaagttatctggtt
taaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttgggggttggggtggggaaaaggaagaaacgcgggcgtattggtcccaatggg
gtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttattgccgtcatagcgcgggttccttccggtattgtctc
cttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtcttcgtcgagtgattgtaaataaaatgtaatttacagtata
gtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgcgacctactccggaatattaatagatcatggagataatt
aaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatccatggtaagcgctattg
ttttatatgtgcttttggcggcggcggcgcattctgcctttgcggtcgacgacaaagactgcgaaatgaagcgcaccaccctggatagccctctgggcaagctggaactgtctgggtgcgaacagggcctg
cacgagatcaagctgctgggcaaaggaacatctgccgccgacgccgtggaagtgcctgccccagccgccgtgctgggcggaccagagccactgatgcaggccaccgcctggctcaacgcctactttc
accagcctgaggccatcgaggagttccctgtgccagccctgcaccacccagtgttccagcaggagagctttacccgccaggtgctgtggaaactgctgaaagtggtgaagttcggagaggtcatcagct
accagcagctggccgccctggccggcaatcccgccgccaccgccgccgtgaaaaccgccctgagcggaaatcccgtgcccattctgatcccctgccaccgggtggtgtctagctctggcgccgtggg
gggctacgagggcgggctcgccgtgaaagagtggctgctggcccacgagggccacagactgggcaagcctgggctgggtgagctcgacttacaagtgccacgagctgaacactgcaacgcgcaaat
gaaaacgggcccatacaagataaagaacttagatattactccaccaaaagagacactccagaaagatgtagaaatcacgatcgtggagactgattataatgagaacgttattattggatacaagggata
ctatcaagcctatgcctataacgggggtagccttgacccaaatactagagttgaggagacaatgaaaaccctcaatgtggggaaggaagatttgctaatgtggtccattcgacaacaatgcgaagtcggc
gaagaactcatcgacagatgggggagcgattccgacgattgtttcagggataatgaggggcgcggacagtgggtcaaagggaaagagttggttaagcgccagaacaacaaccacttcgcacatcatac
ctgcaacaaatcgtggaggtgtgggattagcacgtccaaaatgtactcaagattggagtgtcaagacgatacagatgagtgtcaagtctacatcctggacgcagagggcaatccgataaatgttaccgtg
gatactgtcctgcacagagatggagtgtcgatgattttaaagcagaagtctacattcacaacacgccagataaaggcagcatgtttattgattaaggatgacaagaataacccggagtccgtaacaagag
aacactgcctgatagataatgatatttacgacctaagcaagaatacctggaactgtaagtttaaccgttgtatcgcgcgcggtggcggtccgaagagaaaggttgagcatcgcgtaaaaaagcgtccacc
gacctggaggcacaatgttcgcgccaaatacaccgaaggggacacggccacaaagggggaccttatgcacatccaggaggagctgatgtacgagaacgacctgcttaagatgaacattgaactcatg
cacgcgcacatcaataagctcaataacatgttacatgatctcatagtcagtgtggcaaaagttgatgaacgcctaataggtaacttgatgaacaattctgtgtcctcgacgttcttgagtgatgacacattctt
attgatgccatgcactaatccaccagctcacacgtccaactgctataacaactccatttacaaggagggcagatgggtcgcaaatacggatagttcgcaatgtatagattttagtaattacaaagaactgg
caatagacgacgacgtcgagttctggattcctaccattggcaacacaacgtaccatgattcctggaaagacgcttccgggtggtcctttattgctcagcagaaatcaaatctcattacaactatggagaac
acgaaattcggtggcgtagggacgagccttagcgacattacgtctatggcagagggtgagctggctgctaagctcacatccttcatgtttggacacgtcgtaaactttgttataatcttaattgtcattcttttcc

tatattgtatgatacgcaacagaaacagacagtactaatagctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcc
cacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattc
tagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaata
actatcataacccctaggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggatctgaacaggagggacagctgatagaaacagaagccactggagcacctcaaaaaca
ccatcatacactaaatcagtaagttggcagcatcacccgacgcactttgcgccgaataaatacctgtgacggaagatcacttcgcagaataaataaatcctggtgtccctgttgataccgggaagccctg
ggccaacttttggcgaaaatgagacgttgatcggcacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtattttttgagttatcgagattttcaggagctaaggaagctaaaatg
gagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggcctttttaa
agaccgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgcccgcctgatgaatgctcatccggaattccgtatggcaatgaaagacggtgagctggtgatatgggatagtgttcaccc
ttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacgacgatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctatttccc
taaagggtttattgagaatatgtttttcgtctcagccaatccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacgcaaggc
gacaaggtgctgatgccgctggcgattcaggttcatcatgccgtttgtgatggcttccatgtcggcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaatttttttaaggcag
ttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcggatgaatggcagaaattcgaaagcaaattcgacccggtcgtcggttcagggcagggtcgttaaatagccgcttatgtct
attgctggtttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctgaggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctcccagct
gacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagcagcccttgcgccctgagtgcttgcggcagcgtgaagctgaattctgtcagccgttaagtgttcctgtgtc
actcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccacaaccgttaaaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaaccttagag
gctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatgagagcttagtacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacatgagag
cttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggtaccataacttcgtatagcatacattat
acgaagttatctggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttggctgcgagccctgggccttcacccgaacttggggggtggggtggggaaaaggaagaaacgcgggcgt
attggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaaccccgcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttattgccgtcatagcgcgggttc
cttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcatcagctgctagcaccatggctcgagatcccgggtgatcaagtcttcgtcgagtgattgtaaataaaat
gtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagtatcgattcgcgacctactccggaatattaataga
tcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcgatggtgc
tagcagtgtcgaagggcgaggaagacaacatggcgatcatcaaggagttcatgcgcttcaaagtccacatggaaggcagcgtgaatggccatgagttcgagatcgaaggtgaaggcgaaggtcgtcctt
acgaaggcacccaaaccgcgaaactgaaagtgaccaaaggtggtccgcttccttttgcctgggatattctgtcgccgcagttcatgtatggcagcaaagcctacgtgaaacatcccgcggatatccccg
actatctgaaactgagcttccccgaaggcttcaaatgggagcgcgtgatgaactttgaagatggcggcgtcgtgaccgttacccaagattcgtcactgcaggatggcgagttcatctacaaggtcaaactg
cgcggcacgaattttccgtctgatggcccggtgatgcagaaaaagacgatgggctgggaagcaagctcggaacgtatgtatccagaagatggcgccctgaaaggcgaaatcaaacagcgcctcaagc
tgaaagacggtggccattatgatgccgaggtcaagacgacctacaaggcgaaaaaacccgtccaacttccaggcgcgtacaacgtcaacatcaaactcgacatcaccagccacaacgaagactaca
ccattgtggagcagtatgaacgtgctgaaggacgccatagtacaggcggcatggatgaactgtacaagtaagatcccggtccgaagcgcgcggaattcaaaggcctacgtcgacgagctcactagtcg
cggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctcccacacctcccccctgaacc
tgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaac
tcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactatcataacccctagg
agtccaagcgagctcgatatccgtcggcttgaacgaattgttagacattatttgccgactaccttggtgatctcgcctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggccaagcgatcttc
ttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggccggcaggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgccggttactgcgctgtaccaaatgcg
ggacaacgtaagcactacatttcgctcatcgccagcccagtcgggcggcgagttccatagcgttaaggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctccgccgctg
gacctaccaaggcaacgctatgttctcttgcttttgtcagcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaattgcagttcgcgctt
agctggataacgccacggaatgatgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggtcgttgatcaaagctcgccgcgttgttt
catcaagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgccatccactgcggagccgtacaaatgtacggccagcaacgtcggttcgagatggcgctcgatgacgccaa
ctacctctgatagttgagtcgatacttcggcgatcaccgcttccctcatgatgtttaactttgttttagggcgactgccctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaa
cgcgcttgctgcttggatgcccgaggcatagactgtaccccaaaaaaacagtcataacaagccatgaaaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgag
cgcatacgctacttaccggtttattgactaccggaagcagtgtgaccgtgtgcttctcaaatgcctgaggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcatttattctgcctccc
agctgacattcatccggggtcagcaccgtttctgcggactggctttctacgtgttccgcttcctttagcagcccttgcgccctgagtgcttgcggcagcgtgaagctaattctgtcagccgttaagtgttcctgt
gtcactcaaaattgctttgagaggctctaagggcttctcagtgcgttacatccctggcttgttgtccacaaccgttaaaccttaaaagctttaaaagccttatatattcttttttttcttataaaacttaaaacctta
gaggctatttaagttgctgatttatattaattttattgttcaaacatgagagcttagtacgtgaaacatgagagcttagtacgttagccatgagagcttagtacgttagccatgagggtttagttcgttaaacatga
gagcttagtacgttaaacatgagagcttagtacgtgaaacatgagagcttagtacgtactatcaacaggttgaactgctgatcaacagatcctctacgcggccgcggtacc

8.2.4.3 pACEBac1 polH TAF5 p10 TAF6 polH TAF9 p10 mCherry

accggttgacttgggtcaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatcctttttgataatctcatgaccaaaatcccttaacgtga
gttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatc
aagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctc
tgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcacacagcccag
cttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagc
gcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaac
gcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattgacttgggtcgctcttcctgtggatgcgcagatgccctgcgtaagcgggtgtgggcggacaata
aagtcttaaactgaacaaaatagatctaaactatgacaataaagtcttaaactagacagaatagttgtaaactgaaatcagtccagttatgctgtgaaaaagcatactggacttttgttatggctaaagcaa
actcttcattttctgaagtgcaaattgcccgtcgtattaaagaggggcgtggccaagggcatgtaaagactatattcgcggcgttgtgacaatttaccgaacaactccgcggccgggaagccgatctcggct
tgaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccactgcgggatcgtcaccgtaatctgcttgcacgtagatcacataagcac
caagcgcgttggcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgctcgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacttgggcagaac
gtaagccgcgagagcgccaacaaccgcttcttggtcgaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccg
aactcacgaccgaaaagatcaagagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgttttagggcgactgccctgctgcgtaa
catcgttgctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgaggcatagactgtacaaaaaaacagtcataacaagccatgaaa
accgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagttgcgtgagcgcatacgctacttgcattacagtttacgaaccgaacaggcttatgtcaactgggttcgtgccttcatccgtt
tccacggtgtgcgtcacccggcaaccttgggcagcagcgaagtcgccataacttcgtatagcatacattatacgaagttatctgtaactataacggtcctaaggtagcgagtttaaacgctagcatcaaca
agtttgtacaaaaaagcaggctgtttaaacactagtatcgattcgcgacctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataagtattttactgttttcgtaac
agttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatcccggtccgaagcgcgcggaattcaaaggcctatgaagcgacgatggaaaaagaatttcatag
ccgtctcagcagccaaccgctttaagaaaatctcatcctccggggcagctagctggagccacccgcagttcgaaaaaggcgccgacgacgacgacgacaagggctccgcggcgctggcggaggag
cagacggaggtggcggtcaagctagagcctgagggaccgccaacgctgctacctccgcaggcgggggacggcgcaggcgagggtagcggcggcactaccaacaacggccccaacggcggcggc

gggaacgttgcggcgtcgtcgtccactggcggggatggcgggacccccaagcccacggtggctgtctccgccgctgccccggcgggggcggccccggtgcccgccgctgctccggacgccggcgct
ccgcatgaccgacagactctactggccgtgctgcagttcctacggcagagcaaactccgcgaggccgaagaggcgctgcgccgtgaggccgggctgctggaggaggcagtggcgggcgccggagcc
ccgggagaggtggacagcgccggcgctgaagtgacgagcgcgcttctcagccgagtgacggcctcggcccctggccctgcggcccccgaccctccgggcactggcgcttcgggggccacggtcgtct
caggttcagcctcaggtcctgcggctccgggtaaagttggaagtgttgctgtggaagaccagccagatgtcagtgccgtgttgtcagcctacaaccaacaaggagatcccacaatgtatgaagaatacta
tagtggactgaaacacttcattgaatgttccctggactgccatcgggcagagttgtcccaactttttttatcctctgtttgtgcacatgtacttggagctagtctacaatcaacatgagaatgaagcaaagtcatt
ctttgagaagttccatggagatcaggaatgttattaccaggatgacctacgagtattatctagtcttaccaaaaaggaacacatgaaagggaatgagaccatgttggattttcgaacaagtaaatttgttctgc
gtatttcccgtgactcgtaccaactcttgaagaggcatcttcaggagaaacagaacaatcagatatggaacatagttcaggagcacctctacattgacatctttgatgggatgccgcgtagtaagcaacag
atagatgcgatggtgggaagtttggcaggagaggctaaacgagaggcaaacaaatcaaaggtatttttggtttattaaaagaaccagaaattgaggtacctttggatgacgaggatgaagagggagaaa
atgaagaaggaaaatctaaaaagaagaagcctaaaaaagatagtagtattggatccaaaagcaaaaaacaagatcccaatgctccacctcagaacagaatccctcttcctgagttgaaagattcagataa
gttggataagataatgaatatgaaagaaaccaccaaacgagtgcgccttgggccggactgcttaccctccatttgtttctatacatttctcaatgcttaccagggtctcactgcagtggatgtcactgatgatt
ctagtctgattgctggaggttttgcagattcaactgtcagagtgtggtcggtaacacccaaaaagcttcgtagtgtcaaacaagcatcagatcttagtcttatagacaaagaatcagatgatgtcttagaaag
aatcatggatgagaaaacagcaagtgagttgaagattttgtatggtcacagtgggcctgtctacggagccagcttcagtccggataggaactatctgctttcctcttcagaggacggaactgttagattgtgg
agccttcaaacatttacttgtttggtgggatataaaggacacaactatccagtatgggacacacaattttctccatatggatattattttgtgtcagggggccatgaccgagtagctcggctctgggctacaga
ccactatcagcctttaagaatatttgccggccatcttgctgatgtgaattgtaccagattccatccaaattctaattatgttgctacgggctctgcagacagaactgtgcggctctgggacgtcctgaatggtaa
ctgtgtaaggatcttcactggacacaagggaccaattcattccttgacattttctcccaatgggagattcctggctacaggagcaacagatggcagagtgcttctttgggatattggacatggtttgatggttgg
agaattaaaaggccacactgatacagtctgttcacttaggtttagtagagatggtgaaattttggcatcaggttcaatggataatacagttcgattatgggatgctatcaaagcctttgaagatttagagaccg
atgactttactacagccactgggcatataaatttacctgagaattcacaggagttattgttgggaacatatatgaccaaatcaacaccagttgtacaccttcattttactcgaagaaacctggttctagctgca
ggagcttatagtccacaaggtggttgataacctacgtcgacgagctcacttgtcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccatac
cacatttgtagaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatc
acaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaa
ggaagctgagttggctgctgccaccgctgagcaataactatcataacccctaggcccgggcaactttgtataataaagttggtttaaacgtacccgtagtggctatggcagggcttgccgccccgacgttg
gctgcgagccctgggccttcacccgaacttgggggttggggtggggaaaaggaagaaacgcgggcgtattggtcccaatggggtctcggtggggtatcgacagagtgccagccctgggaccgaacccc
gcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttattgccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatg
catcagctgctagcacttatcacggagcaggctgaggggagccggagttgggctgggagccattggcttttggagtccctggagctggagggggactgtccccagcctcctgcttccccccacaaaggg
cactgccgctgagtggggacggggacgatgccgggggaggaactggagaaggatgggaggtggggcctcctttgccctcccctgttgggggaagtgagaccacgatgtacttctggacactcccagga
ccagagggagcagtgctggggggtgcggtggtggcggtggagaccaacttgacgatgggctgcacgctggggacggtggtggtgacgggcgaagtggtggtggaacctgagccgggggccgaggtgct
gagggacaggacctgctgggtggatggggcgctggaggacgatgacattacaataaacttggttggaggaggggaaggctgtggtggggcagccgctcgtgcagacaccagtgtctggacaggaagtg
cgatggagccaggaaccttcagcaagccaggggtgcgagggccaggctgtggggcctgcgagagggtcagcgtgggccggggctgcgtgatggtcagagtggtcctgttgacctgctgagcctgcaga
gcagcctgggcccgagccttgaccacctgggagcagaggagggggcccaagggacccgaattctgcccgataggcgtcctgattgtcaggcggtgggcgcagctttgccagaacaggagcacagtgttt
cagcaggaggctctgcacatggtctgctccaatccggtcaatgttgctcagcacagggccgtccagcacactgcggatccgctccccttcctgctgcagccggggcagaatcagagtcttgataacatcg
tgtcccagctcagccaagcctgcgatggagccataacgagtcgtccagggcgtcttctcgtccacccagctcttggtgaaggtcttggtgatccgggactggatgttgttagtggttgtgctaaaatgcttgc
agatctgggccaccaggcgggcagcaaagtctcggagtgcccagtgattgtccacatctggtcgcaggcacaactgtctgctcacgatgcaggtcatcacagctggaatcagctcatggacgtatttttct
agatagagcgtggggttgtccatcagcgctttcaccatacgcatcaggtagatgagtagggccaggttgttctgaaccacgttcacacggacccccctccgagataaaggtactgaaccgtggcagcatct
gatacagtccagggtccgtggcaatgctttgcagggcttccgccctcttggcctcgcaggagcccacgcaggcctcggtgatctccttgtagtagagctgctgctccacagacaactcgtggatgctccgg
ggcttcagtcgcaaggggggcccccctccagcaagggcggcgccttcttctctttccctttgccgtcggctgtggtggcccccttgacctttgcccttcaggggtccgtcttcctcctggcctggcttggctgacttc
aggggttctgtggcttcagccttctgttgctctttgggagctgggggcgggttctcggggatagctggctggcagccctcgatgctcagccaatgagctttgaggcagacgtccaggggcacccggggcag
aggggtattgatgatgtcgctcagatcaacctccttctcctcatagaagtaaagctcccggcccccaccagaggcgaagcggaaaggaatgaactcctgggcgtggaagccatagagtggctcgacatt
ctttagcttcaaggcgtagtcaatgtcactggtggtgagcttctgccgcttccccatgtgcatgaacttcaaggcatcctgtgcgatctctttgatgcggtagctgacctcatccgttagcagctggcaggtctc
ctcctgaatctgggcgatgcccatggattcagccaccaccttcatggactccgagggcagcacagtgttgctaagcttcagcttcttctcctcagcatggtgatggtgatggtgcatgctcgagatcccgggt
gatcaagtcttcgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtccgtatactagt
atcgatacctaggcccgggccaactttgtatagaaaagttggtttaaacactagtatcgattcgcgacctactccggaatattaatagatcatggagataattaaaatgataaccatctcgcaaataaataa
gtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatcccggtccgaagcgcgcggaattcaaaggatgcaccatcaccatc
accatgagtctggcaagacggcttctcccaagagcatgccgaaagatgcacagatgatggcacaaatcctgaaggatatggggattacagaatatgagccaagagttataaatcagatgttggagtttgc
cttccgatatgtgaccacaattctagatgatgcaaaaatttattcaagccatgctaagaaagctactgttgatgcagatgatgtgcgattggcaatccagtgccgcgctgatcagtcttttacctctcctcccc
caagagatttttttattagatattgcaaggcaaagaaatcaaacccctttgccattgatcaagccatattcaggtccaaggttgccacctgatagatactgcttaacagctccaaactataggctgaaatcttt
acagaaaaaggcatcaacttctgcgggaagaataacagtcccgcggttaagtgttggttcagttactagcagaccaagtactcccacactaggcacaccaaccccacagaccatgtctgtttcaactaa
agtagggactcccatgtccctcacaggtcaaaggtttacagtacagatgcctacttctcagtctccagctgtaaaagcttcaattcctgcaacctcagcagttcagaatgttctgattaatccatcattaatcg
ggtccaaaaacattcttattaccactaatatgatgtcatcacaaaatactgccaatgaatcatcaaatgcattgaaaagaaaacgtgaagatgatgatgatgacgatgatgatgatgatgactatgataatc
tgcggtcgggctgataacctacgtcgacgagctcacttgtcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagag
gttttacttgctttaaaaaacctcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcaca
aataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttg
gctgctgccaccgctgagcaataactatcataacccctaggcccgggcaactttgtatacaaaagttggtttaaacgatatcgtttatgaacaaacgacccaacacccgtgcgttttattctgtctttttattg
ccgtcatagcgcgggttccttccggtattgtctccttccgtgtttcagttagcctcccccatctcccggtaccgcatgctatgcattcattacttgtacagttcatccatgccgcctgtactatggcgtccttcag
cacgttcatactgctccacaatggtgtagtcttcgttgtggctggtgatgtcgagtttgatgttgacgttgtacgcgcctggaagttggacgggtttttttcgccttgtaggtcgtcttgacctcggcatcataatggc
caccgtctttcagcttgaggcgctgtttgatttcgcctttcagggcgccatcttctggatacatacgttccgagcttgcttcccagcccatcgtctttttctgcatcaccgggccatcagacggaaaattcgtgc
cgcgcagtttgaccttgtagatgaactcgccatcctgcagtgacgaatcttgggtaacggtcacgacgccgccatcttcaaagttcatcacgcgctcccatttgaagccttcggggaagctcagtttcagat
agtcggggatatccgcgggatgtttcacgtaggctttgctgccatacatgaactgcggcgacagaatatcccaggcaaaaggaagcggaccacctttggtcactttcagtttcgcggtttgggtgccttcgta
aggacgaccttcgccttcaccttcgatctcgaactcatggccattcacgctgccttccatgtggactttgaagcgcatgaactccttgatgatcgccatgttgtcttcctcgcccttcgacactgctagcacca
tggctcgagatcccgggtgatcaagtcttcgtcgagtgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaatt
aaaggtcgatatccctaggcccgggagacagctttcttgtacaaagtggttgataaacactagtatcgattcgcgacctactccggaatattaatagatcatggagataattaaaatgataaccatctcgca
aataaataagtattttactgttttcgtaacagttttgtaataaaaaaacctataaatattccggattattcataccgtcccaccatcgggcgcggatcccggtccgaagcgcgcggaattcaaaggcctacgt
cgacgagctcacttgtcgcggccgctttcgaatctagagcctgcagtctcgacaagcttgtcgagaagtactagaggatcataatcagccataccacatttgtagaggttttacttgctttaaaaaacctccc
acacctccccctgaacctgaaacataaaatgaatgcaattgttgttgttaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattct
agttgtggtttgtccaaactcatcaatgtatcttatcatgtctggatctgatcactgcttgagcctagaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataa
ctatcataacccctagggtatacccatctaattggaaccagataagtgaaatctagttccaaactattttgtcatttttaatttcgtattagcttacgacgctacacccagttcccatctattttgtcactcttccc
taaataatccttaaaaactccatttccacccctcccagttcccaactattttgtccgcccaca

8.2.5 Baculovirus Genomes

8.2.5.1 EMBacY

cccactcgtatcgtcggtctgattattagtctgggaccacggtcccactcgtatcgtcggtctgattattagtctgggaccacggtcccactcgtatcgtcggtctgattattagtctgggaccacggtcccact
cgtatcgtcggtctgattattagtctgggaccacggtcccactcgtatcgtcggtctgattattagtctgggaccacgatcccactcgtgttgtcggtctgattatcggtctgggaccacggtcccacttgtattg
tcgatcagactatcagcgtgagactacgattccatcaatgcctgtcaagggcaagtattgacatgtcgtcgtaacctgtagaacggagtaacctcggtgtgcggttgtatgcctgctgtggattgctgctgtgt
cctgcttatccacaacattttgcgcacggttatgtggacaaaatacctggttacccaggccgtgccggcacgttaaccgggctgcatccgatgcaagtgtgtcgctgtcgagcgtcttcgaagcgcgtaac
caccacacccgccgcgcttaatgcgccgctacagggcgcgtcccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaagggggatg
tgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgagcgcgcgtaatacgactcactatagggcgaattgggtaccgggccccccctcgaggtc
gacagccgcgtaacctggcaaaatcggttacggttgagtaataaatggatgccctgcgtaagcggggcatttttcttcctgttatgtgaattcctgcagcccgggggatccactagttctagagcggccgcc
accgcggtggagctccagcttttgttccctttagtgagggttaattgcgcgcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaagcat
aaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagag
gcggtttgcgtattggttcgaagacttaattcctgcagggggggggggggaaagccacgttgtgtctcaaaatctctgatgttacattgcacaagataaaaatatatcatcatgaacaataaaactgtctgctta
cataaacagtaatacaaggggtgttatgagccatattcaacgggaaacgtcttgctcgaggccgcgattaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaatc
aggtgcgacaatctatcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggctgacggaatttatgcc
tcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcaccactgcgatccccgggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctgg
cagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttgatgcgagtgattttgatgacgagcgt
aatggctggcctgttgaacaagtctggaaagaaatgcataagcttttgccattctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaattaataggttgtattg
atgttggacgagtcggaatcgcagaccgataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcctgatatgaataaat
tgcagtttcatttgatgctcgatgagtttttctaatcagaattggttaattggttgtaacactggcagagcattacgctgacttgacgggacggcggctttgttgaataaatcgaacttttgctgagttgaaggatc
agatcacgcatcttcccgacaacgcagaccgttccgtggcaaagcaaaagttcaaaatcaccaactggtccacctacaacaaagctctcatcaaccgtggctccctcactttctggctggatgatgggg
cgattcaggcctggtatgagtcagcaacaccttcttcacgaggcagacctcagcgcccccccccccctgcagcccgggggatctgcagcggccgctccggaattctagaaggtacccatcagcaacta
tatattgatagacatttccagtttgtgatattagtttgtgcgtctcattacaatggctgttatttttaacaacaaacaactgctcgcagacaatagtatagaaaagggaggtgaactgtttttgtttaacggttcgta
caacattttggaaagttatgttaatccggtgctgctaaaaaatggtgtaattgaactagaagaagctgcgtactatgccggcaacatattgtacaaaaccgacgatcccaaattcattgattatataaatttaa
taattaaagcaacacactccgaagaactaccagaaaatagcactgttgtaaattacagaaaaactatgcgcagcggtactatacaccccattaaaaaagacatatatatttatgacaacaaaaaatttac
tctatacgatagatacatatatggatacgataataactatgttaattttttatgaggagaaaaatgaaaaagagaaggaatacgaagaagaagacgacaaggcgtctagtttatgtgaaaataaaattatatt
gtcgcaaattaactgtgaatcatttgaaaatgattttaaatattacctcagcgattataactacgcgttttcaattatagatataatactacaaatgttcttgttgcgtttggtttgtatcgttaataaaaaacaaatttg
acatttataattgttttattattcaataattacaaataggattgagacccttgcagttgccagcaaacggacagagcttgtcgacagttttcatttttggcagtttggtggtttatacacatcttgggatttgattaa
aagatgcatgacttcaccccacacacttggagtcgtggttgtagaaatgcgtcagagccttgattaaacatcggttttcacacttggcacacgaaaccacgtttgcaatgtaacagtttagtataaacctcttc
ctgtacataccgccgcgcggtacattgttttctttcaaatcgttcaatattttttttatacacggaggcatattaatcttgttgttcagcatgtcgacgacatttttctttccttctacgcattctttcatacaaatgtttc
tcttgctgcgttttttaaaattggaatttttaatttcattaatagtgggcggagccgttaacagcatgtacaatcgcagtccggaaattttgaccaaaaccccgccattgtagaataccgtgtatggagttagggt
caatttatcaataaaatcatcgggatcaattaaatatgtagctttcttgtctaaacatgacgctctaatgagcggactccacacgttatacgatgcattcgccatacttcgcggcttctcgcaccagcggtcct
gattcgactaacggtcgattggatggcttaatagttttaacgtcaccgacatactcggtaattttgtagtcggcagttcttttgggcgtgatcgttgtcactacgggggtcatgtcatcgcttcgtaaaaacttatt
acaatcaaagctcgtgccggaacgcgtgcacagatcgtgcacgttgcgtcgcaagctatttaattcgcgcttatttgatctctatcatagtcgtgttgagatgcgacaacgggcgcattgttaattcttgacgt
gcgatacataatatattattggataaagttgcattaaatgaaactaacttacaagatggctagtttgttaaaatacgcgctgcgcttgactcgggaatacaaagaaaacattattccacactttgatcacttga
ctcgattgcgcgatttaatcgacggcatgattaaaagcgaggatgtacaacgttttaatcgcactaatcgcaatgatttaatttcggcttgcatgcaaatcaacgttcggacgtacatgcccaacgccacgat
agatatgcgcaaacaacccaactgtatatattttcgaatttgccaatattgccacttggaggccgacgtgccttcgcccgacgatcattcggtgtacagatacttgtgcgtcgcgtgcggcacgccgctggt
catcgaccacccgctcgacgtgttcggccacacggaggaaggcgtcaacgaactgctcgaggtgcagcgagtcaacgcgggcggggagttgtaggcgtcataactatttattaaataagataatttaaa
aaatcgccgttaatatgcaaatcaaaactgtactattggcttttgcgatgtttgccgcgttgaacgcgcaacacgttttagccgcttgcgccgagaccggcgccgtctgcgttcataacgacgagtgttgtag
cggcgcatgttccccatatttaattattgcttaccacaataagtttagtttgcgttaaaaaacaatagttgttgcaaaatcgacacacgtgtatcaaaatggctcgcgttaaaattggagaatttaaatttgga
gaagatacgttcaacctgcggtacgtgctagagcgcgaccaacaggtccggttcgtggccaaggacgtcgctaacagtttaaaatatacagtttgtgataaagctatacgtgttcatgtagacaataaata
caaatcgttatttgagcagaccatccaaaatgggggccctacctctaacagcgtcgtgaaaaggggcgacccgctgtatttgcagccgcatacagtgctcattactaaatccggcgtgattcagctgattat
gaagtccaaattgccttacgccatagaattacaagaatggcttttagaagaagtaattcctcaagtgctctgcaccggcaagtacgatcccgcgatcaaacaacgggaggaggagagcaaacagttggt
gactaagctgattgcgacattcaccgagcacacaaacgcgctgcaagcggtggtggcgcaaaaaaaccgaggaacttgttaaaaaacaagagtttattgaacgcatcgttgccatcaaggacaagcag
atcgaggccaaagatttgcaggtcacgcgcgtcatgactgatctaaaccgcatgtacaccggcttccaggaaaccatgcaaaagaaagacgaaataatgcagaaaaaagacgcgcaggtcaccgatt
tagttgccaaagtggtggatttgtcggatcgcgcggttcaatacccgcggacaagcgcaaacatccggtgttgtgcgtgacgcgcgacggcactacgtttacggctattaccggccaaaagacgtacgt
ggaaaaccaaaagcataaacgtaacatcaacgttgccaacattgtggtggagaatatccggcctaatcctaccgtcgattggaacaacgccactgatcggctacaagctaaacggagcaagcgaagc
attagttttggttcgctggaagaagcgcaacaatttgaaaataggataaagtatttgttaaagaatgcaaataatattaattaaattaataaatcttgaacgtaattttgtctttcaattttgtgacctctggctttt
cgaatctatctatggcttcctgcggcgcaatacccagggtgtgcattaaatatctgcacaccatgtaaccggtgcgattaataccgtgtgtgcagtgcacgcccaccaacatgccgggacacttttctgtaa
attcttttaccgtgtcaataaattcttgaactatgctttcaggcggcaaagtctggccaggtacttgaattttttttgtataacaggcccgcccgcaaaaaatgcacaccatcataatatttagacgtgttggttaa
atcgattattgctccaatactagggttttgttttactatctgttctgcggtccacacatcttcttcgctagtcacgtacgcaaacaactcgggccgcaaaggtgttttaaaacatatcagattagaatcttttataa
cttggccgcattgtaaatagttgtgccaacgcgcgggaaacatttttacacgtgctttcaatcttatctcataaacgcaactaaatatgatccttcaactcgctttacgagtagaattctacttgtaacgcacgat
taattatgactcataagctgatgtcatgttttgcacacggctcataaccgaactcgctttacgagtagaattctacgcgtaaaacacaatcaagtatgagtcataagctgatgtcatgttttgcacacggctca
taaccgaactggctttacgagtagaattctacttgtaacgcacgatcagtggatgatgtcataaacttatctcatgttttgcacacggctcataacttgagtcataagctgatatcagaagattattaacgtgttt
aggtatgactcatttgttttcaaaactgaactcgctttacggatagaattctacttgtaaaacacaatcaggggatgatgtcattatacaaatgatgtcatttgtttttcaaaactaaactcgctttacgggtaga
attctacttgtaaaacacaatcgagggataatgtcatttgttattctcaaatatgatgtcatttaaatttttattatgcaagaattgtatgttgctgtgtcatttatgaatgttttatcgacaacttgtttatttgtttcattg
acttgttgttgttcgtgattatgattaattataaatcttttgttggggtcaacatcgatagtgtcatacgcatttacaacgtctttatctgcgttattgtcacattgataatgtaacggcatatcgtattttacatctatttt
aatacaattactactatctccatcttaaaagtttcttaatcacatgaaataaaaaaaacataaaaagtataaaaatgtttatttatttattccatcatttttatatatacatttcatatcacattaattttcaaacccaaa
attaacagtaacatttacatcacatgaactagcagcagaagttccggcataaacataattttgaatactatgaacgcgattttcaataaacagctctgaaatagtcgaatagacaagtgagcgaaaacgct
cgttcagctcggctggcgcactgtcgtcacggccaatttgctgattaaccgtaatgccacgttgcggcaccaatgtcctcacctgatgctgcaactcttgaatctgttcattgagacgagtattttcttctatca
gctgtcggttttcagaagctacttgatcactgtgattttttcttcttccaattcagcttccaaagcagctatctttgcctgattctccatatttatgtcacaattattgtcgttgtcttcgacactaatcacgctccgtt
cttgtagcactttaataatgtttttatactgctgcacgggcggaactctttcttgagttttcttgtaaaacttacacgtcacaacagcgttaggctttaaggaacgccacatttttacctgggtattgcacaatgga
cagcacacagtggaatgcggaatattcatcgcgttgctttgcagattgatgacgcatttgaaacaaaaaccgtggtcgcacgtagtcggaatcaaaaaaggacaagtatcgttagattgttgggaatatgtt
tctaaacacacactacactcaaatttaaaattaatcaccggctcgttgtttaattccgccatcattcgctgctgttccatttgtatggagcggagctggttttgacgctgccttctacgcacttgatagcgcaggt

tttcttcttgttcttcttgcattctgtctcggcggtttatgctgtccaaattgtcttgcaactgtctgcggagatcttcaagcgtaggacgttcatatggcgtagaccggtggcgattgttggtgtcccttggcattttg
cttattggcaggctctccgtgttcacttcacaacgaactgtgcttacgagttggtgagtgcctgcttttatagctgcacgcctgagtgggggacagataacagaaactgcagcctgtgatatgataaatttaaa
atgaatttttttgcaatgcaaaaaagttcacttttgcctgacactccatatacagtacaatctctacaaatcgtagactattttattagaatagtctacactgtacgatacgctcccaatatactactacactatc
aactttttgcattacaaaaaagttcattttgcctggcaagttcccccaccactattgtcttatcagtcgtgcagtactgataaacagtataaatacagctgccgttctactcgtaagcacagttcaagcctca
cagcctagtgaacagtatcctaccagcccagccatgagtcgccaaatcaacgccgccactcccagcagcagccgccgccacaggctgtctctcagccgtcgccgcatcaactttacaacatctcccg
aagcccagccgtcttcaagcagtcgcagccagccgtcttcaagcagtcgcagccatcgccgtcaggagcggcgtcaggagcagcgtgtcagcgaagaaaacgtgcagattatcgggaacgtcaacg
agccgttgacgcgcacctaccatcgtcagggtgtcacgtattacgtgcacggtcaggttaacattagcaatgacgatccgctattaagtcaagaggatgacgtcatactaattaatagtgaaaatgtggatc
gtgaacggtttcccgacatcactgcccagcagtaccaggataacattgcgtcggagacagctgcgcagagggctctgcaacgaggtttagatcttgaggctcagctgatgaatgagattgccccaaggtc
tcccacttatagtccatcttattcgccgaattacgtaataccacagtcgccagatttgtttgcctcgccgcagtctccgcagccgcagcagcagcagcagcagcaatcagaacccgaagaagaagtaga
ggtttcgtgtaatatttgttttactacttttaaagacactaaaaacgtaaattcctcgtttgtgacttcgattcattgtaaccatgctgtgtgtttcaagtgttatgtcaagataattatggacaattctgtgtacaaat
gtttttgcagcgctacttcatcagattgtcgcgtgtacaataagcacgggtatgtagaatttatgcccattaacgtcactcgtaaccaggattccatcaaacagcattggcgcgagcttttagaaaataacac
ggtcaacaatcacaccacggacttgaactatgtggagcaattgcaaaaagaactgtccgagctgcgagccaagaccagccaagttgaacataaaatgaccatgttaaacagcgactacattatgctta
aacacaagcatgctgtcgccgaattagatttacaaaaggcaaactatgacttgcaagaatctaccaagaaatcagaagagttgcaatcgactgtgaataatctgcaagaacaattgcgtaagcaggtgg
ccgagtctcaagccaaattttcagagtttgagcgcagtaactctgatttagtttctaagttacaaactgttatgtctagacgttaagtgttgtaaataatataaatagtataatatattaaataaattttattttatta
atctatataaattttatttcattaatctatataaattttattttatttttagttttggttagcggtacatccataatctgatataggaacacaatttttaacatctggatcaaattcaaatccttcagggcagattaactct
aatttttgatttaaaccgacacacatgtagaaagcgtcacatctagtaggatgcggaaaatttactccaattggtttgttatagcaactgaacccgtcgtctgattcgtcatcgtcacctcttttagatttcaaaa
cagtaaaaataattattgaaaacaaacaaactattatgaatattataattaacatgttgggttttaacatatttattgaaacggtatataaatttagcgatcatacaatggagagatatcatccgtatcaacgcg
cgccgtggagcgcatcaatatttaaaaaaaattataacaatgttaaagttgcttcatacgtcgagataccttacacaacaaataatgcagacgtcgaaataacttacacaacaaatgcgccttctcgccca
gaatctaaagcaaggagaagattagattttaccaacatgacgccgctgccgtgttgtattgacgacggcagttttgcaaaacctttactgtgatatatttattaaataagtttgctttaaaatgagtttttttcaa
atcttcgcgcagtcaataaattgtatcctaatcaggccagttttattactgataatacgcgtcttttaacaagcactcccgccggtttcacaaatgtgctcaacgcgcccagtgtacgcaacattggaaacaa
cagatttcagccgggctatcaactatctaacaaccagtttgtgagcacttcagacataaacagaattactcgtaacaatgatgtccccaacatacgcggagtatttcagggcatttcagaccctcaaataa
actcattgagccaattgcggcgcgtggacaacgtgccagactttaattaccacaccaagcagacgcgatccaatgcagtcaaacaaaacttcccggaaaccaacgtgcgcacgcccgaaggtgttca
aaacgcactgcagcaaaaccccccgtctacatagttacatgcaaagcttaaaagtaggaggaacgggcatactcttggccaccggcggctatttcttgtttagcgccgccacactagtacaagatataatc
aacgccatcaataacaccggcggaagttattatgtgcaaggtaaagacgccggagaaatcgccgaggcctgcttgttattgcagcgcacttgtcgtcaagaccccaatctaaatcagtcggatgttacca
tttgcccattcgaccccttgttgcctaacaatccgcccgaactaaccaacatgtgccaaggattcaactatgaagtagaaaaaacagtttgtcgcggcagcgatccgtccgccgacccagattcgcctca
atacgtagatattagcgatcttcctgcgggccaaacactcatgtgcatcgaaccttactcgtttggtgatttagttggcgacttgggtctagattggttactgggcgatgaaggtttagtcggcaaatcgtccaa
cgttagtgacagcgtcagcggcaagttaatgcctatcattttgttgataggcgcggtcttgtttttaggtttaatattttattttatctacagatacatgatgaaaggaggaggaggaggtggtgttggcgcagcaa
cgtcgccaactcccattgttatttctatgcaaaaccccacaccaacaacggcccctcgataataaaagacaaaaataatataaaatatatgtataattaattaaattcaaaatatatgtataattaattaaat
tcgaatttttttatatttacaatttagtttttgttccgcaaacgttatagcgtcggacaacggaaccagaccctgtaatattaaagctaacaattttaacaaattattgtgcaatgtagtgctctctcttcggttcactt
taccgattacaaacatgtgatgcttaaatctattatatttttgaattacttgactaacgtctacatctttaatttcgccagaaatccaataaaactctttgtttttcttagctatagtcaatcgctcctcgtttttgaaa
gacaatactataaaattgtgaccctttacattatctacattctgagtcaaatactgttcgacaatgtgcatgctgccgtcctccttcttaaccttttttaaattttctgcgttattattactcgcaatattgtcatgata
tttataattattaaacaaaagattagcgacgctgctgtatttgtacgtgagcgtactttttttgttaacaattaaatttaaattgtccaccacatatttgtttgggggattgtcgggaaactgcacactttccgaata
ctttaatatttgactcacatacggcgatacaaagaaattattagatgcagtctcaatttcattactctctttacgactaagcataataggcaaagtaaataaatttttatcttgatacatttcgtacaacttgctca
aaagaaacccacattttctttcgcccaacgattgtaacaaagtcacaaatgtggtttgcgcgtaatacatatctaaattaaaatatgaagtcagagcagctttaaacgtgtggtgcacatcgacgaaatggc
attttttacaattttgtgcagccgtctcgccgttgcacacatcttgagaatgaggaatttctatgccggtttctttaaccaaattgtacgaaatcataaacctaattttatcaaaagttaccacaaacacgcgatt
atctaccatgtaataattgtttgtatattcgtaaaccacattgctcacgtacttggcaaatataatttcaaacggcttcacttcactttttttaaccacaaacatgtaataaccagtttcggacatatggtcggag
aacctgttggaattgtagtcgttgtcgtcgaaacgcattaaatacggcgcaaaatcattagtaaaataatgcgtaatttcttgagttgaagcgaccgtgcaaatgttcgtgttgtgattaattgtctgttcaaggg
ttgcacagctttgaattgtgcttttcttgtatttaggcctcaatttaattttgttaaactggcccaccacaccttgtgaattgtccaagtattcgtccagcttccgtttagttccagttgccgatggttggttcacacca
gcagaatgctcaaaagattccgcatgataagcagaactgggcgatggttgctccacaacaggctgctcaagggattccgaataataagcagaactaactgcttcttcgagattatcagtggtcttgagcaa
acattccactaaatcattatcagttaacgagttgacgcttgccaaaaagtttgaagctgccgcagtctcgctgtcagatatcaccgtgtcggctccatccggggtgggatggttataataacttaaatagtcgt
tgggttgtttatcacaaaactctgaatagctgttgtcgaacgacgctcgggacggcgtcgaagcgctggtgtacgacgcgttaaaattaatttgcgtcatagtcacttggttgttcacgatcttgtcgccgcca
gtgtcaacttgcaactgaaacaatatccaacatgaacgtcaatttatactgccctaatggcgaacacgataacaatatttcttttattatgccctctaaaaccaacgcggttatcgtttatttattcaaattagat
atagaacatccgccgacatacaatgttaatgcaaaaacgcgtttggtgagcggatacgaaaacagtcggccgataaacattaatctgaggtcgataacaccgtccttgaacggaacacgaggagcgta
cgtgatcagctgcattcgcgcgccgcgcctttatcgagatttatttgcatacaacaagtacactgcgccgttgggatttgtggtaacgcgcacacatgcagagctgcaagtgtggcacattttgtctgtgcgc
aaaacctttgaagccaaaagtacgaggtccgttacgggcatgctagcgcacacggacaatggacccgacaaattctacgccaaggatttaatgataatgtcgggcaacgtatccgttcattttatcaataa
cctacaaaaatgtcgcgcgcatcacaaagacatcgatatatttaaacatttatgtcccgaactgcaaatcgataatagtgttgtgcaactggaaacccgctcttcatagtaacaagtttctatacatgcgag
ccgtacacccgctgcccgtgtgatcgtacacgataggcttgcagctggcggagtccaagtcaaattcgtgatttaattcgcaaaacatttgcactttatgcggacacagataatacgccgtgcaatcatagg
gatcggcgttgaggccaaaataccctttaggacaaattttatcgctgtgacactttagatgaatttgattcatttttttaaatattaatactttaattaatattaaaaatataattaataaaaacattgtatattattta
attatttattaaaattatatatattaaaccctgatgttacagaatttgagtttattttgtcaaaatcttcaatggtaaatttgaatttgcttaacagtaaattagacgtttttgtgctcattgtcgatcgcaatttgttggt
aaaaagcgaatggtccgttagcgatttgcaatattcaaacaccccgtaatttttaaaaatataataatgtaaataggcatgttcgacggtcagcaacagagttaaatattgagtggcttcaacctcagttatg
ctaaaatcgctcaaatccatatcgtaattgggttgtttatttgactcgtaatcatcgctgcgcgtcgtgtttttgcgcttagaggcgccaaacgtcttggctatccgcacgttgcagttgtcgttttcaaagtttact
aaagcctcggtgtcaaactcgcggcttagcattttaacgatagacggtctgtccacggaacacagcagcacaccttcgttttccgtaaacaaaatgggttctccgggaatgcttgtgtcgttggtttcagtga
ccacaaactccattttgttagtaaaattagtaacaaggggatggaatctattgtgtataaaagctaatgtgctaaacaccacggcggcaatctgccgagtgtctgctttgtttctaaatatttccaacagcggtt
gagtcatttccagggtgctgagcatagccatgtatttgctgagcgccaatttaatttttagagtttcatagctttctaaactgcttagattttttaaatcaaattcagccaattcgtaggtcttttggatttttttcatcgc
tgtttacaatctcggtgaccgttcgaactttgttgcatttgatacgtttcatttttttatacacttatctacaacatttgcctttgaggcgtgttcataaagggattagctcgcatggtagcgtttaaaggatttacaaa
accccccagggcgtttggaggtacattacccccccgagctgtttccattactgctagattgaacaaatataattattaaagcaataattactactacagccaagatggtcaagaacatgtttggagttagcctg
tttaaatagtttgtggacggcgcgtctgtgctagtcgtagcgccagtagtgggatcagtgtaaatcatattattgtaccgagtcggggattaataatttgagaagcgtgttctgaacccacggattcaaatcgcg
aagggatttcaactcagacgaactttcgtaatcacctcttattaaaagatatgcaggcaccgtgtttgagaacacgtgtttgaccaaaaaaactttttcgttcttatcaacaacatacacgtcgcccgtcaaa
cggaccggcggcgaactctgccggtacacgagcaggttgggatgaaagtcgatcgtttccgtggcgccaaacaaacgattcaccgccaaaatacttaatatacgctgcttgggcgcgtagaacgcgttc
atagtgcctttcaattgcaaaaagtctttgctcacaaaaatgtacgacgattcatgatcgtacatgacaggaaactggttggttatttccatttttatgtacagcttgcgataattgtctactgagtagcgatcaa
gaataatctcgcgagacatttcaccgctcatagcggggttttcttcaatcacaaacttggtaattattttcatgatcgagcttttggtcatgtattcgtttagcaaatcaagcaaatctggcggaaactgggcgtc
gaatatgtagtcgcgctgaataaacttgataaacggacagttgccatttaattctttatctatctcattaaatatatcgttgggttttcgggtcacgattttagtgctgttcaccaccgtgtactttgatttgaacagc
ggcagacctttgtaaaagtttttcagcataaacattgaattgttgattttgcgagtctttgcgttagaagggtgcggcggtagtatattattatcgacaaagtgcttagccatgtcctcgcccagtagatacagtc
ggaacggatgcatgttgttatctaatcgcggcgccgcacacattcgcacaccattccaatccacgtacgcatcgtcgaacaagaaaccggcactgcccgccaacacgcagccgtagttgttgatgtttac
gaatattttatttttaaacaccttgccaaattcagcgtacaggaaagccatcatcttgcggggattggacgtgaacaggttcgtggcgtacacgggcgtgccgggttttatgtacatgcgcgtgtcaaattcta
aaatgtccacagttgtgcgatcgcacacaaatctaaattgtggctttacaaacttgaaaatgtttgtcgacgtgctgcccaccaaatgtttcaattgcatgctttgcaagtaattgaaatagttcatgagcacg
gcatcgtctagcaaatcaaacgtgcaattcaagtaattgcgaataaatgccataggctctgaaggaacatgttcattatcttttaaatcaaaatagctggtcaaaaataaatatttaaaatgatctctttccag

agtcaacaagttgccgccaccactcattttgcgttgcgtgcttatttatacgatgtcctgcacgctggacattttgcgtgcgtggtggccgtcttccacatgttaacgcagcacgcgttgcatatagcgtattcg
caacattcttttggctttaaaaatctttcatccgtagtatttctttacaaacattgcattcgtacaaaaccaaattgtctacaaacaaattcataacttaacagttttattaatagtttccagcttttcaataaac
aacaggcaatttgatattaactcgttcacgtacatttgcatagaatttataattttaaaacaacacgcaaaagagtttttaattagcttcaacgcgtctcgcaactgtttcaagtacggataaaacactattaca
ttgccgcgatcacgcacgccgtacgcgctttcaatctctttaatacattgaatatacgttgggcatcaaaaataaaatggtgacagcacgcgccgtcgtcatcgggcagacggtcgggtttggcaatgtcg
gcacattgatcatagtaataaaaaaatttattctctattctatgcttgtacatttccaaatgtttatcgtcgatcattgcaaacgcggccgagcgcacgccgcattgcagctttgcgtcaatttgtatgtcggcgtt
gtacatgtgaaagaaaacaaagttatcgatcattacctgctgagcttctgcgcaagcggacgttgcctcgcacgaatatggcgatgacgcacagtttgacgtcattatattttcctggacgttcagcacgtga
ctgctggttcttatcatgttgcgttgcccgttatcaattacttatactatccggcgcgcaagcgagcgtgtgcgccggagcacaattgatactgatttacgagttgggcaaacgggctttatatagcctgtcccc
tccacagccctagtgccgtgcgcaaagtgcctacgtgaccaggctctcctacgcatatacaatcttatctctatagataaggtttccatatataaagcctctcgatggctgaacgtgcacagtatcgtgttga
tttctgagtgctaactaacagttacaatgaaccgttttttcgagagaataacatttttgacgcgccaaggaccgggggcaagggtcgtgccaaatctttgccagcgcctgccgccaactcgccgccgtcgc
ctgttcgtccgccgccaaaatctaacatcaaaccacctacgcgcatctctccgcctaaacagcctacgcgcacctctccggccaagccgttggagcacagcagcattgtaagtaaaaaaaccagtcgtc
aacagaaaagatggatattttgtgccgcccgagtttgggaacaagtttgaaggtttgcccgcgtacagcgacaaactggatttcaaacaagagcgcgatctacgtatgcatttcatgagcgatttagaacg
cgacatcatgaaagccacgctaaaattttccaccaattacattatgggctacataaacagcaaagatatgcgcatgacgggcaaatttgcaagcaggcctgtgaaatacaagaagactatggagcacat
gagcgattcgcgatgcaccacttgcaattatagattcaaagacaacacgcgcgaatggtttttgtatgttgtggttcacatcgaaaagccgttggacgatcctgaccgcatagacatctgctgtcaaaaatg
ctatctatatcacaacgttccaaaaacctcgtacgaaatatatccttctatcaatttggtcgacctcagctatttagccagagagaggtttttttaccaatacattttccctgtaagtttagagcacacgacgga
agttaaagaattgagaattgacgatcacaactgcaaagtgtttgaaataattcggcgcataattcgaaatcacaaagagcccaacgagcgtattcaaacaatcgatcttagcaccaccggcggtctcgtt
ctcagggaaacatacaccaatatagtgttgcagcgctaccgaagcatgtgcaccagacccgacgtggtggatgacgtcaactgttttatacttcaagagccaagcgaaatgatggccgctctgcaagac
aaccgattttcgggtatcaaaggaaccgtgtttgccactgttaaagtcaaaaagtttacccaggtattggacggcgcaatcacgtttcctttgaaacctacaactaacaactattgcaaattgtgtaaaaaaa
ctaaactgtattataaaaatcccgtgttatattgtaccaaatgcgggtttaccaacgtttatcatttcctgaatactctaaattcatgtattactttgaagctatcaaaagctttgaaatgcacaacgaaatgatt
atatattatgatttaaaaatgtataaaaagttaattaatattgtaaataacaatgtctaaataaaaatgtatcgttataatataatcagattaaataaaaagtattgttataataaatgtaaatgtattgctataata
taatctaataaatatttcattatatattgttgtttattttaagtcctatataacatggcggttttaacagccgtcgatttaactaatgccagtaggtatgccatacatatgcatcgtctcgagtttatttcgagatggcg
caccaggtttccgcatatactaatcgactacacgctgcgtcccgcgtcaagtgacgacgattattatgtgccgccgaaattggccgacaaagcgctggccgtcaaactggctttcagcaagcgcgggtgt
gaaagcatgagctgctatccgtttcacgaaaccggcgtcgtgtccaacacgacgccgttcatgtacacgcagacctccgagactagcgtgggttacgcgcagcccgcgtgctaccacctggacagggc
ggcggccatgcgtgaaggcgccgaaacccaagttcagtctgccgagtttagatacacgctcgacaacaaatgcatcttagtggactcgctgtccaaaatgtactttaacagcccatacttgcgcaccgaa
gaacacacgatcatgggcgtggacgatgtgccggcttttaacgttcggcccgatcccgacccccttgtttcccgaacgctttaaaggcgaattcaacgaagcctactgtcgacgtttcggcagagagctatt
caacggcggttgttcgtttagatggtggggaatcgctcatcggattcgtgctgggcgacacgatctttgtcactttcaaaatgctggccaacaacatttttagcgaactgcgcgatttttgattacaaagcgccttc
cagcatactaccgccacgacccaacgtggactccaacgcaatactagcgcaatggcgcagcgtgcgcgataacgccacagacttggaatttgaaaaacttttaacaaaaatccaacactaaacgatt
tgggcatgatcgtcaacgggtcgcccgtccagatcacgtacacggccgaaaccggctttactaaaactccgatcgcatacaattaccgcggtaacgaacgtgcacgtgtcgaacatttcgaggcgctcg
accgttcaatcagcgaccaagacttggagtcaattataacatcgtttttggaagattacgcgctcgttttcggtattgctacggacattggctttgatatgctaatgtccgggtttaaaagcatgttgaaaaaaat
taacacttctctcattccggccatgaaacacatgttgcttagcacgactcggcgcgtgaccgtgcgcatgttgggagagacttacaaagccgcgttagtgcactcgctcaacgtgatcgccatcaaaacg
ctgaccgtgactgccaaagcgttaactcgaatcgctatccaagccagctccattgtcggcatcgtgctcattctattgacgctggcagatttggttttggcgctatgggacccgttcggttacaacaacatgtt
tccgcgcgagtttcccgacgacatgtcgcgcacgttcctgactgcgtactttgagagtttcgacaacaccacgtccagagaaatcatagagtttatgcccgagttcttttcggaaatggtcgaaacggacga
tgacgccacgtttgaatctctatttcatttattagattatgtggcatctttagaagttaattccgacggccaaatgttaaacttggaggagggtgatgaaattgaggattttgacgaatctactttggtggggcaag
cgttagccactagctcgctatacactcgcatggagtttatgcagtacacgtttaggcaaaacacactattgtctatgaacaaagaaaacaacaattttaatcaaataatcatgggtttatttgcaacaaacac
aattgtggcgtttacagcatttgttatacacacagaactcatattttttatattttcgtaatcttcctaatgatcacatttttattacataatcaaagaatcgtacgaatattataaaacaattgatttgttattttaaaa
acgattcatttacttggaactgcgtttaccacgacgagcgtctgaatcgagttcaaaagcttggcttttcagcttaggctttagtgagtctggaagatccggaacaatgtcgccggtcaatatggattgaatttc
tgagatcttagtgttaagatcggtcaattgagcgggcaaaccgtccaaaagctggaatgattcttccagcccgtttaactgagtttgaacactgtcaacctttgtgttagtttccgtaacggcgtctaaaatttg
cgtcaaaacgttaggctttgacatgattgtaaataaaatgtaatttacagtatagtattttaattaatatacaaatgatttgataataattcttatttaactataatatattgtgttgggttgaattaaaggtcccggc
atcctcaaatgcataatttcatagtccccccttgttgtaagtgatgcgtatttctgaatctttgtaaaatagcacacaagactccaacgcgtttggcgttttattttcttgctcgagcgcatgctgtttaattttatcgt
acggcaattgcacagtcccgtagacggacatgttgggtcgccatttttcaacacggacgccgttcagcaccttcatatgtcccgacagctgggacgcctctcgaattcccgtcaccggtaacagccattcg
ttttcgccaacacgatgaaacgccgtcacgaccgaaaccagctttttgccggcaaaaattggagctccaatgcgcagtttttgcttcgtttcattattcacgaccgccgctggcaaattgccaaacacaaa
gcgtttattgcacacgtgaaaattaaaacaagtgttttcgagtttcaacgtgagcaaactgccgtcttcgctcaaaacgctaattattgtgtttaacacgagttgcgggaaaataatcgaactaacgacgcca
ggaaattggtggtactggcttagcgtttcgtcgaacacttcttgacccggttctaaaatatgcacaaacgcgtccacattgtcgacttgctctattacaatttttatttgttggatcaattgcatatttaatattatac
aattccataacgtccgtttgatttaacgtatagcttgcaaatgaattatttaattatcaatcatgttttacgcgtagaattctacccgtaaagcgagtttagttatgagccatgtgcaaaacatgacatcagctttt
atttttataacaaatgacatcatttcttgattgtgttttacacgtagaattctactcgtaaagcgagttcagttttgaaaaacaaatgacatcatctttttgattgtgctttacaagtagaattctacccgtaaatcaa
gttcggttttgaaaaacaaatgagtcatattgtatgatatcatattgcaaaacaaatgactcatcaatcgatcgtgcgttacacgtagaattctactcgtaaagcgagtttatgagccgtgtgcaaaacatgac
atcatctcgatttgaaaaacaaatgacatcatccactgatcgtgcgttacaagtagaattctactcgtaaagccagttcggttatgagccgtgtgcaaaacatgacatcagcttatgactcatacttgattgtg
ttttacgcgtagaattctactcgtaaagccagttcaattttaaaaacaaatgacatcatccaaattaataaatgacaagcaatgacaaaataatattaggcaataaatttttaacatttatttaattgtgtttaata
ttacatttttgttgagtgcactagttacattgtgcaatttaacatacaacactttgatttttatcataaataaaaccgttacaaaaaatttccaattatacttttcggatttgccccagctcgattctgtagtaaac
tctaaagccttgtagatcatgcttttgtttttcgacgcttcgttgtaatataaaaccttttccgaatttttgaacgacgacggcaatattaaactctgcacaaactcttcaaattgttcgccatcaaactgcacact
gtccacgtaagctacgacgtcgtactcgtaaagtcccgtgtcgttaatttcgtgagcaaacggcacaataactttgttaccaatcgggttgtactcgaaagatagcacttggttgctgccgttctcaaatccc
ggcttcaacacgcataccgccacgtagtagtcgttgcgttcgtcgaccaatttgggcaaacagtatttctcgtaagcttcaatagactttttatcggtataatcatgactctttaaaattttagcaaacttgctttt
caacacgctgccattttggcaacttaccgaatagtgttcatctttaaaatacttgatgctatcgtggaatccatccatttgatcgctgtaatcgcgttctagttgatcaaattgttcatcgactttgattttattctgt
cgcgcaaattgttgttcttgcgcgtaacgcttcgtataggacccgaaatgttaaacatcatgagaacgggttttgtcaattgcgtgttcataatcttgttaatgtacaccaactctctggtttgtttgtccacctga
caatctcgaataatcgtctgggacacgtcgatttctaccggaaaaattacacacatttttgcaatggtaaagctcaaatgctcacttaatacaagcaacgttgaatatttataatataaacgttcacgttacga
cgcgcgtgtaattttcgcgaccactatcgcgctgccagagataacacacacacgatcactcatgttacagaacgtgaccgctccgctcgtgtatataaaaatccgtcgccaaaattaaatatcaatcaac
aagatgattaactactacgtgtacgccaccgacgattccttgtcgacaaactcggactattactttaacaagaacgcgttacaaactttagaacaatttcagaatgaaatcgaaaacatcagtagttgcga
caaagttttgtatttgcattggagtgcttattgccgccaaaaagaaatcggagacgtcaagagtcgttatttgcgccgcaatggcgagggtactgacaccaggcccagcgaattcataaagtggattcaca
agaatataaacttgctcgacggcaagttgaaactttttgtacatggtgaccgacggccaaatttccaaaaatgaggcaaacgtttgcaaaaatttgctaaacgaaaaaccgtttagtttcgaacgcattgtgtt
ttatgccatcaataataacacggaacagattgatttgtccgttgcctcggcgtttgtcaacaactcggactgcaagatttatcgcaacgatgaaatggtcgaatgggtcaatttaaccaaagaatttaactac
gatatcattacaaccgaaaattttatatcgaaaaaagatgaactgctgtcttttgttcggttcaaatttatcaattccatgccgacggacgccaacgtgttaaacgaggtggacaagctgaagcgtctccgcc
agcggctattcagcgaaatcaaacaaactaacaattctagtatgaattttgaccagattaaaaataaaaatgagtttgtaaacacgttcaagtcaacagaattttataaaacgttgtataacgccgatgttct
caattttgataaaattatcgacagcaccattagcacggccataaattatttgcacaatcgcaacaagtcatacgcttttgatgtaatgaaaaatttacactatcagaacaagttagcgagcgtggacgccga
aaccgacgatgttgttgccaacgacgacgacgaagcgtacgattacagcaacgtagaaaatattcgattccccgattgtattttagccaacgatagcggtgtgccggccatattgttaacacattacaactt
gtttgaaacaattcagggcagtttgaccaagtttaaaagccgactcgagtttcctcttttatggagccaaaataaagagattaaaaattcaattgaatattgttataatttggaaagtttgaagcagttgattcag
catggaacgcgtttgtcaccgcgcagcagacgcccttttacgggcgccatagtgcccaacgagcaattcgacgaatacaacgattacgtgttggcgtgcacgtattttgacgctaaaaaggtcgcgtttaa
cgcgggccttatgtactatttgctgtacaagcacataaacgacgcagaatacattgacgataatgtaaaggactacttttaaacgctatgtcatttaccgcattaacaacacggaatgcatgatcgggttcag
caacctggctatggagccgctgatcaaagtaaaattgcccacggctttgtggtatgtctcggaaatttcaacgttgcttttaaacacgacaatcaacatttggaaaggaaaagctgcggcagttcgcgcat

tttgccgaggacatgttgcaaatcctgcaatggtgcgactacacggacgtgaacgttgaggcggtaaaaaaacgagcgtattgtctaaaacgtatcaacatgtttaaacgcatgtccgttttggacgctgtg
gagtggattgcaaacagagcgtttgaatgcaaagacaaatttattatcaacaagttgaccaacgcggacgctttgcaagatttgaaattttttgaaagtaaatcataacggagttgttgacgaacatgttttaa
acgacacgtccataaacgcggagagatatttgtatttctatcatataattgaagattttgacaagtacataagcgttgtagataacacgatgcgccccgcgtttgtgcttgaagaaggaaaaactttttacgac
agtttattgaaacaattgcaaagtgtacatttcaacggtcaagaaataacttttgagaaatgttctcgtttggattttaacagaattttgtcactgcacaaattgtacattgaatgcgtcaagagtttaaacaaat
atcctacgctggaagaatatcaaaactatgtatacaatcaaaaacatgtcaaatttaatagaattgccatttttcccgaaaacattctacaaaacctggccgcagtgcacaatgagtacgcaaacaaaatt
gttaatttgccagttgaggagtttattgttcgcgccaacaataccgtaaaccggattacgcgcattcaaaacgaacgcgttggcagcccttttgcaagccgaagaaattgacaagcttattaaactttctgaa
caacgagtcaatatttgtcgtaaataaaattgtataaacaaatgtataaataaaagaattatatataatataaaggtttattttatttccaaataatatacacagtcgtttaaatgatcgtgtttggtgggcgacga
gttttgcgatccgcagcaaaaggtcttcagggttcctgtatgagttaaatagtaaccaaaataggcaagtgttttagcgtcgtttctaaagtggtgactctttagcaaactttgcaatcttttaacataatcaaag
tatgtaggattgtcaaagtttttatgtttcaagttaaaactttttgcagttggtgtgctcgcgttttaaaacagaatcgcacgtttcccgactgtctgagctaaaatcacagtacgcacatttaaatgtccataatct
aaatatattccgcgatcgaccatagcatcgatttcttgaccactatagttgtgattgcagctgtgcgttttgaatgaaagccgtcgcttgtcggcgcattggtagcgcgcgcagttggagttttccacggcaaa
cgccacgtacgcgttgttttccgtagacagttctttgttgcagaacgtttcgtcgcgcggcacagaaaccaccaccaaattttgtttgaatttgaccacgtaaaagcccataggcgcgttcatcacatacatc
tgccgttgcatttgcttgtaatgcgcatccgttttgaccatttcgaactggggcgtgcctttcttgttaaccaacagcgcggtgtgtttcacgcgatacttgcgattgccgttccccaactcgacacgcatctgct
ccacggtcgtgtctcggtaattgtacggacattttatttccactgggatccacgttccgtcggcgagagaaaaatacgcatcgggcgaagcggcgtgcaaacccaaggggctaaagaacatgccgcaat
ccaaaaccgtttcaacgaccgcgctcataattttactttctatgatttcgcgcatgtggtctaacatgatgccgttggtttctttgacgtgactttcttgcgcgtttccaaacaaaagcgctggtcgttgtaacatg
ttgccggacgagttggacgatgcagacgctgtgttgcgatccaaacgcaatagcgtccacagctcatttttgctttgtcccctggtggccttttcgacacgtaaaatttctgcacgatccaaaggttttggctg
aatatctttgtttgaacgccaatgagccaactgcgcttgactcaactcgatcgttttcacataattgttaaatttatattttttaataacagcttttgctctgaggtcaacgacgcaaacatgatgtcgacgtgct
agccttaccaaatctaaagacctaatggtaaaatgacgggaggttccaattgcatttcatcaaacaatttttcatatactttggtacttcccaactgtcactatttaaaatattaaaataaataatggattttctt
agttcttgatcttgaacaggtgcagtaattgtgagaactttattttgtaaatcgtattcaatagtttctatagtgtctaaaattacgttgttcaattgcaatttaaaagacgatatcttaggcctgttgatagcgacttc
gtttttgtttttgtaaagtaaatcgcttctgattctggcacattctagttttaatttgtaatcggtattgacaaaataggccatgggatacgtcggattcacgaatgcgacaataaaattatcaagaggcgtatcta
atatgggatacaatcgtttttgttcattttctaatactacaactgcattctctaatctttccaactcttgaaaaacggcactctccccggggttctgcaaagaagaaacatctgctgtggtgcgcagcgtttgctc
ggcagcgatggctttcttgtattgctgcaacttttcagactttttgggggttttgttacgccgcgctctcgcaacgtcatggcatggctgccaccctttttgttggtgttttgtcggacatgcttattttttgtcagc
gttcccgtgagtatatttttgatttcgttaaccaaagtgtacacgaggtccagtttttgtctgtcggcttcgttaaaattgccaccgccgccgccgcctactaagttttcgatagcggctagcaaattgttgagcgt
accgttgatgtttgtgacgcttgacgctaaattggtcaaaatagagttggtgttgttcaactcgttgcgcagcgtttgctgcaccgtgttgatggcgtccaacaattgttcagctaatttatctaacatggacgtca
agtcgggtaacgagtctgtaatctcggcaagcaatgtgtttaatctggtcaaaataacatctttagcggtttccaacgccgcgaggagctgcacgttttgattgcgcacgcccgtcatcatgttggacaattct
aaaaattgattagtttggttgagatttatttggttgacgttgctcatcaacatgtcgttttgatgacgaatcttttccagcgcgtcgaatatttgtcgacgccctcgcctgggagacgatgaacgtgaccgcgatct
agaccgggaccgagatcgcgaacgaggtctgcaatgaggagagcgcgaacgaggtctgcaatgaggagaacgtgatcgcgagcgtgatcgtgagcgtgatcgtgatcggcgacgacggtaaaagta
gcacaaataattgtttgcaataaatgtgctcaacaactggtcggcgaacgagcaaggcgttcgcaaaaccagaacggccgctccgtaaatgtccaccatgtatttgaccgtgttatgacacgatccgccc
gcatcgttgatcatttttctatgtttagatggaatggtgaacaaaagttgttgaaacgcctgcaaaggcatttttaaattttgataaatgtagtccgtgtcgatccaaaggaccgacgcgtcgtcgaacataacg
ttattcgtcggcttcatattttgaaatattgtagtgactatgtgatttttagtatattaacttagtgataagcgtttattgaacgtttagcacgttgttaactaccccgtttttgttaataacacgttaatttttcgtcga
gattggatatttttttgccatttttacaactgcgtcccacgttcgaatcgtgttattttgcaagttgctcaacatcataaacgtgtccgttttaagagaatcaatttcgtctaccacattggaaaaatgcttttccat
agattcgtacatgaccactaatattgattttatttcttgtagctcgttattcaagtgtcccgcgtacaccaaataacccaccagacaaatgctaaacgtggcccaaaagttcattgttgaaatcgtacaaaac
ccttattattatttattcaggcacattaaatcctttatatttttaaattctccaaaattgtatcagtggtaaaggtaatttttttattgtattcagactgcaaatcttgtatcagttcttgcaacttgttaatgtctggcgc
aggcgggttggcggccgcatcgttgctatgaacgttcaaaatcgaaatagcctcgttgagactctcctgaatcgcacacagttgcgtttttacttcgcaaagcggatcaaatactttgccttgagcgcccata
agcaagtcgctaataatcgatttaagggtgtaatactctttaggacgctttttattatttaaattgtaatttgacagatatttaaataaaccaaacacatgaatataatttttattgtttttaatcaatttgtgcgaggg
cgccgcgttggtccacacggccgaactgccgttgctaagaaaaggtgacatgactttggcggcggcggttagttccagatagccatcgtgatcgccgtccgaattggtaatgatgaccacctccagactg
ctgtcgggagagtacatgaagttgcgcgtgttcgatgtcgactgagcgtccgtgttcatgatcccgtttttataacagccagataaaaataatcttatcaattaagataaaaagataagattattaatctaaca
acgtgccttgtgtcacgtaggccagataacggtcgggtatataagatgcctcaatgctactagtaaatcagtcacaccaaggcttcaataaggaacacacaagcaagatggtaagcgctattgttttatatg
tgcttttggcggcggcggcgcattctgcctttgcggcggagcactgcaacgcgcaaatgaagacgggtccgtacaagattaaaaacttggacattaccccgcccaaggaaacgctgcaaaaggacgtg
gaaatcaccatcgtggagacggactacaacgaaaacgtgattatcggctacaaggggtactaccaggcgtatgcgtacaacggcggctcgctggatcccaacacacgcgtcgaagaaaccatgaaa
acgctgaatgtgggcaaagaggatttgctcatgtggagcatcaggcagcagtgcgaggtgggcgaagagctgatcgaccgttggggcagtgacagcgacgactgttttcgcgacaacgagggccgcgg
ccagtgggtcaaaggcaaagagttggtgaagcggcagaataacaatcactttgcgcaccacacgtgcaacaaatcgtggcgatgcggcatttccacttcgaaaatgtacagcaggctcgagtgccagg
acgacacggacgagtgccaggtatacattttggacgctgagggcaaccccatcaacgtgaccgtggacactgtgcttcatcgagacggcgtgagtatgattctcaaacaaaagtctacgttcaccacgc
gccaaataaaagctgcgtgtctgctcattaaagatgacaaaaataaccccgagtcggtgacacgcgaacactgtttgattgacaatgatatatatgatctttctaaaaacacgtggaactgcaagtttaac
agatgcattaaacgcaaagtcgagcaccgagtcaagaagcggccgcccacttggcgccacaacgttagagccaagtacacagagggagacactgccaccaaaggcgacctgatgcatattcaaga
ggagctgatgtacgaaaacgatttgctgaaaatgaacattgagctgatgcatgcgcacatcaacaagctaaacaatatgctgcacgacctgatagtctccgtggccaaggtggacgagcgtttgattggc
aatctcatgaacaactctgtttcttcaacatttttgtcggacgacacgtttttgctgatgccgtgcaccaatccgccggcacacaccagtaattgctacaacaacagcatctacaaagaagggcgttgggtg
gccaacacggactcgtcgcaatgcatagattttagcaactacaaggaactagcaattgacgacgacgtcgagttttggatcccgaccatcggcaacacgacctatcacgacagttggaaagatgccag
cggctggtcgtttattgcccaacaaaaaagcaacctcataaccaccatggagaacaccaagtttggcggcgtcggcaccagtctgagcgacatcacttccatggctgaaggcgaattggccgctaaatt
gacttcgttcatgtttggtcatgtagttaactttgtaattatattaattgtgattttatttttgtactgtatgattagaaaccgtaatagacaatattaaatgtaataataaaaattgtatcattattaatgtaataataaa
aaattgtatagtttttaattgtatattataaaataaaatttttaaaattacaaagtttttaatatctcaggcgagacttgaactcacaattctagcaagatctatacgatccagagcgctacgctctaccttgagc
tactgagacgatatgattatgttccaaatagcgagtgtgttgagattaataaatgactgcagtagacgcaagttcgtttctcataccacaggcgtttatgttttgttgtaccctgaaaaatccgtcctctccccaa
tccgtgccccaagtgtttttaaaggtccaatatggaatgttgttttcaacaccataacccactaaaagaaccgcatggtttagaccgctgttgaaacaatattttataataccctgtttatagttaacaatgtcg
gcagcgtctatggccataggaatagggccgacaaggcgtaacaaatctttaagtttttcctcgtacacggtaatgtatctataacaatcttttacttgaactagaaacttattggagttcatacggcaattgtta
ttgtctgcttcgtatggatagtcgctttccagctgtacgccgcccattttaatgatggcttcgaacgcggtaccgcatgctatgcatcagctgctagagtttaaactaacttcgtataatgtatgctatacgaagtt
atggtaccgcggccgcgtagaggatctgttgatcagcagttcaacctgttgatagtacgtactaagctctcatgtttcacgtactaagctctcatgtttaacgtactaagctctcatgtttaacgaactaaacc
ctcatggctaacgtactaagctctcatggctaacgtactaagctctcatgtttcacgtactaagctctcatgtttgaacaataaaattaatataaatcagcaacttaaatagcctctaaggttttaagttttataa
gaaaaaaaagaatatataaggcttttaaagcttttaaggtttaacggttgtggacaacaagccagggatgtaacgcactgagaagcccttagagcctctcaaagcaattttgagtgacacaggaacactta
acggctgacatgggaattagcttcacgctgccgcaagcactcagggcgcaagggctgctaaaggaagcggaacacgtagaaagccagtccgcagaaacggtgctgaccccggatgaatgtcagctg
ggaggcagaataaatgatcatatcgtcaattattacctccacggggagagcctgagcaaactggcctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagca
atagacataagcggctatttaacgaccctgccctgaaccgacgaccgggtcgaatttgctttcgaatttctgccattcatccgcttattatcacttattcaggcgtagcaaccaggcgtttaagggcaccaat
aactgccttaaaaaaattacgccccgccctgccactcatcgcagtactgttgtaattcattaagcattctgccgacatggaagccatcacaaacggcatgatgaacctgaatcgccagcggcatcagca
ccttgtcgccttgcgtataatatttgcccatggtgaaaacgggggcgaagaagttgtccatattggccacgtttaaatcaaaactggtgaaactcacccagggattggctgagacgaaaaacatattctcaa
taaaccctttagggaaataggccaggttttcaccgtaacacgccacatcttgcgaatatatgtgtagaaactgccggaaatcgtcgtggtattcactccagagcgatgaaaacgtttcagtttgctcatggaa
aacggtgtaacaagggtgaacactatcccatatcaccagctcaccgtctttcattgccatacggaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttat
ttttctttacggtctttaaaaaggccgtaatatccagctgaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacggtggtatatccagt
gatttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtgatcttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaa
agttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatga

tggtgttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagatccggatatagttcctcctttcagcaaaaaacccctcaagacccgtttagaggccccaaggggttatgctagttattg
ctcagcggtggcagcagccaactcagcttcctttcgggctttgttagcagccggatcttctaggctcaagcagtgatcagatccagacatgataagatacattgatgagtttggacaaaccacaactagaa
tgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggag
gtttttaaagcaagtaaaacctctacaaatgtggtatggctgattatgatcagttatctagatccggtggatcccgggcccgcggtaccgtcgactgcagaattcgaagcttgagctcgagatctgagtccg
gacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccatgtgatcgcgcttctcgttggggtctttgctcagggcggactggtagctcaggtagtggttgtcgggc
agcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgccgtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatata
gacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgatgcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgta
gttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaagaagtcgtgctgcttcatgtggtcggggtagcgggcgaagcactgcaggccgtagccgaaggtgg
tcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagcttgccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtc
cagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggcgaccggtagcctttgaattccgcgcgcttcggaccgggatccgcgcccgatggtgggacggtatga
ataatccggaatatttataggtttttttattacaaaactgttacgaaaacagtaaaatacttatttatttgcgagatggttatcattttaattatctccatgatctattaatattccggagtaggtcgcgaatcgatac
tagtatacggacctttaattcaacccaacacaatatattatagttaaataagaattattatcaaatcatttgtatattaattaaaatactatactgtaaattacattttatttacaatcactcgacgaagacttgatc
acccgggatctcgagccatggtgctagcagctgatgcatagcatgcggtaccgggagatgggggaggctaactgaaacacggaaggagacaataccggaaggaacccgcgctatgacggcaataaa
aagacagaataaaacgcacgggtgttgggtcgtttgttcataaacgcggggttcggtcccagggctggcactctgtcgataccccaccgagaccccattgggaccaatacgcccgcgtttcttccttttcc
ccaccccaaccccccaagttcgggtgaaggcccagggctcgcagccaacgtcggggcggcaagccctgccatagccactacgggtactccagataacttcgtataatgtatgctatacgaagttattcta
gcaccatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatct
cagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctcca
gatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttg
cgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctc
cttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaag
tcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctc
aaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagg
gaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatggctcgagacggatttgggctaccaaacaacagtctacgcgccaagttg
gaactcggaagagctatacactacacactacgctgtcgatgcgttactgaaacaaggcgtcgatcccaacaaaataattgtgggcgtcgccatgtacggccgcggctggaccggcgtaacaaattatac
gaatgacaactatttttccggcactggcaacgggccggtgtcgggcacgtgggaggacggtgttgtagattatcgtcaaattcaaaaagatctcaacaattatgtgtacacgtttgacagcgccgctcaagc
gtcgtacgttttcgataaaaagtaaaggtgatttgatttcgtttgacagcgtcgactctgtgttaggaaaagttaaatatgtcgaccgaaataaattgggcggcctgtttgcttgggagattgatgccgataacgg
cgacttgctcaacgcgataaacgcacagtttaaacctaaagatgaactgtaaaagcaatatattgagtatcattttagttttgtgttagttgttattaggatgtcgagcgttacaaagcgcccacgagcaaaa
cgcatacgtttaccattagaaattatagacacgattctgcagtatttggatccgattttacatgcaaaagtcgtgggattaacaacgcgtgtaaaatgcagacttttgcgcgacaacaacgttgaagattattt
aaagctgacacctgccagttatcaccccacgacggatcagtttatttgtaattatcttggaataacaaatcagccaatggcgccatatttagtacctttgcttcgtttggtaaagcgagctgtgtgtttttaac
aaatgcatacccgaggacgtgcgcatcgtcacgttgaattggccgttaccattgttggaaaacttttttgtccaaacaatttttatggtacaagttggcgagaaagttgatcgaacacgaacgtcgaatggatc
gctgtgtaacaccgtctacagtgcaaataaacttatacgacgacaatgaagattatttaaattgttttaattgctttaattgttgtaaagataatttggttagttttaattgttgtatagtagactgtaatatcaatgat
atgaataggtgtccagatttacaaattgatgtttatttagacgacaacataataagtttgtatttatatttttttgtttagaatttacagaattgtgaaagaataaaaaacatttttataaattatttttttattttttatgtca
tccattacaatgttagagaacggtggcggcggatcgttgtatgaaaacggaacaacttgaatactttgtaaatttgaaacgtaataactagtcaaattatttttaaagaccacattattgcaatcacaatgta
caaagcgcacacaaacgcgttaaacatcaaactgttatcaattagtttcgttttcaaacactctaaacccataacaatagtaacactcataactgaaacaaacaatataaccatataacacgatttgacat
attgataatccgtgtacgtgcccattgtgtggcccaacaggataagcaattcaacaaatatcataaaaagtacacaacagtgaattaaattgttaacaaaaacaaaatttgtacccatttgattgtccaaaa
acaccacataattcaacagcacgttactaaagtacatggcgtggaacagcttcatttcgtaaattacttcatgggcttctgtggttttgagcatcgtccacagagtcggagcgtttagcaccacggacatcaa
aaacgacaacagcaaccagtattgtaaatatttcaattcgctggagtcattgaatgttaccatgtatcccaccaacatggcggtgtaggtgcacagcgtcacgttgatgtacgacaacgcaacatagggaa
acttcatttttgtgtcggtcgacagccttgtgtatttttgtgcaacacgcaaacaaacccatttttaattcttatatcagtgcacggttcgattgattaaatacaaatgcttgctaaaatttattgcagtgcgttttagtt
tataaacagatcgcgttaagcatcttgtgcgataaattgtatgtatatagggcgtgcgctcgtgtctgaatttagtgacgcgttaacacttaactgtataccatgaaacccgaacaattggtttatttgaatccgc
ggcagcaccgcatatatttgcgtcgcctctaaacgagtacatgttgagcgactatttgaaacaacgcaatttgcaaacttttgcaaagaccaacattaaagttccggcggattttggcttttatattagcaag
tttgttgatttggtgagcgccgtgaaagcgattcattccgtaaatatcgtgcaccacaatattaatcccgaagatattttcatgactgggcccgactttgatttgtatgtgggcggcatgtttggcagtctataca
aaacgtttatcaaaaacaaccctcaaaatataactttgtacgctgcaccagaacaaatcaaaaaagtgtacacccccaaaaatgacatgtatagtttgggcattgttttattcgaattgataatgcccttca
agactgctttggaacgcgaaactacgctaactaactttagaaacaatgtacagcaaatgccggcaagtttatcccaaggccatcctaaattgaccgaaattgtttgtaaattaattcagcatgattacagtc
aacgaccggatgctgaatggctgttgaaagagatggaacaattactgttggaatacacaacgtgttctaaaaaactataaaagaggcttttggggacaaggtgatttagaagcgacccatatattgtcgaat
atagaacaccatgaaactgattatcctgctgtttgtcgtggccgcgtttgtcacactcgccatgggcaaaaaatcgttgaaagatccattaacaattactacggcgcattttgaccatcccgaggggctgaa
acacaagattcgaagaccttacaacaaacacgaagacagattgattcaatgtaaaatttatttgtaataaattatatttacacacaaatatgtttattttattacatactattatttgtgatgatgacatcatatct
caaacttgttttaccagtagaattttacgcgtaaaacaagtttgaaatatgatgtcatttatttttgcacacagaatattgaacccgttatgatgtcatcatttgatagtgtaaatttttattaaaacacgttgtgcaa
tttttttatctatttttcgtttcaattgcaaacactctgtagttttcttgtttaacatgactgccacttcttttatgttagtctttcgacgcaatagttccaaattatctaaagcgatatccaaacaagtcttggccctaa
ataaataacccaatttcaaaaaagtgtgtgccaaatcgcacgcttctacaacttttaaaatgctcatgttgttatacagagtagttggggtaagtttcaagtatttgggttaatgcgtctatttgttgcccttgggt
catgtcgcctgtagtgaccgctcttattgtgtttcgttcgcgatcgttgttcagaaagaagccgcgcaatcttaacgctgcaggcgcgtctccgcgtggatatcgtgacgcgcgactccaaacttttatatgcg
taccttcacgactcagtacacaccacgagttggtgaacgtttctgcggtccacacttgaccaccagcgtgtcgcctgatgcacgaatcggaattgacggttatgcaccgtcctacgcctggaaagaagca
taccggctgagcagtgccactttctttaaatctaaatagcctaaaaacagaagatggtaatagtatatttttcaatgtggtatcgtatgacacggaaattttcatgacacccatgttgactgttacgatttctgtta
cgtccgggtgcgactcgagcaacgggtacaaaatcgctcgataccgttcgtgactgagttggttgacattcgcttgaaacaccatgtccgcgtcggcaaactcggtgtggtcgctgcggccccagaaaaa
tttgtaatccacatgacgcaacgataacatgtgcacgttaaacggctgcaagcaagcgttcgcgacggtgcggtcgctttgtctaaccatgcccgtgtcggcggtgtgtcgattaaacacgggcagtaacc
cgtcatcggcccggacaattgcatccgttgattcccacaccatttacggttggttggtgaaataagcgtcccgatgtgcgttgcccgctaatcggtccaccgagcaagggtcgatcacgcaaatgtcttcga
gtctaaagtgtctgcggtaaaaatcattgagcgccggatgatccagacgaacttggccgtctgcgcatggcgcccgcggaaaaaaactctcgtcgtacattacgtcgcgcacggtgcgcggacggcaat
acggagtttcggtgtcggcgttatagtcgctcacgtacccgtcgtcgcacacgcaccggatcgaagcgctgttgatattgtcgatacggccgtgaggcgcgcagcccacgggcacgttgcagtcttcgtac
atgttgagctgcgtaacaagtccgggccgtaagcagttgcacaaaagagcgaaaccagtttcagtttcggccaacaaccacacacccgtgttggggttacacgatcgggcgcgttctcgatccaaagcc
aagcagtacgcttctcccgctcgaataagcgtctcgtgttcttggccgtcatctccacgcatcttgactattgtatcttcgtcaaacaactggcagttggctaatccttcgcggcacagatcgcaatcgccgtg
tgtgaagcacggcgtcaaagttttgtgacattcgtgtgcattgccttcgattactatttcggacggcggttcaatgagcggcgcgcgtgtgttgtcaaaaaccgtaataggataacgcacctcttcatggtgca
catctatcaagtctacatacgtgtacactattgcaataatcactagaagaaacaacagtataatagcaaaatgcattttacttattggagaatgtccgagtatttttgtacgtacgagatatttgcggccaacg
cgattatgtgcgcgctcatgccgtgtttagcaattttcttcattatagaattgttatgtaaaaactttaacactatgtataacgtcatgtacgtggtttttgtttgcaatttttcaattgtttgaaaaaggctttgatatcg
caaaaacaattgcggacaaatggcatattagaaaaatttttgttatttgtgaataatgcgtatcaacatttaacgttgtacgttgaaaaatatcaaagtataaattacgctatcaaaattgcagctatttactata
atatttatgtattgtattattctgtagtatttttacactgtttgttatttttttaatgtgcactcaaaaataataatcaaattatttgttgtactaggaatagatgtggcgttgtattctttaattattaaataaaacaaaatg
atataataaattttggtattttatttatacacaatttttttggtaatacttgtcgtaattatcaactatcacacatatttttttcgcataacagttcaatgttgtcactagtgtattctaaactctccaacacttccatgtcc
gtgttttgcagctgatttcggtccaacaccaccacggccaaacaatcgccattgcaacacatcacaccttcaattacactgtaattttcaaataggtattataaatgtatttcacgtcccgctttttgagcgcg
ttaggaacgagcaatacattagccgtcagatgcatgacggtaaaaatgaacgcggtcgacgccgagttatatagaccccggtttatttttgcgcaaccagtcatttcgtgcgtcacaccacgcttttcaca

ctcagttaagacgatcatgttgaacttttggcaaatacttattctactggtaattatactaatcgtgtacatgtacacgtttaaatttgttcaaaaatttattttgcaagatgcatacatgcacattaacgagttgga
agcgccgccgttaaattttacgtttcagcgcaatcgaggcgtagattgttcattgaaccggttaccgtgcgtgacggatcagcagtgtagagacaactgcgtgatctccagcgctgcaaacgagttgacctg
ccaggatggcttttgcaacgcctccgacgctttggtaaacgcgcaggctccggacttgatagaatgcgatcccgcactgggtttggttcacgttttctccgccggcggagattttgtcgtgtctcaaacatgc
gtgagcacgtacagagacttggtggacgatacgggaacaccgaggccttatctatgcaacgacggccggttgaatatgaatttgaacacggtgcaattttcgccggacgcgtgcgattgttcgagcggcta
tgaaaaaatgctatttcgccaaacggcactggctcgcacgatccccgtgtgcattccgaatcgcatggccaatttgtacagacgcgtttacaactgaaatgtaagagattctgtcaacatgagcaagtacg
cgcttttacaaaaaatgatcatcaacgagttgctgtttttgaacgacaacgtgaattacgctaccaataaactgtttagcaaagatcaagctaacggcgagttacacaaactgttggcgatgttgttaaatta
caaaaagtccaacgaaaacgcgccaaatattaaattcgacctaaaaaatttgtcgtttatgctggagaacaaggacaagatcgatataattcaatttgacgacataaaaaactatgtacaacctgcgata
gtaaacttgtttgaatcgcacaatcgatcgttgaataactactcgaccgaactgagcacgctactggaagacggcaatgaaaatctggtgccaaacatcaccgacattgataacattaagttgtcgcacat
gcaactggcgcgcttgttatgttacacggccgtcgttgagtcacgcaacagtaaaccctggaaagcgatatttaacaacgacacaaacgttttgacagactctttttcaaccatatcatgaatattttgaata
tgattaaaaccaatcaaggatcgctggctcacaacgtgtcggtggtataccatattgaaaatatccaaatgaatttacaaaataaactcaaacctcgctcaatcaccattgaaactgtagacaaggacaa
atttgaaaacaagcattccgacattgaagtgtgttacgttatgaataacagtttacatccccaggacgttaacagccaacaaactttaatgtgctcgtcgtttgtcgaattaaacgttttaccattttgtttgtata
atgacacactgcctgacgaccagtctatgagcgtttttaatttgtacaaatttgaacagacaaaaaacaagtcggtgtctaaaccgtcgcgcttgggcaacgtgatgtttgttaacagtttaatcgacaaacc
gataacaagagaaactatagtaaatattatcaattcgtaccataacgcctgtcaaaatctaaaaagatctggccacagagtcgttggcgattacagagcgtatgaacgggattacaaattagccgctttag
attttatcattttgatgctcgtcaccagcattacacatcgcactctgaaatacaacatgctaaacatacacgaaaaaatgtttcaagaactaaaaacggtcgtgtgcaaacacaacgccgccaaattgtac
gacatattaatcaattacgacattaacaaagagccattgaacaattttcgatacaactacgaagtgctgtaatcgtgtttacactagaattctactcgtaaagcgagtttgtacgatgacatcattctacaaa
tgacatcatccctcgattgtgtttacaagtagaattctactcgtaaagcgagttcggtttgaaaaacaaatgacatcattctacaaatgacatcatccctcgattgtgtttacaagtagaattctactcgtaa
agcgagtttgtacgatgacaccatttttgcacaaacatatgagtcatatattaatcacgtttatcaacttgcacaaacaaatgagtcatacattaatcatgtttatcaacttgcaaaaattaagttcaatgtacc
atgtttaaacatgtttataatcacgtacaaaaatgacatcatttcttgattgtgttttacaagtagaattctactcgtaaagccagttcaattttgaaaacaaatgacatcatttcttgattataaataatatcaaat
gtgagtattgtcaataaaaacgcttgtttaaattaatttgttttatttaataaattttggcatatttttactttgtaaacataattgcatacgggttgcagttgagttttccaactgccgccacgtctaaaagttaaatt
tacacaaggtttcatcgtgttgtcctcaaaattgataggaaatctgtaattgtaaaattgcaaatacactttataatcatgttaatataattttgaacaccataaaattgtcatttgctttcattgtaagaaaatcc
acacaatgcagtattaattcttttcgatgttgctgcacttgtttgactcgaattagagttgcatagttttcaagataactatcaaaactttgcccgcttttgggcgcgtacaaccatttaatttgagagtcgtaaaa
cggctcaaaacaatacgtgtcggttgacaattgtttccattgcacgctttgcaccacacaatcatgtccatcaacagcaaataattgtacgcgacgtccaacaggtgggggtttaacattgcagttactttc
atataatcaattttgccagcacgtttcactatactgtcaacacagtttccataataatgaggatacgctttttataatttatgagatttatgttttttgcaaacgggcagttgtagcgatccgtcgttttccacacag
tacaaaacgatgcatttatagcgccgtgcatggtattcaatgttgtgccaaaacgggtattcgtccgtgttgcaaacaaattggttgtcgtggtgccacacaaacagcaaagcaaattgtcgtttagtcatcg
cgttgataaacgttatgttgacggtctcgcaattcaaacgcttttatatattgaaagttatctttaaccgataaacgtgctattgtgtacgtttattaaatcatataataactgaatgcaataaaaaaattaaagcat
atgcatcagtgacattgcctatgctgcactatggataattattcggtgcaaaatttttacaacaacgaccgcaaaccctctcaagccgacaacgcttcacgatggtaacatcaagaaatcggtgtacgaaga
tgtcacatatattcgcaagctgatgtgcaaagaaattatgcctggcgagcacgatcacaagttttacaattacggttataacaaagaaaacaaatataaataaacgtataaaagtactatgtattttattcaa
ctagcataagatttaaagatgaaatattttctatctgctacctttctaatcatagtgtttatgtacttgatgtatttttgcatttcaatcgtagtgaacaacgcacatgttcaacaggatctgttttaccattacaatta
cgtgcccgacaccctattgaatacggtacgagttcacaaattaaaataatctaaagtctaatgaagtaagacatatacatcgcgatggagtgcccgtttcagattcaagtttgtattagcgatcgctttttttgc
gttccctcacaatttggtggaaccgcaaagcgacgtaggcaacaaattaatcgaaaacttgatcgtgtacgtgcccacagacgacgaccggttgtacattgacaaaaaacaatttccaaaattcaattct
gtgctagtgtacaggcacgagcatgatgttaacatagacagtcgatcgccaaaaaaaaacagcgtccgccaccatagtgtactggaacccgttggtgcccatcaccgaaatcggtgcgggcgagacgcg
cgtttttagcgtgctcctgaccaacaacttgttttattgcaacactatgatcgttcaccatgaaaaccccaaatgtcccatcgagtttacgtatcccgaaacggacatgcaatctgcgtgcagcgcgttacta
aaaaaccgcaacggtcaaagcgttccgcctccaataaaaagcaatttgcgaccgatcgcttgtgaaattccgctgtctcattttaaagaattggtggaaagcaatgacttttgttatgtttcaatctggaaac
gtccaccatggtgaagatattgtcgttgaaacgcatttttttgcatctttcaatatcgcaaacagcccgcgcgctatgtaatcaatctgccgcacgaggagattgataatttgtacaacaaattgaattgggaga
gaacgcgccggcttatgaagggcgacgtgccatccaattgcgctacagtaaaccgctccagtctcaaatacattaaacaggcgcagtctttgctcggcattcccgactatagccaaacggttgtggacttt
gtcaaaatgttccagaaaataattttccccctatcaattggtacccaacgtaatcataaaattgaacaactttgaccagatggtgtcgtctgcgccgaataaagcagagccctacaaaaaaaattcggctgtttt
gcaagaacgacagcattgctatcagctccagcggcattgtgcccatcaacatgcccgattttagtccgcccaacacgtttgattattccgactacgctaatcggacaaacattaattttgtaacgcaacgc
gtcttaacggacggaggttttagcagcggcatcacggtgacgcccgtcaagtacaactattatttgtagttaagcacatcatgaaaccgacggcggccgacattatatcgagagcgacgggcggtcgcgc
cggcaacaatatcgtcgacatcatccaagcccacaactcgccgactgaaggcgatcagctcggccaatttgtaaatagaaaccgatcgctaattaaagaatttgttctagtagtatgcggctttttaattttt
gttatgattgtgttattctttatgttattggttgtgatcttgttaaaccaagaaacaattactgtacaaaaacaaaaatatgaaactactttattagaaaactatgatattagaaatagaaatgcaacaatataatt
acaaagaaaacaaaggcaataaatgattctttttttcaagtttacgcaattgtggcggtggtgctggtagcaaacattttgccgttgagttggtgttttgtaattttctttttacgttcgttcatcacgttggccaaaa
gcgatggcggtcttgcatctggcaaaggtttcaatgccggtttaataaaatcaaattcatctgttttgtccatgctgatggcattagtgttggtgtaactgtattcaaacgcttcgataaatatttttgtagcaactt
cttttccaaatgttaacaaatgatgttcgtccgtgactggattgtctagattttccaaacattcattgtaatgctgaagcattgcgcgcggagaactctttaaatcggggttaattttgtttaaacgtttaatggcca
cctcgatgatctctttgtacgatggaaagtagttggtgctcttgtttaaagcgaatttaaaaactatcaataggatgcgtcgttgttaaaatctttataatctatcgaatcgtccatatattttgttcgtaaaaataatttt
ttaatgtattcatagtttttttgagttggctcctcaaaatacttgtctctggccttctctacaatctctaaaatgtttgccggcaacagatccgcacttttgattagagcactacatttatctgtaatcaattgtcttgc
aaacgtatccacatccacaaaattattggaagtcataatgtttcacttaaatattgtaaaaaaatattacacgttgtttcattgacgtcacaaaatttttaaactcgctttacgagtagaattctacttgtaaaaca
caatcgtcggatgatgtcatattttgcacatgattataaacatgattagttatgatgtcatgttttgcacacggctcacaaccgaacttgctttaccagtagaattctacttgtaacgcacgatcgattaatgagt
catattttgtacacaattatatgcatgtatattttacgatgatgcaaaaatgtgaaatattttgaggtatataaggcttgatacaggcacagtaagcagcacattctaacatggtgtacatttgcattgatacggg
cagccatgccaagggatacgcggtcgaaagcagcgacactgactaccatatttacactaaatgtgatcgagagacctttgaaaaaatttatcgacaataaagagcttttgaaaaatcgacatgcaaaagat
gaatcgggcaacgatgtcaaatacgttgacttgtacacgggattaattggcattttgactggaaaatctcccgaattaagcatgttcagtaaacgtgaagattttaaagacaaatacggtattgaaaacttgc
aattgtacgagtttgtaaccaagttgatgacggtgtcgatggtgaaaattatttatactttgatgagatacaaaattttgaacaatgccaaaggattgttgcagctcatgttcaattacgtgtacgttgaatattatt
tagattataaacgtgcgccgaaatccaccaaaatttttaaacatgttgttcaacgtgggcgatgagataaaaattacgatggataaaaacaatcaactggtagtgaacgatttaaatgtattggatttaaaca
aaaataatggaggatacaaaatcgacgaaactctaacgttatttgtaaagaatgtcaaactttttaaaactgtacgtaaaactcatgcaacgcggcgagtatcaacaagaatggaccgaatattttcaacaa
tggaagcaacaattgcaggacagactgcaccacgtgccagaaccgccggagcgcaccgacattcgccacaacattgtcatgtacgcgttaaatgaacgcggacccgtaatgcccgaagatgagaat
aaaattgtgtatcaaatttatccatctgtgtcgcatctagatcaaggtaaaaaaggtacgttggccgacaaagaaattattgttcaagaaaagttggacggttgtaatttcagaatcatttgcaaccaaaataa
aattacgtacgggtcaagaaacacttacagaccggacggcaattttatgaattattaccgcattcgaaaagatttagaaacgtgcatgcgatcgcttcaggcgcggttcaatgacggttttattgtctacggc
gaattaatgggttggaaagacgacgcaaaaacgacacccattaacgtaattaactacgtggaccagaaagagagtttgaaatactacgcttacgaaattcaattatacggcggggagtttgtgccctttgt
ggaggcgcaagagttgttgactaacgtggggttaatactattccgtgtcacaaatatttgtacaacgatttttgtagaaaggttaaatttcaagtctttaatgtttccccaaagcccgctggaaggatttattattc
gatgtggaaatctaatttacaaactgaagtctgattataaagatttaaacaaattaaaaatagaaaagggtcctttgagtggttaacttgcgattacattaaaagcaactgtgacgccatagacaaatcaga
catgatgaaaattttaatttttttgttataacatgtgcaaagtaaaaaattacaacgaaaagttattgtttaataaagtgtttaatttgtttagacaacaattcaatttaaatcataatgattataagaatttgtataaa
caatatgttaatatgtgtaaatgtactgaatataaataaaaaattaaaaatattttataatttttttattatataacattgtagtttgcgttcatcaacattattagtctttgcaaatgttttgtgttgcgcgacttgtctttctt
tgaaaaggtttgaaattgtaaattactaaaagctagtttaacaattttatcccgcatgaccaacacgctgtgaattctcatgttgggctgtttgttgggcaccagctcgtcgattaatttggcaaaattacgcatg
tcgcataaaaaagttaaattttataacaaacatgaccataaagtcgatttgtttaaacgggcccaaacacgttaaaagtttcgcgcatatagttgaagtttttagcgcagttttgaaaatagttactggccaaaaa
tatgtaatttaccaattgatgagaaagcgttttgtcactgttatcattgtcattgtacagtgtcaatggaaacagcaactcgtaaaaatttacattcatttcgatatgatcgagattgtacaggtttagcctttgag
caagctgccatagttagcggcgctcggtaaagtctctaaagaatttttgcagttgcagtattcgtacagccgtttcataccgtcggtgaccacaataatttttgaagcttttcgttgacaatgtccttgctaaatt
ctaatgcttttaatttgagaaaatcttcgtcttctaattttctgcgtttactttcttttcgtagtcgacatcttcaacgtcatcatcttcctgctcgtcttctgacgatattacattaatttcgggtctattggcacgtttg
ctgctgcgagtcggcaacaagtcgtcttccgcagatctcttgccggaagttttgcgaactggactggttgcggacgtgtccaagaaattgattggcagcctgtcgttgtacaagtcgaacaaggttttaatgtt

tttgtttttatggcgtcataatcgcttttgtacatggcaatgttgtacagcggcgtgctcgaatctttaatttcttgcgtcatttttatgactattcgccgtacggcatcgctttcaggctgcaccaactcaatattatt
gcgaagcgcttccaacattacctctaacgtgtcggccacgcaacgaacgttgaaaaactttgacgctaaatttgacaaatttttcagtatcacacaatcttcctggtttaaaaatatttgaaattctccagcat
cagcatcgctattcagataatctctaaactttttcaattgatcaatagttgttatgtgcaaccgatttttagagtggccagcgccttttcgattaaacgcaaaagaaccgtcttgaccagggaattgtactcgtc
ggcagtgtacgccgcgttgtagtctttgtcctgatcggtgtaatgttcggcggtacgtcgggcggtttccatgtacgtgtccgactctgaactttcgtacatgtaccgcggcgcgctggggcctggaagcggat
cgctaaatgccgtgggaatattggcgtcggcacgattaacgacatcgtctgcggacgtctgcgctcgcctctccagaagagcgtttgcggcagtttgttcggccctcaatctctccgcttcctgttctctttcg
aacgcttccgtgtcgtccatgatttttttaatgctctctttaaatagtctatcaattttaacataacactcttccgccttatttaataaactatcgtaaccctcaatgttctgataattgccttggtttacataattttga
tattgcaacagattactttctatttcggcaactaaaccggcggtcgtctcgttgtaatacttgggataatgcttaacaaaaacgacggccatacgcaacacgttgtatcgcatttcaaaatcgttgcccaccat
tttgctgttatttttaatgatcgatgcggccaggtcgctggatgttacccgcggcggcgattgcggcgggctatttacagccaacaccatgctggtcaacgtgtctattttttcttccaaactgagcgttttgtcgt
caacaatcacatccaccgtttgcatgtttctggacagtatttgcgctgccaaacgcgtaatcaacagagaattggaatcgttcatattataaggttatatcgaatgcaccatgtgcgcttacagattacaatac
agtttaaggtttaacatttacgaccgctttgaaaacgtgtgttttgaagcgcaattgttacaagatgaaatcgattcgttgtgtttttttattttcaaaatatttcaaccaaagcttgatcgtggacagtaaaggtctt
actttttttcacagaatttaacaaatgtattgtttccataaagagtagttttgagaatcaagccaacaacacagacaacatacacaatgtaaaaaacatttttttcgatatttttgcgcgacgagtttatcaaacag
gtaccccattttaggaccattatgcaatacttgcaaaaatattacaatccgacgccggcacccgacgtggatgcgatcatgtgtcagtcgtgcaagcccgcaaacaaaattcaatgttttgagtgcaaatg
caaatacttggcgtcctcgctgagcaccctcgacgcgggcttgcaaaatggttgggacattttttttaagacccatgttcggcatgcctctgatgttgtacgtgttgcttcgaaccgattacaaaaacgaaagt
gacattattaacgaaaacaacttgattacgcagatatttgtgcaatttttctataacttgatttgcgacaaagcgtattctttgtacacgaaacgcgacatgtgcgtgccatttgtcaaagaatgcaaaaaagc
aaccgtcggattgcgccaagaagatcacgagcgcgtgttgagtattttgagcgcgcaatgcaacggctgctccactgttgctaacggagacaggctgttgctaccgtttaaaaacttcatgatagaaatgg
gccgcaatactaaaatgaaaaaggttaacaaaatagcgtccaccgtattgatcggcttttatttgagacactatttagagagtttgcccaacaaagcgtatccggtagcggagttggaactccgcaacgttt
gtcgtttcatcatgagcaagtactcggacgaaaacattaatctgttaattcataaattgaaattgatcaaaatagatatttgtaacgtgctaatgacagaaatgattgtgcccgaaacatttataaggcacatc
attaccaagtaccaattggacaacgaaatatctctgcttatcgaactcaaccatgattgcttcaataaatgataccgatatggacactgacgacaacatgtcgcaagcacgaagaaaccgccgcaacag
gccgccagcaagaccttcggcgcaaacgcaaatggcggccgtcgacatgttgcaaaccatcaacactgcggcctctcaaacggctgcgtcgttgttaatcaacgacattacgcccaataagacggaa
agtttaaaaattttgtctactcaatctgtcggagcgcgcagtctgttggaaccgatgcaagccaacgcgtccacaattaagttgaatagaatcgaaacggtcaacgtgttagactttttgggaagcgtttacg
acaacacgatccaagtgatagttacagagtaagtgtgcgtcgtcggtcacgatgagcgctatcgcgttgtatttggaaataaacaaattgcggttaaaaatagacgagcctatgcaattggcgatatggcc
ccagctatttcctttgctctgcgacgaacatcaaagcgtgcagttgaacaccgacgtgttgataaattttatgatgcacgtggctcgcaaatcgcaaaacacaattttaaataacaacgctgccatagcgtc
tcagtacgcggcgggaaacgcggacgtcgtcgccgcgccagcgtccgcgcagcctacgccgcgtcccgtaataaatttgtttgcgcgagccaacgccgccgcgcccgcgcagccttccgaagagttg
atcaacatgcgccgatacagaaatgctgctcgcaaactaatccatcactattcgttgaacagcaccagttctacagaatacaaaatcagcgacgtggtcatgaccatgatatttttgctgcgatcggaaaa
gtatcattcgttgttcaaactgctcgaaaccacgtttgacgactacacgtgcagaccgcaaatgacgcaggttcaaacggacacgctgttggacgccgtgcgatctctgttggagatgccttcgaccacaa
tcgacttgacgaccgtcgatattatgcggtcgtcgtttgcacgctgtttcaacagtccgattatgagatatgcgaaaatagtgctgttacaaaacgtggctttacagcgcgacaaacgcaccactcttgaaga
attgttgatcgagcgcggcgaaaagatacaaatgttgcagccgcaacaatacatcaacagcggtaccgaaattccgttttgcgacgatgcagagtttttgaacaggctgctcaaacacatagatccgtac
ccgctcagtcggatgtattacaatgcagccaataccatgttttacacgactatggaaaactatgccgtgtccaattgcaagttcaacattgaggattacaataacatatttaaggtgatggaaaatattagga
aacacagcaacaaaaattcaaacgaccaagacgagttaaacatatatttgggagttcagtcgtcgaatgcaaagcgtaaaaaatattaataaggtaaaaattacagctacataaattacacaatttaaac
atggtttatcgtcgccgtcgccgttcttcaaccggtaccacatatggttcgacacgcaggcgcagaagctcgggttacagacgccgcccgggtagaccccgcacatacagaagatccaggagccgttca
tcgacaggtcgccgaagttacagaacacgctactattaacaaccagacattccacacagccgacagtagcgaatgaacgaagcgatttcgtcgcctgccctcgtttggctttcgactgttacaaaatcat
gtctgcaagattttaaactaagcccgctaagctcaaatagtttatttttattactgttttgtaaataaataactttatcattcaatatttgcctgtgtttgattttctttttgcgtgtacacccagtgctggtatgtttttcc
tttttttcaacaacaaactttttgtaatttagggaatacttgtaaagaatgttgtcaaaaggttcctttttaaatttgggcggctttatgtcaaacagtttgcaatttaagttttccacgtaatttttggtgttgagtagac
tcttgtgcaatagctggcacggacaacagcaatcgacgccttcggtccgttcgataaactcgtacagctcagtgtacagtttgaaatcgtttattgtgttgttgaacagctttgccaaatattctgtctgcaatc
gaatttgtttgcgctctttcaccaaatcactgacgctgggcacgtacgcgtgcaacgagtgaaacagatggccggtagacgaaaagttgaacgttttgtttttgacgttgttgggaaaattttctgttaagaaat
ctattagtttagagtagtttttgtttatacgaaactctttaaacagtttaaataaggcgtaactggtccatttttgtacattatggcctcctcgcttcaaagcaaatggatctgtttgcgccttaacgacgccatcat
caaacgacatgttctcgttttaagcgagtacgccgacctcaaataccttggctttgaaaagtataaatttttttgaatatgtaatatttcaattttgcaacgacccgcagctgtgcaagatcatagaaaacaatta
caattactgtatgcaaatattcaaagcgcccgccgacgacatgcgcgacatcaggcacaatattaaacgcgcattcaaaacgccagtgttgggccacatgtgcgtgcttagcaacaaaccgccaatgta
ttcgtttcttaaagaatggtttctattgccacactacaaagtggtctctctcaaaagcgaatctttgacgtggggatttccgcacgtggtcgtgtttgatttggacagcactctcataaccgaagaggagcaagt
tgagattcgagatccgtttgtgtatgacagcttgcaggagttgcacgagatgggttgcgttttagtcttgtggtcgtatggcagtagagatcatgtagcacactcgatgcgagacgtcgatttggaaggctatttt
gatatcattatcagcgaaggttccaccgttcaagaagaacgctctgatttagttcaaaattcacacaacgctattgtagattacaatttaaaaaaagcgtttcattgaaaataaatttgtttttgacattcacaatc
atcgatccgacaacaatatacccaaatcgccaaaaaattgtaattaaatatttgagcgataaaaaatgtaaactttttcaagtcgatcacacttgtcgacgatctgccaactaataactacgcgtacgatttcta
cgtcaaagtcaaacgatgccccacccccgtgcaagattgggaacattatcacaatgaaataatacaaaacattatggattacgaacaatattttattaaataaaaatataaactctattcattgtcgctgtct
tctttatcgccgttttccattgcaaaatttagtgactctttcacccattgcagatgtatgtttagttcaatcaaatgtattaaaatgattactacgatgaacgtacaagcgtctatatacgcgttattcggccaaca
aatccacgtagccattgtaaatcaaataattgataatgtcgagtattgaaaactgaatctgttggtcgaccacatcgagaacgagcgtgtgatcgggcacgttattttttaatgttgcaaaatgttcaattagcg
tcaagctgctgaccggcgtgcatgcggtggcatatattcccgtctcgttgatcgttttaatagtggattgcgtgaacgtggtcacgctaaactcgttgggacaaaacacgttcgccgagtccaagtaataata
gagcgttttgttaaaattatacaacaccacattttcgggcgcaataataaacaatcgttcggtgcgggcgggcgtgcttaaatcgtaaatctgtacatacgcgccaaacaacaatgtgttgttgaaatcttgta
ttagcgtttctattctatgtaagaatggatggtgatttttgatcgaaattattaaataaattaacacgaaaagtacaaacacgatagccaacatgattgacaacattttacaatttttttgaaaaacgtgcctca
agacaaaacgtacgagattaacaacttgcaagatgcaaatcatttaatcattagaaacacgcgaacaggaacacgtagattgtttgagtacgtcaacaactttcaacagttcttaaacacaattagaaac
aactttaacggtccgtgcgcgaaacacgacatgggcgcatcttgcgaagacaccgaagaacctgcagagaaacacgcagctcaaacattagacggccacgactgggtgttggaaagtaacgatttttg
catttttgtaaaacctttcattttgaagaagcactatgacatcatacaaaagtatattaattttgaagatttttttcaagagcaccgaccctggatacattaacaaatgcgtgcaagcgggcgattactattattgg
cccaactggcccaaaaagcaagctttctcttttaacggatggcagctgtttttaaatataaaatttggaattgtcatcgaacccacgataccaatcatacacaacaaaaaactaggccctgtggatttgtttg
tgtttgaccctaaatgttttctcaacgtcgagctcagcctgcgcacgaatcacgatccgccgcaaacgctgtttgtaaatggaaaaaccaagtttgacgattctcacgaagacttgtttatactcaaaatggc
cgacggcactgtggccacgtgcaaagttaatggcgagctggtcaattcggataaaaatttttttcaattacattagggacgacattaatttggaggaatgcattacggtgcccaagtacaagcatatcgtcaa
cgtcaacctcaaaagcctcagagtgttcgagaacaacaattttgacaaaaacgacgtggacttgagcgacacgcgttccagaaaaccgcgcatcgtacccattatatcggcgagcagcgaaaatgcag
actacattcaaacgcaaatcaatttggggcttatagcaattcacgagaacatggttaaagttttggcgacccacgaacgagccaacgatccaaaccttctacaacaatactttgaaaaaagcaaatttaa
aaattttgatttttttaatctatgtactgtggaagatattgaccaagaacgaaaatttctcgtacagagaaaccgacatcaaactgtttttggaacttttgtgtgaatcattatttgcctgcgacaaagaagctttga
atgaagctttgaagcgatgtgagccttacaaaaaacaagagaaaatagtgtttaatagggcctgcaatcattggtttgactttgacgataccaaattgtgcgtatcgttgggctattattttggcatacactaca
tgatatatttgactcagtcggctaaaaacgaaattttggatcacgatgaattgtgggcgtacacgtacgagaatgtgatggcgctaaacttgccgcctgacattgtgtgtaaaggattctttagaaaattggaa
aacgtagtgaccggagtcaatttggttttcaacggcaaacattatcaaattgtaaagaaagaggatgacctattcaaattgaccaaaagcaattgttacaagttgagcaacataaaatttaacaattggaaa
tacttgtacttgacaacgcacggtgtgtacaacgtgttcaccaacagctttcattcgagctgtccatttttgttgggcaccacgttgccgcagacattcaagaagcccaccgacgaaaagtatttgcccgag
gacgcgtttaattacatgctatctactagcgccgacgagctcagcatttatagaacttatcacatcgccaaaatgtgccgagatgtaaaaatgctaaaaactaacacggccatagttaactacatgggcaa
ttgcaacacatgccaagccgatatgcgagtcgcgttaaacaacctgtttcgcgatttgtggaatttggacgatgaaaatctgattacgctcgctctgtacgtaaacaagaacagggtttccgacatgttgcac
aacctaaaatgcaaaccgtgtcggtcaacggtatcgggatctcgaccaaaatgcaaatgctacaagaaaataaagatcaatagaaaagcgctaaaagtctgcttgatggcggacatgttcggtaacgat
gcagaattgagcgaattgatttggatgctaatttttactaacaaaacgtacgtgtcgaccacgttgattcgcacgaacagcgaatttgttaaccaacacggagagtttttttccaaagaacacaacaaaatta
ttcagtacttgtatcgaacaattcacaaaatcgaatatgtcgacatgttgatggacaagtttaatgataaacgtttgtttttaaccgaattgcgcgatgatgttgcacgtgaacccgatgttcagtttgaagaatc
ggataacatttccaaattttacacgcaccacgccgacgcgttaatgatcttgaagaagtacaatgtgtggtgggacaaaatcatattggcgcgcagcacggacgatttgcccacttggctgactcgattcta

catgcgcatcatcatgtccaaagtggacttgaaggagtactcgtacaattatttgaaaaagattgtcgaaggctatttgtactttaaacgcttcaccaattttaatcacgcgaacgcgattatgttgatgcattt
cgccgctagtttggccattcccgtagattacggaaagaaggctatttacatgcccggcgaaccgggttcgggcaagtcgtcgtttttgagctgttggactatttggtgttgatgcacaagtttgacgatgaca
accacagcggagagtccaacaaggaaacgagcgacaaggaagtgagcaaactcaactcgcagttgtacactatcaacgagttgaaacaatgcagcgagtcgtactttaagaagcatgccgattcgag
caagagcgattccaagagcagaaaataccaagggctgctaaagtatgaagccaactacaaaatgttaattgttaacaacaagccgctgtacgtggacgattatgacgacggtgtacaagacagatttttg
attgtttacacgaaccataagtttgtagacagcgtcaagtttgccggctccgtgtacgaacacatcaagtcgaaacagtttccaatagaaagcatgtattacgaatcgcttgtgacgccagtgcgtttgttttt
gtcgcacgtattaatgtacagacgcgatcccaaaacgggctttgtagtatacaaaacgctgctgagcaacgaccctatgcataagcacaatctaatgtgtttaagcaccaacaatagtccgctgtatgcgc
tcatttacattctaaacattaaaacggtgcgcagtgctactattacaataggagaggacaaaatggaggaaatgataggtattgccgttcagcatttaaaaaactttttgcacccttcgtttgttcagtacaatt
ataaaaaaaatataaacgcaagtagttcaaagtcttttgtatttaatgaacaagttttattacagcaaattaaaaataaatttaaaaataattacaataaaactacaaatgtgttttataacatgacaatggcgt
tgaacagaaacgatttgaacactagtgtaccaaattttgtatgttaataaataatatattatataaatagactttttgttttattaattcatttctctgtatgatttatagtttttgagataggcggtttgttcagcattaa
tgacggtaaagttattttccaataaaacttgctgagctgccgcgtttttcgtgtacaccagagtgcagttagatttgcgcaattcgcgcacctctcgttcggacagcggcgccgtgttggggtcgatcaaacac
acaatcatgttgtcgacttcgaatcgcaacatgccattgtcctctacagttttagtgggcgtcacacccttgaatacgataaacatggacttgaattccttgatatcaaagtttgcggcgccgcgcgtcatggtt
acattactagatgtgctggccgcgtttggtgatttgattccatagtcgtatacggtgcccagaaaaacgctgttagtgcccaccttgttggagccctctttgactatgtcgtctatagatagcggtttttcggcaat
atacacttggctcaatttattatcgccgtgtgcgatgcgcaagttggccacccggccgttattcagctttacgtttaattgtttgttctcgtctaattccaccgagtcgctgctctgccctaaactgggcgacgac
tcggttaaggaattgaaatttaatgcattcgtccaataaagataaaacagtatgagcaaaacgataagtaacacgattccccacatgatttgttttaatttacaatttcaattccaatgagatttaggttgtgca
aatgtttcaaaagtacacttaaattatcgttttcccaattgacgggcacggccacgacgattttactgctgagcgtaaaaatttcaaaaatggccgtgataaactgttcgcaatagttgataacgttgcttccg
aacgcagactcgtcgaccaaatgcacatatgtcttatacataatcattgcatcgaacaatacgtggtgtttgttaaaagcaccctcgtttttaagcaccaagtgtttgtatactttttcataaataaggcttgct
cctgtttgagttcagcaaaataagtaaacttgtcgcattggtcgtccgtaagacattctgtaatattttcgaaataaataataggctcgtcggcgtccgtgtctgcgctaacgtacaacgacgctgtggccggc
gcctgacacaagtacagcacgatcgttttgctagtcgccatgataccgctgacgccgctttttctcgttacaaagacagctatctgttgtattcgtttcgtctcatagaccttttgcgagcctctaaatcgacg
cacctcaccaaactgttgtcttcgcaagcaacttacttataccattttgcgtgtttgatgaaatacaaagacattcaaaagtacgaggtgcagcagttgatcgaatgggccattaacgcatcccccgacatg
gacttgcaacagtttcgcatagaatttatggacaaaaccacagaactcaatttgcgttcgtgccagcccaagagttttacttacacgtttactaccatatgggacaccatgcactttctcagtctcatcatcga
cgatatggtgtatacgcgagacaaaagcagtttagattttgtcatgcagcagttgaaaacaatgaaggtgctattttacaacgtatttttcatactacagtgcgccatgtgccgcgatcattatatgaacgtca
aaggtttcatcatttaccacattgagcttatagaaattgcacttgacaaagaaaaaatacggcacagacattacgtttgtcgattcgtaccagcaagaaacggcaggcgcagatgttgcagttgtgtcaaata
atatgctcatgaaaaatttaatggcctacgtcagtatgacttttcacaaccacattaacgattacaagtggatacaaagaaataaaaaaaccgcccgctcattacgaacgcatgacatggggagaatacaa
aaaattgttaaatttgcaataatacaacattctttgtttaatttttatatttaagcaacatgtggtacattattgtgataataataataataatattaaatttatttataataattttattaactttaaaaatattagacga
cattcaggaattgtatccaaatccaccaattgtaccaccacctacatctccaccaattgtaccgttacctacacctccacctactcctacaccttctccaccttcgcctacacctcctccaacacctatacc
tccaacacctacacctacacctcctccaacacctatacctccaacacctacacctacacctccaccttctcctattcctcctacacctacaccttcgccgccgccttctcctattcctccaacacctacac
cttcgccgccgccttctcctattcctcctacacctacaccttcgccgccgccttctcccagtcctttgggcgaacctatgtattatccttctaatatagatacaaatgaagcattaataaatttttcttaggccgtt
gtgtgcaactgatagaactcgagcaacatacgctgtgccgtggaattgtaacgggatatttcattgcaattacttctcttatttcttttatatcctcagctgtcacaacaatgcgtattcgtttgttaacgatggttg
tattgactttaacagctcggactgcccgttgtatccgcttaacgtgctttaatttggcacgattcggtcgcgacgatgtttcaacacgtttattgtcgtgtccgtgattacacattcgtaaatgttttcgtgcagtaa
ctccggcaagtcggttaaaatcatgagaccgttcaatggcccattgagtgtgttaaacgacttattcacggcgtcatactctaactcggttgttggcatccatttatatttgacgtatctcaactcggtgtcgag
cacgacatacccgtccacggacacggtcatgtaattgctctgctgtagcggtggatcaaaaaattgctgaaaccgcaattcaattgcggggcaagtgtcggtcaacgtgacgctttgcacgttgttgttcat
gtagtttatacattcgatggccgtcaccggatctatagcgtatgacgcgttcacgccgcattcgtactgtgttcgattgttgtatttgtatttaaacacttgcagcaaatctgtaatgtaaatcttttgttgtccata
acctcgcattgaaaggccacaatgttgtttagcgcaaacagattggccattttggttttgtagaattgcatatcgtcggtttgaataatgcaaaaattgcgcataaacagacctctgccccgcataccgtcca
atttgagcgcccatttttttacggacgcgttgtcgaacgcgtcgtcgccgtacactatttttttgttcccgttcaaattttcgaggggccatcttgtcaatcaacgtggtcaacggcacgaacggtgatatgtttgc
gcgtcggttaacgcttccatgtccgcaattatgttgcacatcacgtttagctgcgcgtcgtcgggcgcgtcgtcgtcaaattcatattcgagacgtatgcgggccaaaatttcatcgctgcccacttgcgagtt
ttgcagcgtctctgaagcgtccccatttttccaacaaagtcagcagtttggatatcttgttcgccatcgtggaatcaaatagatcaatgtcacttttcgaaaaatatacatgttcaaatttgatttcaattttatcgt
gttggtaaacgtacactttaattatttttactcaagttgtgcgaaagagtcttgtaaggcagtttgatttctttgctttctctccacaccaacggcaccaacgcgttggtatctttaggccaataaacaaatttttttgt
gtttggaattagtctttttcacgcttgatattatgttattgcaagcgctctgaataggtatacgagtgcgaaagccgttttcgtcgtacaaatcgaaatattgttgtgccagcgaataattaggaacaatataaga
atttaaaattttatacaacaaatcttggctaaaatttattgaataagagatttctttctcaatcacaaaatcgccgtagtccatatttataacggcaacaatatggcgctagtgcccgtgggtatggcgccgcga
caaatgagagttaatcgctgcattttcgcgtccatcgtgtcgttcgacgcgtgcataacatacaaatcgccgtgttcgcccgacgcgtatcatgacgatggatggtttatttgcaacaaccacctcataaaac
gttttaaaatgtcaaaaatggtttgcccatttcgacgaagacgacaatcaattcaaaatgacgatcgctaggcatttagttggaaataaagaaagaggtatcaagcgaattttaattccaagcgcaacca
attaccaagacgtgtttaatctaaacagtatgatgcaagccgaacagctaatctttcatttgatatataacaacgaaaacgcagttaacactatatgcgacaatctaaaatataccgaaggtttcacaagca
acacgcaacgcgttatacacagcgtttacgcaactacaaaaagcattctggacaccacaaacccgaacacgttttgttcgcgggtgtcgcgagacgaattgcgtttctttgacgtgaccaacgcccgagc
gcttcgaggcggtgctggcgatcaattatttaacaattacagtggatttttgcaaaatttgattcgacgcgcagtagcgcccgagtacttgcaaatcgacacggaggaattgaggtttagaaattgcgccacg
tgtataattgacgaaacgggtctggtcgcgtctgtgcccgacggccccgagttgtacaacccgataagaagcagtgacattatgagaagtcaacccaatcgtttgcaaattagaaacgttttgaaatttgaa
ggcgacacacgtgagctggacagaacgcttagcggatacgaagaatacccgacgtacgttccgctgtttttgggataccaaataatcaattcagaaaacaactttttgcgcaacgactttataccaagag
caaatcctaacgctactctgggcggcggcgcagtggcaggtcctgcgcctggtgttgcaggcgaagcaggtggaggaatagccgtctaaaaatggagtttgttaaattgcaatgcaacatttgtttttcggtt
gcagaaattaaaaattattttctgcaaccaatagacagattgactataatacccgtattagaactagacacgtgcaaacatcaattatgctcaatgtgtatacgcaagattagaaaacgaaaaaaagtacc
atgtcccttgtgtagggttgaaagtttgcatttcaacgtttacagcgtaaaccgaaacgttgtggatgtgattaaatgcagcgcgtcgagtgtcgcacaatggaataaaataaacgctaattttgatgcggcct
ctttggccagtgtgttgtttgaaaaaagtctgttggacgacgcagaagacaacaacgcaaatgccgacgacactatgctgtctgaagcacaagcgatactaaaaaaactacaggtagacattgccgaac
aaacgcagctcaacattaaacaacaattagatttagacaaattacaacaaacaagcgtttctatgcaagaaaagttagacaaaattaaaagcgactacaacaacatgcataaatcttttaaagaattgc
aactgaaacgaattacaactgaaaaggcgctaaaatcgctcaatgacgattacgcaaaacttgcgtctaaaaacgccaaattgagtagcgaaaataaggttttatcaaataaaaatattgaattgattaa
acacaaaaatttattacaaaacgagtacacaacattacaatcatataaatgtataaccaacgccactattaccacaaatgttacaataaatgtagattaattaactgtttacataacattctactttaatgtaa
taatattcttcaatttcttgggttttagaattaaaattaactctatatatatgtctcttgtttattttcacaccttcttcaatatcgtgatgccaattatttgtgtcgcgccatcggctatcgtttaaagcgcactgattga
gcgcctcgttgtattgttgcaacgtaagattgtatgtttttagttcattcttgatcgcctttacgcccttgaaacgcgtgtacgcgtcacgtttacatttttctagtattttatgtgcaataattttcaatttactttgaggt
tgcaagcccaaagcattcataaatatcattttattattattaacacaaccccttgtattcataatgttgtatgcgatggtactagcaaaaactgttacattgcgtttatatatgtcggcatcgaacgcgttgataa
caaagttatcactgcactttatataatagacagtttagcattggtatgcattattaaacagcaatcatgttgcacgtttcgcgtctacttgcaaacggcggagtgaaaaatctttgtgacaagtttaaagtcaaa
attaaaaattataccgaacacgatttgatggttcttaattacgaatcgtttgaacgcgatcgagatcatcctgtggtcgtggaatgtcgcggtttaattcttaacagccgcacgtacgccgttgtgtcgcgttctt
tcgatcgtttctttaattttcaagaattgctacaaaatatcggcggcgaagatgctcatcacaaactgttccaatcgaaggagaattttaagttttacgaaaaaatcgacggaagtttaattaaaatctacaag
tacaacggcgaatggcatgccagcacccgaggaagtgcttttgcagaaaacctttgcgtgtcagacgtgacgtttaaacgtttagttttgcaagctcttcaacttgacgaggcccacaatcaatttcaagcg
ttgtgtaacgaatatttagactgcgcgtcgactcacatgtttgagttaacaagcaaacacaacagaattgtaaccgtttacgatgaacagccaacgttgtggtatttggcaagcagaaacaacgaaacggg
cgattatttttactgttccaatttgccgttttgcaaataccccaaatgttacgagtttacgtctgtgcaagaatgtgtcgaacacgcggcccagttgaaaaatttggaagaaggatttgtagtttacgataaaaa
caacgcgccactatgcaagattaaaagcgacgtgtatttgaacatgcacaaaaatcagagccgtgctgaaaatcctactaaattggcgcagcttgttataaatggcgaacacgatgactttctagctctttt
cccgcatttaaaaagtgtcattaaaccgtacgttgatgctagaaacacgttcaccaacgagtctacaataaatattatggtgtcgggtttgactttgaatcaacagcgtttcaacgaactggtgcaaactttgc
cgtggaaatgtttagcgtacagatgtcgtaaagcacaaacaatcgatgtggagagcgaatttttaaaattaaccgagcctgaaaaaattaaaatgatcaaaaatattataaaatttgtatcgactaagcagg
cattaaataataaacttgctccaacaattaaattgccttcttcaaaacaattgttggtgttgataggcatttcgggttcgggcaaaagcacgtatgccaaaagtttaaaaggctatacagaaatcaacagaga
cgacgtgcgcgttaagttgtttttaaacggcgattatacaaaactgaacgctttctataaccaatcgcggaaatgtagacaaactaaagaagaacaaattaccaaaatgtgcattgaacagtttttaaaagc

cgcaaaatgtggtgccaacgtggttgtcagcgacactaatttgaacactcaatcggtggacatgtggcaaaaaatggccgcaacgcataattaccatttttttaacgcgtttgatggacgtttctttggaaacg
gctcttgaacgcaattataaacgctctgataaatttccattaaatccagaaactattaaaaagcaatataaaaagtttctaaaggtaaataatttcgaatattacgttcctgtcggcgacaaatttccccgcgc
tgtattatgtgatttagacggtacagtagctttgcccacaaatagatcgttttatgattttgataatagagttgctcaagatgaagcccgcttagacgttattacttgcgttaagtatttggcaaattgccacgacg
caattatcgtgtttatgagcgggagaagcgtaatttgtgaacagcccactcgaaactggattgaaaagtattttgatattaaatcatataaattgtttatgaggccttcagacgatacttgtaaagattatttatta
aaattaaaattgtttaacaactatatcagaggaaaatataatgtaatagcggtgttcgacgacaggccgtgcgttgttcgcatgtggcaagatttaaaaatacctacagtgtttaatgtttgcagggattacttg
gaattttaaataaaaaattttataaaaaatatattcacttttatttaacaccatttccaatatacgttgggatcttggctattgtcatcgttccatccaatgcgagttcccagcaacacacacctatttgcttgatgg
taatctggtagcgcttgcaaattcataaaaaatgaataaaataaaatcaaaattgttaacagcttcatgtttttttattgttgcacacgccgctgacccctcgaccgtcgacaaacaaataataaaactttgca
caattggtgtgattttatattgtaaaataacaggaaaataatgtcataatattgttttattattaatcattgcaccatgataacctttaatgtcaaacaaacaacattatccatgtaataaaacaatgtgtaaattgt
tattcttttattttaaaaacatttaaaccacaaatcatacaaaaataatctatatacaaattactctgcaattgtttctcttcaacgacatgaccccgatcacaattggcaaacctttccaattttgaattcatga
aatatcccacgctgcgctgcggcacttttctttgtagcaaagccggattaaaaataattggtaaacccagcgaaaatataatgtctcgcatcaaacaaacagcatttaacgattttaactgacgaaaaggtt
tcttgagcgaaacggctattttgtccagacacaaatcttgcaagcttttcagcgatttgtttttgcgttttattactaactggttgttaggtccatattcgtttacaatcgcaacgttaggcatcaaagataaataat
ctgacaaaaagcgaatatgtgatacatctttataacgaacttgaataatgtcacccactaaacaaccgggcacagttgtgttattagggtatattcgaaacatggtgacaagtaaatgaaaaaaaatttaact
ttgttcggttatttatacttgctatcataaacaaaatatttgtaatcaaaattcaactcgctttacgagtagaattctacttgtaacgcatgatcaagggatgatgtcatttgtttttcaaaactaaactcgctttacg
agtagaattctacttgtaaaacataatcaagagatgatgtcatttgtttttcaaaactaaactcgctttacgagtagaattctacttgtaaaacataatcaagagatgatgtcatttgtttttttaaaactaaacttg
ctttacgagtagaattctacttgtaaaacataatcaagaaatgatgtcatctgtaaaatgatgtcatttgtttttttaaactgaacatactttacgagtagaattctacgtgtaaaacacaatcaagaaatgatgtc
atttgtttttcaaaaccgaactcgctttacgagtagaattctacttgtaacgcacgattagtggatgatgtcatcctttatacacgattattaacatgtttaatgatgactcatttgtttttcaaaactaaactcgcttt
acgagtagaattctacttgtaacgcacgatcaattattagtcataagctgatgtcattttaaaacatgattattaacatgatcatgataggtagatgatgtcatttgtctgatgtcatttttgcacacggcttatttta
acttgctttacaagtagaattctacgcgtaaatcatgatccggaatgatgtcatttgttttcgacataaaatgtttatacaatggaatcttcttgtaaattatccaaatagtataatttatccgattctatgttacattt
aaattcgttgttatcgtacaattcttcgggacacgccatgtattggccgtttttgacgtgcaaccaacggttgtatttgacgccgtcgttggattgcgtgttcaggttggcgtacacgtgactaggcacggcttctt
tttttaccactatcgcatctccgtcgtacgcggatctacaaccaatcccattgcccacataagcgtacgcgttcaaaacgtgcgataggtctttggccaattcgcaatcagcgtccactttaacgttgttgcgt
aactcgtttaaagcatcaataatgacgtcattttccgcatgacaactggttagcttgaaaaacggaaccgagtagtggcatgaataaaataaatctttgttgtctaatattagcggtgacggtaactctggctc
gggttctggctcgggttctggctctggttctggctcgggcggtggaggtggtgcggttggcggcgccactacccaatggtcgtcgtcgccacctggttcgctgtaccatacggcgtcggtcgtttgaaaaggg
ttaaacaggggtagaatagaatcgttttgcctgttatttggctggttttcaaagggattaaacacgggcggaatagttgcagcgtttttttggaatcgtagcgtttttttaaacgtgtatttgccccatatatttcgatat
ttttgggccgttttacgagagtgtggttatattgaagtggagttgtagtatatttagagtttagtgtagtatatttttggtaaagtgtagtatataggtcggcacaaacgtccgactgtagtatatttgtcaatatacgt
cgctgaaaatgttcgttttctacgatttttggccgttttcgtaggatatagcacaaaattggcgataatcctcatcttttccggtcagcgtgctgctcaatacgtttaaaaactcgccatcgccaaaattgagcg
tttttaaaacgctcgttcccgaatcgttacaaatgtttgcgcgtctagcgtcaaacacatcgtacaagttgtctccaaaacagtcgataggctcgccggtgaacgggttaagcccgatggtgtctttgtctctc
gcatacaaaagccattgagcaaagatggcgtcgttgtttaacgacaaatctttggaaaccaactgtttaaccatttcggtcgtgcccaacatggaggtttgcatgtcgatgccataatggtttggaatattttc
aatggcaaacgtggaacgtaaaaagttgacattgtccgcggtggctggcacgcacgcggtgccgtcaaacacctcagttgggaattgcatgtttaatctaaatttgtccataaacaacgcgttttcaaaca
cgtttgattgatcacattcgatgtcggaaataacttcaaaattatcgcacaccaattggccactgttgtacgaaatatcgtcgtcaacgtgtttgaatacatgttggcccgtaccgttgggtaaatctatgcatct
ggagtcgccggaacactcgtactggttgtcagagtttctgatccggttgatgcacgttatcagttgtgactcgttattattcaaacatttgaaatattgcgtgtcgccgatatcggccgttatgtacgtgtgtccgg
cgccgttaaacgcgcacggatgcgcttccacgcacgacattaagttgcgatcaaatattttattcgcggggcattcgcccaccacgtggcgcccatttacgcactgcataaactggttgacgagcaaattg
gagggaaagtatgatagtatatagccgtctggcctgttttcacacaattcgttaactttacactggccggtttccgcgtcaaacgtgtaattatctggacattcttcgactgcgtgcgctccgtttgcaaaacac
ctaagatagaacgtgggatgatacaagtgcgcgttggtagaataatctttgtccaagtgttggttcaacaccaacgtgtccagcaaacgctcgtccatgggataaagaccggcagacttgttgtcgcacgg
cggcacgggaacacattttagttgtgcgtaatcaaagttaaaatatgcggggcatttcatggtcacgtcggccttgtcgccgctcaaaataaactcgttgggattttcatcatttgctctaacgcgatcgtgtac
gattcgatcaacaggttgaaatttttgatttaagaaatcaaaaatttcaatccggtcatcatgcacgctttcgtgataggtggaaaggtcgacggtgttgaaccacgttacaatataagtgttttgcataatatc
cgacacgtagcctattacgtcgggtgtgggttcgtctgcgttggtgcgcttcacatattcagtcatcacttggagccgcttggtgaaagtcgtttcgtcaaattcaaaataaattgccaaatacattaaagtaaa
cgctattataagaaaaatggcaagcatgagcaacattacgccagacatcatcgtgaacgcccaaatcaacagtgaagatgaaaacgtactagattttataattgaagacgaatattacttaaaaaaacg
gggcgttggcgcccatattatcaaggtggctagctcgccgcaactcagattgttgtacaaaaacgcctattctacagtgtcgtgcggcaactatggtgttttgtgcaatttggtgcaaaacggcgaatatgatc
ttaatgcaataatgtttaactgcgccgaaataaagttgaacaagggccaaatgttatttcaaacaaaaatttggcggtccgataatagtaagacggacgctgctgtacacacgtcgtcgcccaaacgcact
gtagaaaccgagaatgacgacgacggagaggccgcctccgccgccgctatcgatgagcaagaaggaaacgccgacgttgttggactcgatttcgaagaaaatatcgacgacggagacgctccaacg
cctaagaaacaaaaacttgacaacgctaaacaagattaaatacgatagtgaactgttgcttcattatttgtacgatgatcaccacaataaaaactctgattacgcaaataacaacattaatgttattaaaat
atcaaaagtaaaggttaaaaaaacgggcgcctcgatattagcccattattttgcacagatacacgtgtccactgggtactcgtttgaatttcatccaggcagtcaaccgcggacatttcaaacgatacaca
ccgacggcctcattataaaggtgttaattttgtgcgacgaatgttgtaaaaaagaactgcgtgattatatcaaaggtgaaaattcgtttaatgtggcttttagaaattgcgaaagcattttgtgtcggcgcgtcag
ttttcaaacggtgctgttgacctgcgccattttgttgttgttattcaacgttgaaaaattttccatgattaacttgttgataattttgttaattttattatcgctttttttgtcacaacaactatattataagtaatccgtatat
tgagttttgtaatcataagagtacaaataaaaagtatgacagatgaacgtggcaatttttattacaacacccctccgccgctgaggtatccctctaatccggcaacggccatattcaccagcgcgcaaactt
acaacgcgcccgggtacgtgccgccggcgacggtgcctacgaccgtcgccacgcgcgacaacaggatggattacacgagccgcagcaacagcacaaactcggtagcgattgcaccgtacaacaa
gagcaaagaaccgacgctcgacgccggtgaatctatttggtacaacaaatgtgtggattttgttcaaaagattattagatattacaggtgcaatgacatgtcagaacttagtcctctcatgattctttttataaa
cactattcgtgacatgtgcatcgacacgaacccaatcagcgtaaacgtggtaaagcggtttgaatcggaggaaaccatgatacgccacctgattcggttgcaaaaagagttgggacagagcaacgcgg
ccgagtccctgtcgagcgattcaaatatatttcagccatcgtttgtgctaaattcgctgccggcgtacgcgcaaaagttttacaacggtggcgctgatatgttgggcaaagacgctttggccgaggccgcca
aacaactcagtctggccgtccagtacatggtggcggaagcggtgacgtgcaacattcccattccgttacctttcaatcagcagctggccaacaattacatgactctgttgctcaagcacgccactctgccg
ccaaacatacagagcgcagtcgagtcgcgtcgctttccgcacatcaacatgatcaacgatctgatcaacgccgtgattgacgatctgtttgccggcggcggcgattactaccactacgtgctgaacgaaa
aaaacagggcgcgagtcatgagtttaaaagagaacgttgcgtttttggcgccgctatctgcgtctgccaacatattcaattacatggccgagctggcgacgcgagccggcaagcagcccagcatgttcca
gaacgccaccttcttaacgtcggccgccaacgcggtcaattcgccggccgctcatttgaccaaaagcgcttgccaggagagcttgaccgaattggcgttccaaaacgaaactctaagacgttttatttttc
aacaaataaattacaacaaggacgccaacgcgattattgccgcggctgcgcctaacgccactcggccgaacacgaaagggcgcactgcataatcatggcgacgactctgtacaccaacaaggtgtg
gtgcgtgtacattctgcggcaagacaatggaaaattgtacacgggcatcacgagcaatcttaacagacgcataaaacagcattcgaacaaacaaggcgccaagtgtttgcgcaacgcaaccaatttac
ggctcgtgtatcacagcgcgagcgcgtacgactacaagaccgccgcccgcatggaatacaatcttaagcgtaaatgcagcaagtatttcaaattgcgtctcatcaaagccaaaccacttttgttgcataa
attcctgttagcaaataagttgtaaacatgaatttggacgtgccctactatcggttgggcaaccacgaaaaggtggaatacatcccgttgaaattggcgctcaacgacgatactgccagcgaaccgcaac
aatttagcgagcccgttcacaaaatgccaataaacgacatggtgggctatgacaacacgacgagcaacgtgtcggcgggaattataattttaattagtgtagtcgcttttatagctttattcttactgctgtatg
taatatattatttcgtaatattaagagaacaacaacaatattcggatagtattgacaccgattctccttttgtttttaataaatttgattaaccacaatgaacggttttaatgttcgcaacgaaaacaattttaattc
ttggaaaataaaaattcaatccgctccccggttcgagtccgtgttcgatttggccaccgatcggcaacgatgcacgcccgacgaggtgaaaaacaacagtctgtggagcaagtacatgttccccaaacc
gtttgcgcccaccactttaaaaagttacaagtctcgattcattaaaattgtgtactgctcggtagacgatgttcacctggaagacatgtcgtactcgttggacaaggagtttgactcgatagaaaaccaaaca
cttctcattgatccccaagaactgtgcaggcgcatgctcgaacttcgctcggtcaccaaagaaacactacagttgactataaacttttacaccaacatgatgaacttgcccgaatacaaaattccccgcat
ggttatgctgccgcgcgacaaggagctcaaaaatatcagggaaaaggaaaagaatttaatgcttaaaaacgtaatagataccatattaaattttattaatgataaaattaaaatgctcaacagcgattatgt
tcacgaccgcggtctaattaggggcgcgatagtgttttgcatcatgttagggacgggtatgcgaatcaacgaagcgcgccaactcagcgtggacgatctcaacgtgctaattaaaagaggaaaactgcac
agcgacacgattaatttaaagcgaaaacgcagtcgtaataacacactcaacaacatcaaaatgaaaccgttggaattggcacgcgagatttattcacgaaacccgaccattttgcaaatatctaaaaac
acctcgacgcccttcaaagatttcaggcgactccttgaagagtcgggcgtcgagatggaacggccgcgcagcaacatgataagacattatttgagcagtaacctatacaatagcggcgtgcctttacaaa
aagtggccaaattaatgaaccacgaatcctccgcaagcaccaaacattacttgaacaaatacaatataggtttagacgaaacgagcagcgaagaggagaacaacaacgacgacgacgacgcgcag

cataatcgcaattcgtccggttcgtcgggagaatcgttgttgtactatcgcaacgaatagagtaagggaataaaaatgaatttatatttgttgttgggcgcactggccatatttagcctagtgtatgacaaaaa
ggaaaacagcatctttttgtatctgctcatattgtttctcgtgtttattatcgtcagcccggccattataagtaagaacaccgagtctaccgtagaagacataccgagtcataaagctaagagcgtccgaaaa
aaattggagatcgagcaagcgcttgatgcgattctaaacaaaaataccagctcaatagattgataagtgtttgttaaaatgtccaatttaatgaaaaacttttcaccgaactggtcaaatcgacgacgttca
ccaccaaggtgtccgtggtcaaaaccactcttagcaattggctgtgcgaacaagtctatccggacaaggatttttcgcttaaactcaaacgagtagttaacatgtttttaaacaatgaaattgaaaataaca
aaatctacaaactagtagaaacggtcgactcttctaataaactcagtaggcgacaggttgatttttaatacatgcattgttaaacaatgtcagtgtaacatttacattgcatagattcgtggacgacaacgttt
taacacaggacgagctcagtttttagcaaattttctagtcacaaaaatggatgaagcataccaactgccagcgtattaatagtaagatcaataaaaatgaaaataaatttgtgcaatatacaatttcagtctt
taattaatttattaaatttacaagcgaccaccatgtcgttgctcaattctaataaaattaaagccgatttgatacccgatgaagacgatcccaaaaagattacgttacgtctgaacagcgtgcccgaagagc
acacggacgacaactcctcgatagatttaccattgaccacgccaaagcaaaaagatgattttatgaatgccattaaatcttttgaaactttaaatattgaaccggatatagttaaaaccgaacaggccgac
gcgccggcaacaagcggcgataataatcgaaaagtggtcgatgcgaacgtggacgaatacacggtggacggtttaaaacttaaatcaaaatacgtcgcctattacaaatgtttgaaaatacttgtagattt
tttagtaatgtacgttagcaaagaggtcagcatgaaagaatacgagcaagtgtacacgttaggcagacagttgtacgaggtgttgcgcagcatttttgtcgacgagccgttcaaattgtggttggaacgcaa
cacgcacgaatttgacaacaataaagacaaaattttagaaactttacagagcgaattgaaacttgcgttggccgataaggacaaattgaaaacgtgcacgttcaaagatatcatcaccaatttgttgaaca
caaaattggattgcaaatacgattgcgccgacgagtatattaagccgaattgtatagtggacacgtacaattgttgcaatttagttttcaaaaaagaaacatgaacacgtccgtggacgtggtgacaaagct
cattcatttgcagaacaacgtcttggacataatgcgcgaagtggaccaatatctaaacagcgacacgcccgactacaccatagaatccctaaacgcgccgggcaaacagtttgatttttagatgaaatg
ctaacgaaaaaactaatcgaaagcaacgccatagtgtttgacgaaaaaagtaaaaatttaaaaattattcacaacaacatcaacatgtgtttaaattggtgtattaatttaattacaataaaacattatgtac
aataatatgtcttttattttacaatatttatgtatatgaaattacaattgttattatactgataatagcgatcaccgtcgctgtcgtcttcgtcgcttagatgcgacacgacgcagcgacaaaataaactcgcaaa
ttttttgaaaagcctttacaagttttacttgcagtaggttaataatttttttgtcatacaatgcaatatttatatatagcttatttttaaagaatgtcgtagtcgtttgcaatgtttactgaggtaatgtttcgatgcgct
gcataatttttgcgttgcacacacaacaacgcttgcaccgacgcgaacattccgtgcacacagccaggtgacggcacggcatgaaacacaccgatttttctctatcaaaacaaactttgcattcggagac
ctgttgagttagcggttcggcgggcggcgccgaaggtttaacgcttagccgcggaggaatcaaatcggcagccagtatttcttttccgcaacatcgagtttgccaaattgttcattgacggaatagtcttcta
ttgcgttgagaaacttgcaattttgtttgtgccggcgttgggcgtcgtctacgcttttatatttaaaaactatatggcatccggaacaacgcaaatggccggctttgccaaaataatagaagccgcgacgagc
caacatgtcaacgaccacggattgtgatgcaaactgacgccgagagcttttaaaactcgtaaatgattttttctcatcgattcattgaacatcaacgcgttggttgacgatatgcaataattcgaatacgtgt
gtcgtttaatttgcttggcgttaatcttgtccaaaatcgtgcgacagcctatacatttaatcaggttattgtgagaaaacagccccgttttggccaattcaagtttgtactcattgtctagcgacatgtttgccatcg
ttctgaatctaccgtctgtggaaagtatcaaaaaaattaaattgcatcaaattcattttatagtttattccataacaatatccttaaaattaaacatttattaacttgtatatacttttgaaagtataatctacactatg
cttccattttggaacacgttttattttcattttattaaaacaagccgctttactttaacacacaatacattatttttaaaatatttagatcgtcggcatatacattcacacttataaccctgtcgaggacaaaaaca
ttgcatggtattggtacaaaaacagcgtgcaaatatatgacattgagcacacaatatcatttttctagaataattgtcgacatcgtaaatgtctaatatttttaaatctacatcggccaacgttgcgtacagcga
ctgcatcttcaaacgctcttcttgcttctcttttgtacattttcaataaaggttgattgcaaacttggcacgttttgattcttttatacatgaaaaacaaacggctttaaacgaaggacaatgaccatcaatcgtat
acaaatttacattgtataaattcttttatattccaaacacatgtaacagtttttattaaaccaacaacaattgagctctttacaactttcatttgtaaaataggaaattaaacatctccggtttgattgttgattaca
ttcgtttttgcacgttttattaaacttgttgctggtaatatatatttctttgttttcttcttcggtttaatttttcgatatatatgatgatggacaatctatattttttggattgttgcatatgttgcatatgcattgtccacaaaa
tttttcatcagaaaattttggtaaaaatctgtagcgacatgcgtcgcaatattttgttcgtaaaaacttaggcggctcaaacttgtaaggcagggttttcaagcacaaattggctagcgaatccatgttgaatga
tgcctgatttgaagccgctggctttttatacacacgttatattttcaacaagtttatcaacttgtttaagtttttttaattgtatgcgataatatttttcaaactgtttttccaactcgttgacataattgttacaatacggt
ctaactaatttattgaaacaaattaaatagcgttcggaattggcaggccgagaagaaggcggtttgtataaaacccatttttcgaaatggttaacaaacttgtttagcatttggatcgtttcgtgttcaaacgcg
tcgaaaacttttaaaacgcaattgccgccgggacgcaggcaaattaaaattagctgcgtctcgcacatgatcaaatcaaagttgagacgttcttgttcgttttcgcgtccattaacgtcaaccgagccatctg
ccaacaccagatcgcacgcgttgccacacttgatgctaatctcaaatacaacatttttatcaaacacgtcgcctgacttgtcgggccccgtaatggttgtgaaatttttgcgtttgcgcactgtcggtttgtaca
cgcacaccgagttgtttgtcaacgtgacgccatacgctttgcaaagcgggttcaacgacatggtatagttggcaaactcgcccggtccgccgcacaaatccaaaaacgtgtcaacgtgtcggcaaacgt
gaaactttttgtcgatctctgatagttttcgccaacatctaggtctgcgcgttgggcgtttgtcaaataatttttgagcgagcgcaaaccaccgacttgctgctgaacgtgttcaaaccatctttgagtttatttaatt
tttgctgcaacatttttactcttcgtgtcggtcgcaatgtttgtgtcgaaaaagacggccaacacgctcagcaaaactatacaaataaagaacaaaaaatacgtacgcaatattaacattgaccgtttgatcgt
taaatcggacgggtctgttcagagccgctcttattctctcgttgtacattgttaaagtttttgtttttaaattgtacacaatcggcgtgttgtagtcgaaattttcaaaatcggctttttgaaacattgttctgaacgtgt
tgtcgagcggcgtgttgctggccacgtttataatcaactccctccacgctaacgaacggtgctctggcgacacttcgatttcgtcgccattcagtatttgccatcggatagattcccacatatcgacaacagc
aatatggcgaccaaaagatctttgtctggagaaagtagcggtgagccgcttataaaacgcatggctatggccagtagccccaagaagattagagaaaattacaaaaggattagtggtaaattgatgagca
aaatgacgttgagtatcgataatgaataccattacactttcgaataatgtccgacaacaaaattcaagaatactacggcgattcgcaatctttcaaagacatggaagaaggcaagtgctatgacattagttt
aaattatgtgaaaacaaagtttagtcaaatgattcagatcaacgaatacaaagaatgtgaaatggaaattgaaactgcaaccccccatgagcgactatttgactaacaaacatttttgaaaatgaagatggc
gtcaatattattgtaaagtacaaaattatctacaagaaaatcaacagcggtttgtacaaagttgtttttgaggttgtgtacaaaaatctcaacgacgatcccgacgtggttcaagtggaatgttctgtcaatgcg
aaaacactaataaatctattcaagaacaatataaaggggtctgacgacatcaatgaggtgtttaaatatttaaaagacaatgaaaaccaaatctttaccatttacagtataaagtgtcaacaaatttttaatg
gatcaaacgtgtacatgaattggaacgtggtcaactctacgcgcatcgaattgtgcgaggctaaagaaagcgaagcatattcaaatttgcaaaattgcacaaacgcaaagatcaacattagtcgatcgaa
caaacacgttgcttcgtacaatgtaaatgtgttaaaaagcgaattggaagaaaatgacatgggcgataacaagttattgtgcagtttaaaagcgacgagttgaacattgccgattcggatgactgttctacc
tcttccgacttgggcaaatggaacaaatctgttttctacgtcaacacaaataaaaaaaccgaagcggattcgctacaaaagctatgtgcggatttcaatcaaatctcaatgctgctggaagatatttaatc
aaagtaactatttacgtaactgtcgagaacggcgagaatcacaacatgaacgtgttgggactgttgaagtatgacgaagatgaaaatgaatataaattttgtaattctattcgacgtttggttgaacgctggc
cgctgttgttcgcgccaacgtttcgtacccttttaagcatgctagttttccactcaaatatgtcttgagctttacgattatatttaacaaataaagagtttaattcatctagttgtgtgtttaatttactcatttcagatttg
acagcctgatctgcaatagactccatctctttaattttagtcgtttccaaggctgaatatttactagacatttcgtcgtattgttgcctcaagttttggaatcgctcgtcatttaaatttctgttgaaagccggatctc
gtgagataaactctgatattgactcaaaatcagtctgattttgctccaataatctttgtaaatcgttaatttttatttcgaattcttgtctaagtttaacgttttgcttttctgatcgttcatattgctgtactaaagcgctt
acagcacttgtgacagcaacactttccagtaattctgcattgacattttctacttttacttttataatttttcttttaaaattcgatttacaatatttcgtggtacaatagaaattatttgactttttaaatcactcaattcg
gaaactgattgcaaattggcaacgtcggcgtttgacagtttagacctcaatttttgcaaccaatctttgatataatccgtttgatatatacttcttgtgctgtaagctgtggtttagaaatcggcacgtaatcagg
tttttcaatctgcataattatttgataaaagctatctataaacgccaacgtgccttctaaaagcactttagattcttcattttctgcaaataaactttgattggttgtcgcgacttgagacaattcatttttagttgcag
ttaattcatttacagtttgttgatatatagcactcaattctaaatattgagcgtttttcatggcaatttcttgatctttaatcgataccgactgtttgagttccataatatctttggcgtcaagtttgtacctctgtgtagtt
tcgtctatttttgaaatgctgcttcttagttcccggtttaaattttgtttatcatccaattgcttttccaacgattcaattttatcgttcaaggtttcgacgtaatctttatcaaatgtcgattctattcttcttttgagtttat
aaatgtcctcgttcaactggcgattcatgtttttatattcgtccaaatcgcctttttgcaattgtatttcagacaatctttcggcattttcatctctcaatttctgtacggtgtcgttgagttctttgttggatctctctaat
tgtttaatttttcgcacgttttcactttttgttgtgcgaattcactttttgtcttcaaatatctctttatactttaaatccaactcgtcgtattgcgattgcagagtttccactctttcgtttgcaattctatgttgctcgctga
tttgttgcatcaaagcgcgttctcttttttaattgttcttgcaaatttgaaatgacgatatcgctgttgaccgtttccacggtcggaggagaaagaggagccgcgctttgaagtatgctcgtttgaattcttgtgtaaa
gttcgttagagtattgggtcaattggtcaaaagtggtgaactgaaaatttgttcgcatcaccaaataaaagactctagtaaaaaaatctataaacgtaactatgttagtgcacagatactgatctattgcaaaa
cgtatgtttgtttttttggtatacgcatctaataaattctgtaaaattgtaatcaattaatgttcgaacggaagttatagacgatataacgttaatcatggtcacgttgtttactacgtgcacgcgtacgatcctgatt
aacacttcaacgaaattttgaacaagttcccaagtcatgttcgtttgcatgtttttaaataaagccgccagttcagactcctcctcgttagtcaacgtgatttgattcaaactttgcgagccgccgacgcccgc
gacgttcgattctggcggctgtgttggcggcggtgtcgaaaacggatacggcggataatacggataattgttgtaatacggtggttgaggtggcgcggggggttgaggaggcggtggcggtggttgcatcgc
gggcacgggcgccggcggataattgtaatcgtacttgtagttgatgttgtgcgtcaacgtgttgatatctcggccggcgaaatccgataaaaaattgcataatcaacgccggcagagcgtcgggttcaaagct
atgtgtgttcgcggacggtttcaaaattacaacgactttgtgaatctctcgcaaagcgtgctcataccgctccagagttttgattcgagcactcatggtgtttatggtgtttaacaaatcgtgtacagtgcgcgtg
ttgacgtctgtgcctccatatttgggccatcgctgcatatttaaaaatgaaaatatatccttacaatgaactcaaaacgcgctttgcagaatatgcaaaaccaggagaattcaatattacttcagccgatacg
tttcgtatcattcgtttgcattacgatgaaaaacaaggttgcttgtttgcattttgtaatacaaatatcaaagaacgagtgctgcagttttactttaaagtaaaactaaatttgtattcttacaaacaatgctacgac
aagcacatatttccgtcttgccgcaacaagtgcatcagctacaccacgtttgtggcgccgggcgtggaaggaaattatttgaacaagataaacgtgattaaatacgaaagaaacaaagcagcgccatcg
gacaacgcggcgtgtttggacaagtttcttcacaacgttaatcgcgtacacatgcaaacgccgtttgtggaaggcgcctatatgcgattcaagaaaacacaacgctgccaaaataactatgtcggtggatc

gacgacgcgcatgtttaatttgcaacattttaatgaagactttgagttggtcgatgaaatgactttaaccagcggcatcatgcccgttttgtcgtgctatgacattgagactcattcggacggcccacaacatgtc
gaaagcatcggttgattgcataatgtctataggatttgtggtgtacaaaaacgacgagtatgcaaaattttgttttatgtatcacaagctgccgacccagattccagagacctatgacgatgacacgtatgtgg
tcatgtttcaaaatgaaatcgacatgattacagcgtttttttgacatgataaaaattacaaatcccgatgtgattttggattttaacggagacgtatttgatctgccctatatacttggacgattgaataaaaccaa
aatgctattaaagcgttacgatttgccggctgcggcgccgacgacaaagctgtttataaacaagttgggcaacaaagtggacacgtattattttaactattacattcacattgatttgtataaatttttcagcag
cgattctaatcagcataaagttgaaaactttcaattgaacacaatcagtagttactatttgggtgaaaacaaaatcgatttgccttggaccgaaatggtcaagatgtacaacactcgacggctggacgtgatt
gccaagtacaacgtgcaggattgcatgttgccgataaaattgtttgtcaaattgaaaatggccgattctgtatattcgcaatgtatattgcatcgtttgtgtacagacgacgtgatatgtaatatttcgcatttgat
tagcgtggcgtgcttttacgcagccatcacaaacacgcgcattaacgagagcaccggcaaagaggagcctgacccgtattttttaacaaaaacgatctgtcaataatatcgggccaattcaaagccga
caaagccgcagcgggtatttctaatctcaaaagaaaattgatacctctaaaaaacatacctaaagatgccatcaatttgggacccgccaatcaaacggtcaagtacaaaggcggcaaagtgttaaaac
ctcgtgcaggcatttacaagaacgcgttttctctcgattttaattctttgtacttgactataatgatagccatttgtgcttgtttgtccaatttaattttatgcgaagacggcaacgtgtatttgaaccacaattcgcg
tgcaatagttgttaaattgctattaaaattgttaagcgaaaggtgcaaatttaaaaagaatcgcgacaatcaaagcgaatcggcattcttgtacgatctgtacgatcaaaaacaaaattccgtgaagcgcac
cgccaatagcatatacggttattatggcatttttttacaaagtgcttgcaaactacatcactagagttggtcgcaatcagttgcggctggccatttctctgattgaaggtttgagcaacgaccctgagatttaga
aaaattcaatctgggcagcatcactttaaagttgtgtacggcgacacagattctacatttgttttgcccacatttaactataacgaaatatccaacgaaaccgacacattaaaacaaatatgcacacatgtt
gagacacgcgttaataactcattcacggacggttacaaaatggcatttgaaaatttaatgaaagttttaatactattaaaaaagaaaaaatattgctatctcaacagtgagaataaaatcgtatacaagggtt
ggttagttaaaaaggatatgcccgtttttatgagaatcgcgtttaggacggcagtggaacaaatattgcgtcatttggacatggacaaatgtttgcaaagtttacagactagtttttatgaatactatgatgagttt
gccaaatcaaaatcgttgaccgactacagctttagtatgacctataacgataacccgggcaaaaaaacgcaaatcaactgacgataatgagggtccgtcacccaaaagacgcgtgattacggtggctag
gcattgtagagaaattttggttaacaaaggcacggattttgtgcccggcaacggagatcgtatcccgtaccttttgatagacatcgaaggcaaggtaacagaaaaagcgtacccccttgcgtttgttcgatcc
agtgaaaatgcgaatcagctggataaaacacatgggcattttgtgcacatttatgaacgagttgctggaaatatttggcgacgaacaaaaggataaaattgcaaagtgttttaccgcgatcatgcaaaagta
catgcagaatcaattgtacgatcgaaaagaaccagtgttagtgaaaattaaccaaaaaaaatgcagtgtcaaacgcaaacgtgacgacgacgatgacaatgatgatgatgatgacgatgattgtgatagt
tcagacagcgaaaatgacactcaatgtgctaacaatacgtataaattttgtttgtataaaatgaaaaaataaagtttataattatatacgatactatgttgttttattgtatgttactgtaaggtgttattatgtttca
acataacgtttttgaccaagattacgacagcggttattattccgattagtgttctataagtataatcataatgattgcattattaattgcattgtttgctgcgattcacgcgccggcagtgcgctctcacggctattt
gtccgtgccgacggcccgtcaatataaatgttttaaagatggaaacttttattggcccgacaatggcgataacattcccgacgccgcgtgccgcaatgcttacaaatcagtctattataaatatcgagctctc
gatttagaatccggggcggccgcttccacagcgcaatacatgtttcaacagtatatggaatacgcggctgtggccggtcctaattatgacgattttgacctaatcaaacaaagggttgtgccgcacacactt
tgcggtgccggttcaaacgatcgcaattctgtgtttggtgataaaagcggaatggacgaaccttttaacaactggaggccaaatacgctatatttaaatcgctatcaaccggtttatcagatgaatgttcatttt
tgtccaacagccatacacgagcccagttattttgaagtgtttattactaaatcaaactgggatcgtcgcaatccaataacttggaacgaattagaatacataggtggtaacgattcaaatttaattccaaatc
cgggcgattcattatgtgacaacagtctagtctattctatacccgtagtaatacgtacaggtcaaatcaatttgtaatgtatgtgcgttggcagcgcatagatcccgtgggtgaaggattttataattgcgccg
atctggtgtttgaaactctagacgacgaatgcagatacgctcaaatggctaaagtggtcagaagccaattacaaaagcataaacttgatgcacgcattgatcataatgacgaagaatcttgttggcgtgca
cgcaaatcaaattattcgtcatttttaatccaggattttaataaaacaaacaaaattttaattacatattatatttagcaagaagtaataacagatctctacaatcaatacgaacaatttcaaacatgtcttcaa
attcgtagtctgctatgctaaaattgttaattaaaaaatattcttccatattttcacacctgtcataattatgtaataaaaatagtttaaagccatcttgatctcgtttggatatttcgtcatcgtcaaaaagatattg
caaagcaaactgtacttctttggcgtaaggatttaacagtccatttttgcagactacgctgtctacgagcgtattgagactttccgcagttaaagtgttgcaactgattgccggtttcagacgagtttcacaaaa
ccgcttgatttcatcggctacttgattttcaaaatccatttgagacattaaatatttgtagtattttgcttctccgtgcagcccattcaagtgaagatagatcataaatttatcgtacatgttgctctaaaattaataa
aacactgattgcatgcaatcaaatgcgctctatttatattatcattcaatgaacatgtcgagcaaatgttctggaaacagaaactcagagggcaaagcgcaaatattcatgattgtacgttcgtcgttttgtaa
caaacatagttatttttaataaacacgctctgtaacacaaccgtgttatgatatatgtcggcgccacacactttaccttattgtgagaagatatgtaactatttaaactggtggtggctcgtgaagacacagttt
caatatctgtatttaaagtctgtcggtattgttgcgtgcttttaatcagattttgactgcctttagcaccgcttttaatcatgtctacaaattggccgcctaaattgaatatttcttcatctccgcacaccatttcttcgt
tagctatgagctgtgtaagctgtttaaaaagtaaataactagcgtttgaactggcaattaagcaaaaatcacgcaacaacatctctagcctctgtgcaaactgctcatgtttatgtttttgccacagcgcgtaa
accaatggtgtagatttaaacaacgcctccattgcgtcaatatttttgtacaaataataaatttgttgtgacacaaacgacaaacggtttttgtcaaagcaaatgaaattaaaccgagggtcgccgtacaaaa
ggcattccgagtcgattagggaattaggtttgggtaaaaatgtaataatttttttgtcaccgtcgcagtctgtgttagcgccgggaaatattcccagacccactttacattccagtcggtagatgactcgttaat
gtggacgtccgagacttgcgtactcagctggctaatgttgggatgacgcgtggtccatgcacgcacgttgtccacatcacgtccataatagcgctttacgctgctccgcggcggcataatttcgttgacacc
atttaaacattgcacgttggcgtagaacgaggccgtgttgagaaacgtcgagtacaagtattggaccgcgtacccgtttttgctctgcaactgatctttgatgacaccgtgtgtaagtttgattttctgcagagc
gcccgaaatgtcaaccagattgttttcatgtttagaattgaaagctttgtttaaaaatataacaaaattgtgatcccacaaaataaagtggggcaaaattaaataattaatagtgtctgtaaacttgttggtcttc
aactttttgagaaacacgttggacggcaaatctgtgactataacacaagttgatttgataattttagccaacacgtcgggctgcgtgtttttcaccgacgcgtctgtgtacacgttgattaattggtcgattaaac
tgttgaaataatttaatttttggttcttctttaaatctgtgatgaaattttttaaaataactttaaattcttcattggtaaaaaatgccacgttttgcaacttgtgagggtctaatatgaggtcaaactcagtaggagttt
tatccaaaaaagaaaacatgattacgtctgtacacgaacgcgtattaacgcagagtgcaaagtataagagggttaaaaaatatattttacgcaccatatacgcatcgggttgatatagttaatatggatcaa
tttgaacagttgattaacgtgtctctgctcaagtcgttgatcaaaacgcaaatcgacgaaaatgtgtcggacaatatcaagtcgatgagcgaaaaactaaaaaggctagaatacgacaatctcacagaca
gcgttgagatatacggtattcacgacagcaggctgaataataaaaaaattagaaactattatttaaaaaagatttgcgctttactcgacctaaactttaaacacgtcatagaatcttcgtttgacaaaaacca
cattgtggccaagctgtgtgacgcgacgcgcgctaaagaatggcaaaccaagtcgcgcgagcgtcgacttaaaaattttaatttaaacattaattacgacgggcccgtaaaaatatttgtggccgccaca
gcggagcaaaagctgttacttaaaaaaactcgcgacgctctgttctgatgaatcccctaatgattttggtaaaaatcattaagttaaggtggatacacatcttgtcatatgatcaaatggtttcgcgaaaaatc
aataatcagacaacaagatgtgcgaactcgatattttacacgactctctttaccaattctgccccgaattacacttaaaacgactcaacagcttaacgttggcttgccacgcattacttgactgtaaaactct
cactcttaccgaacttggccgtaacctgccaaccaaagcgagaacaaaacataacatcaaacgaatcgaccgattgttaggtaatcgtcacctccacaaagagcgactcgctgtataccgttggcatgc
tagctttatctgttcgggcaatacgatgcccattgtacttgttgactggtctgatattcgtgagcaaaaacgacttatggtattgcgagcttcagtcgcactacacggtcgttctgttactctttatgagaaagcgt
tcccgctttcagagcaatgttcaaagaaagctcatgaccaatttctagccgaccttgcgagcattctaccgagtaacaccacaccgctcattgtcagtgatgctggctttaaagtgccatggtataaatccgt
tgagaagctgggttggtactggttaagtcgagtaagaggaaaagtacaatatgcagacctaggagcggaaaactggaaacctatcagcaacttacatgatatgtcatctagtcactcaaagactttaggct
ataagaggctgactaaaagcaatccaatctcatgccaaattctattgtataaatctcgctctaaaggccgaaaaaatcagcgctcgacacggactcattgtcaccacccgtcacctaaaatctactcagc
gtcggcaaaggagccatgggttctagcaactaacttacctgttgaaattcgaacacccaaacaacttgttaatatctattcgaagcgaatgcagattgaagaaaccttccgagacttgaaaagtcctgcct
acggactaggcctacgccatagccgaacgagcagctcagagcgttttgatatcatgctgctaatcgccctgatgcttcaactaacatgttggcttgcgggcgttcatgctcagaaacaaggttgggacaag
cacttccaggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagttttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctactagctca
aaatttattcacacatggttacgctttggggaaattatgaggggatctctcagcgctctgttgccgttttacaaatacatttccatttgtaaaaacggtgttatggtgagacgcgacgaaaagagtcgcgtgttta
ttgttaaaaatgagcaaaatatagaatatctcaaagccaacaaatattacgcttttcatagcgatagcgtcgataatttgagtctgaaaacgatagcgaaaaaatgcttcaaaatttaatttaatttaattaat
aattaattttagtatcaagttgaatctcgtaatgtaagaaaaaaaacactaatattggttgtgattcttgccagacgtccatttatgtgataaataaaacacaaacacaaagtttttatattgtcttttattgatat
gttgcctaaatattatcaaaaccttccctataacggtaagaggatatttgaaaaattctacgatcgtagtttacaaaaatataagtcgacacagatagctaccaagctagcatgttgtgctgttaggaaaaaa
tacatgttgatagacggtaaatggcagccaaggcctgacgccaacaatagcgacactactactagtacagaggattccactacagatacagaaaccgaatacagtacaaccgaagacgaacttgctta
attctaattacaatgtatcaaattcccgatatgttatacaatgaaaaaatgcctcctcgcgccaaaaagttattcgtcgagacgtttacaaagtatcataaaatgaacggcggcgacgaagacattgctatg
cataaagctagaaaagcgctcgaagaaaagtacgttaaaatagatacattgcaaaattcttggattccgcgcaaagccgcctacgagatagttagggacgacatcgacgaaagcgacgataacgcgg
acaattctacgaaaactaaatcaaacgatcgtctcatagaaaataacaacgaaaacactagcacggattatgacactgaagacgacgagcgcgtcaacatgtttagtagtcgtaaacggcgcgtcgct
aagcgggcatttaatgcgaaatccaaaaaatttcctatcggcgaagtggtgtcgacgcgcaaacgattaaaacaaaacacaaacacaccaccgcattatgacaccagcgaagatgaggacgaagata
attattacaactattaataaattaatttttttctacttgtatagttttatttctagttcacgatccaaggcgtctcgcacttttgggtccgactcgattgggggccgaggtatgcagttacgtttttgaaaatccaaaa
atccggacacgcgctgatgatacgtttcgtcgggcgtcagcttcttggccacgattatgcgcgcgtcggaacacaagtcttattatcatggagttgtcgttgtcggtcgcgttgaaactgataaaattaacaaa
gtttggcggctgctcaaactccatttgtttcacaattaggcgagctttttcgtacaccaacacggtcagcgtttcaagcgacaaatcacaagtaatgtgaaagtttttgtaaaacgcctttttttacattcataaa

aataatgtactcttcattttgagcgttatcaaactctaaaaatattaaactagtcaaatctaacactacatcatcaaactccaaatatctagcgtaaacgggatgtttcatcgcgcacgttcgactgcctgata
taatacctccgagtgaaataaaccgcttgtcgagcgaaaattgtatttatatgaagctcaaatcccttttggtagggaaaatacaactctatttctaagaaccttaaacccaatatctagcgtaaataaataa
aaccacaaaatacaaacaacaaactttattaacacagagggtaaaatagggcagaacaattttcatgattattccaacttttagagttggtcacaataaaatcgttagggcaccgatcggcgaatttcgtgt
taaaatttttatgcatttctatgtattcagccacgctcacttcggccgacgtcacaaaataacactttggacataacactacataataatcggagggtttttgctgagattgtttgtactcttgaaacgtcgcatg
attgatcgcagcatgtcaatatttttcaccaatttgtcaaatttcgtgtaatctcgcttgccgaccacgtcgtacgtagtgtgctggcccactttatttgcttcaaaagttttgtaatattcgcttaaactcaattttg
cactggcgcccgttgacaatgacgacgacacactgttgcttttatatttgctttctactgtgttttcgataatttttccaagcgctactttttcatgattgtggatacaattgtcaccacgtcggaccaattattgag
ataagaatgcttgtttaacaaataggtcagctgttaaatactggcgatgctacacgttgtgtgcgtgcagaccaaaataactgctgaattcacccataaagttttcccaatccaataaatgtaattgagacac
atctcgagcgtgtagctggttgtttacgttttcgcgtaaactcgcagatccgccgccggtttctgattggcgcgcgattaactcgctgtatcgtttgtacttgttgggatttcgaacccactttgcgtaatgaaaaa
ctttttggtgatctcacgcggcgggcacatgacatgtcccggcgggacgccgccgtaattaagatcgcagcactttacccgatctttattgtttggacagctgcccaaattcctgactataattgaaataga
agtattttcggcaattacgtcgaatggattgttttgtttaagaattgcggtcaagactgtattgtacaaaagaaaagtgcgataaatgtattgcgactgacgtgactcgtccgtagaaacaatacaattgttcgc
cgtaattaatttccaaggatattcgagttcattacaaacgtcgacactttgcccgctttcactatcgctttctgcgtaatctcctttgataatgtacgcgctttgaggcaaaataatggccctaaattcggaatag
attaaatcgacatacatgcggtccgctatcaaaatgtactcgtcgtcggccagtttgccaatgtacttgatgaccgcatcgatgatgatctttaaaacgctcatgttgtgttcgccagcctcgtcgacgcttcta
aaatgttccgtacgaattggtctaattaacatgcactcgtacacgctgtcggcgttgagaattgtttttctcgcctcggcttgatcgttgcgcaacaagtgccgataatacaatttcgcgtgctcgttcaaaaac
acgctgtttaacacggtcatgtgcaacatgtcgttttcgttatcgggatctgaataattgttacgtggacgcgttactttgcaattcgctcgtcgaggatgtatccagcattgcatcggcttgggcggccgcatcg
gtttgaagggcgttaatttcaccgaggcacagaggtctaacttgatcggtttcttggtcgaacacatttaaaataatgtaattattatctatcaaaaataaattatctttcagtacacaattgatcaggttgacgt
ccgtcgcgaaccatacgttcgtcatgattgcattttaaaaaaatgcctaaaatcgagttcaatgtccagggcgtcttcctgtgtaacgcgttgataaacgcgtttgtaaacgcgatcaatttcgtgatcgccta
taaaccctttagtttctcgaagcatcacgttgcactcctctagtgtttttgggggaaagttaaaataatattttataacacgattaaagggatctcgagtatgttcgcgacgccaacgctgaatacattgttcgtg
aaacatcttttccaagtttaacatgccagtgtcgggcattgccactatgcccccgttatcgtcaattttctcaagacatatctggcattctgtcgtgcacgttggccaataatctttaaatagtttgaacatatagt
aatttctgtcttttaaacaaaaggtgcagaacataatgacgtggccttactccaacagttttaaacttatcaaagcttatcaaaaacactctaacacgtaccaaagttgtacaaatcacaaatataaaagtgt
cttgttgaaaaagtggacccaagacgtcatggacaacgaacttcgcatagaatctcgtccgtacaaaaattgcgacgcattttgccaaatgtgcttgtcggttgtgacaattaacgagtatttgtgttgtgaca
agtgtttgtttcccataatagactatttggaaaaacagcgacaattaacgccgcgacagcaattgtacatgtttgtgtttctgagtatttgttattggaaggaggccagcggagcgacgtttggcgacaatgaa
aaattaatatggattcatcgttttagaatagcgtgggagagcgcggacgtgccttttgtttgctattttgcgccgtccacaatatgtttgcaatgctcgacggacaacagacgcattgtaaaaaagtacgacac
tagaaaaaaacttgtaaattttaatgtaaattatttttgttacaattgttttttcctttgttcgatattttaatttatgtgtaagattttaaagcgtttcttaaatcctcaatgtccttgcttattatgtcaattcgctggttca
ccatttctgttttggccacgttgtagtactgaaccacggaaaactctgtaactgtaaatcgagtgtcgtcatttacagtcatggctttttgtaccgactctgagtttgcatactcgacaataattcggttgtgtttta
aaggacatacgtatatgtgtaatatagtgccaaattgctcaaaacaaggctttacaatttttttcggcaatttcttgtctagtcatttttttgttcgtcttcggaaagagttttccaatctatttgaattctgttcatagttg
taggacgttttacaattatagtttagtttttgtttaattgccctgttatacaatttttgtactctgccaaaaataaattgatacgaaactacatctgggttgtttatatccgacttgtcgggtttattagtaaaaaaatt
gaccaaagcttttctgagtttttgtatttcttcatttacagattcgccaatttgtaaaaagaccgtaatcaatggatcgtacgattccaagtctgtattttttttcaattttaacacatcgtcatagtattgctttttgac
actattactaagcactaaacgcgcatttttttattaaaatcaacacgttacgcaccgtgttgtaatggtgcttttcgtcgatactgttatacggtctaacttgtttcgtcaccacgttggccattctgtacactgttaa
aataaaaatcgttgtcaaacttttggtccggaacgtattgcatttttaaacattgataataatcaatttcttccaagttgcaaaaccccaaaaaagattcatttttggcatgacagtgttcgttttgtcgaatgcgttt
ggccgtttcatccgaaacaaccgaccgttcacgaacacgtttagacataattagtgcgcagtaatcgcttcacgatgacggacgtggttcaagatttcaacgagttgtacgatagaattgaaaataaatac
aaattaaaatatactttgattgcgctacaaacagcaatgaacgtatgcttttttgcgctattcaagaacgcaagtcttatttgtgttgcgctcttaatgaacagctcagttgcgtaatgcataaatgtgtgcccgt
tatattcggcacgcgcttagacaagcaatttagagaaacggacgatgtcgacgcaaacaacaatattaacggaacttttatgctggacggcagatttttgagttttcccaacataatgatgaacaacaacgt
tttggtgcacaatttctacgacaagctatacgccaaacattgcaaacgcatgtttttgtacggcaacgtggatcaagagaagcacatcaatcgcgcaatccaactagtttacgacaaacaagacgacgtc
ctgtttgctcgagacgtgtacgctagcgattacgtagtgacggaagatttaaattctgtgctggaaacatatctggccaacagcggcaaatggaaaccgctcgattttctcttcgaatacaacacgcttcaca
agcaacaactcgtggaacacataaaaattattatgaatcacgacataaactattctatagacagtttggccaacaaaattgtatacaagcacgcctatttgatagaactgttattaacttcgacaattttgca
aaactatcaacgagttctcgacaagacggccgacgacgacgcgtacacggtggcgaacaagcgtcgtaaaattcaaagcgtcttgtacaacaaagagtcgaaaaagattgtcgattgtatagtcaacg
ggcgtctcatatattgcgtatcgaaaacttttagcaagcaacgcaaagtctttcccaatcaacaggacaacagcagcaataacaatattgaaatttcgttgcccgtgttaaaatatcgcgtcggcaacgag
gtggcgcgaatcaccaacgacagcatgcgccagaaaatgttgaagcaaaaaaaggattttgtgaaatttattggcagttttttttcacggcgaaatgaccgtggcgggcaaaaagtttttccttttgccgcaac
gcatgtttgcccaacgtggactacgaaatggttgcgcaaaaatttcaatatttgctaaaacacaatctggtcgcgttcgtggacgacctcaacgacgttcaagacgactcgttgctgatcgcgtttaatgata
gacccacaaatttaaaatgtttaaaatcaaacacgtcgtttatcgtgtacactatgaaaagaaacatggcgcccatcgaattgaagattacagacagaatactttacgttaatcatcacgaaggaatgatat
gcatcaaaaaaaaagttgcgggtgaacaacgaagccgacattaacgtattgttgacaccctacgaataccactataaacattctatatactataatcccattattcagtgtacgatcgtagaaaacgatgac
gtcaaatcgctcatgtccaagttggagcagtattattatcgtaattttatacatttatttcataccacgcccgtacccaaattgattgtgtccctaaccaatctaaagaatgccatgcccgtgtttgagtacaaa
gaggaaaattgcgtatcgggtttaccaaatggctactctgtggcggtaaacaagtccattttgctcaacaataaaatgtttaaattgtggacgttggtgcgagacaataaactcatgaccgccgaagatccgt
acattccgcatattgcgctgcccatttgtttatacaacaacaaagtgaacaaattgaagggcaagcttgtggtcggccccaagcagtcgtgtttagtaaagtttacgaatagtagcgacaaaaattacgttg
cgctagacgacggacttgtgctgtacatggccggtgtgttggtgagcaacgcaaagattaattgggtgtacgacggccgacggtacaaaattgaaacgtgcaccaacggcaattttaacgtttacaaagtg
tacgtttactttagacaaattaaaaatcaaaaaattgaaaaactcgacgccagcatggtggttaacggcgacaacgttatgcttaaaatagtcatagtcacgtccaccaacgatttggaaggaataaaaat
atgtggcattcacggtcagaaaggtgtgtttaaccgcagcgaagacctgaccgagtggatggccgaagatggtacgcacgcccaaatatgcctgtcgcccgtctcgtttttgtctcgccagtccaatttga
taaaattgagcgcaaatatgtggtgcgcggtggaaatcatgacgacccgcacgccaagcgttatcccatcttcaacataccatacatgctatttaataatacacccgacaatattttcaaagaatttattaa
aacaagccacactggacacgaaaaagtcgaaggcactcgcttcgatcaatggacaaagaatcaatcgttcgtaggcaacagaatgtcggaaagtttgcattggatgcgcggcgggtccaacttgccgc
aaaactgcggcgagttcaacgtggtgtccagcctgttgatgtgcaacaatacgataatgaaaaattgataacgcttgcacgattgcaaacatgcacgctcggttgaataaaagctcgcatcgtcgtcgtaa
aattagttgtatcaaagagcagctgcaattagaatcactgctaaaaatgttcgaagcggaatttaaaaccggcgcagttttgaagaggctggtggaaacttttaaggatctgctgccgcacgctacttttgatt
gtgataatcgaggcgtaagcatgcaggttatggacacctcgcatgtcgcgctcgtgtcgttgcaactacacgccgagggtttcaaaaaatatagatgtgaccgcaacgtgacgctcaacgtgtccataaa
cagtttaagtaaaattgtcaaatgcgttaacgagcgcagctctgtactgatgaaagcggaagatcagggcgatgtgatggcgttcgtgttcaacaacgacaatagaatctgcacgtacacgctcaagttga
tgtgcatcgacgtggaacatttgggaattccggatagcgattacgattgcgtcgtccacatgtcgagcgtcgagtttgcgcaggtgtgcaaagacatgacgcaattcgaccacgatattattgtgtcgtgttcc
aagaagggtctgcagtttagagcaaatggcgacattggctcggcggacgtgcagatgagcgcggacaacgaaaacttttccgttctcaaagcaaagcaaaccgtaacgcacacttttgcgggcgattac
ttgtgtcattttgctaaagccgctcctttggcgcccaccgtgacaatatacatgtccgaggaactgccgttcaaattggaatattgcattaaagacgtgggcgtgttggcgtgctttttagcgcccaagattgtt
aacaacgacgaggaaatttttttaattgcgccgcatttgcagtacgaaatggatgcgctgcatggaatttgcgtcaacattaataaatttttataaatgccgcagaattgtaatagaatacgacaattgtagcgc
ttcgtttacttttaattgcgtacacgataatagaaaagtgaattgtttagaaatagttggcttgagaagaaacgagtacgtgtgtttgggcaaaataataaacggtgataaaattatatctgtgcacgaaaacag
tgtcaatgggaaaataatcgtacctgtcgaagacacctttgattttggtctgtttactttaaaaaaataaaattaccgacgctgttataaaactaaatgtttacatcaatgaaacgagctaacgtgaaagtaaac
ataaacgtttaacgatatatatgcaacaaccttttaatattcacattaaaaagtgccaggatgatggagcgcacagtcactcgctggcatttggtgtcggataatgtattatttaaaattggcgaggtcgccca
acgactagattattatttgcaagaatacgcaaacttggaaatgcaaatcgaggaagaaatcaaatatatggaagtggatgacggcgaagaaatcgacactatcaagacttttctgcgcaattccatgtcca
cgtcggaacagcgtgatttgtacgcgttagcattaaaattaaacagtttaataaataaataatatttaatttatttttgttttattatctattttgtaacaacaacaatatcagtctttgtgttaacagtcagtcttttg
aatatcacattctccctaaatataattacacaatttcaaaaataaattgtaaatcaaacgtgctgttgttcaatttgaaaccgtttccatacacaaagcaatttgatggagagtattgcaaagctggttcgccg
atagaggtgaaggaattcattataacgctaaatggtatagtagtgtcttcgttagccgacatgaaagacactttgttattttctttgaccattacccagggataattaaacagggctttaacgccgttgcccacg
tttacataaatgttttcgtcggaattaactctaaaagctctcagggaatccgaatgaccttcgacgtttaatttgtaacaggccgaagccgtggccgtgtacattgtttgaaaactaattatctgctcgactagtt
cgacgctagaattgggacttcgaacggtcagagaaggcacgttgatgtttgtgataggaacgcgatgacagaccacaccgttgaacgtggtgccttccgacgagttgttgttgttaattctttgcgcttcaaat
gacaaatcgttggtgttgaccgacgtgtctctgcttagcacaatcacccgcccgtttgtgttgttgagcgagtttggttccacgccaatgttgtaatatagctggaacatgccgtgatcgtagaacaacgtgca

cgatttaaattgcaaattgttcatttccgcaaacttggcgtacaccaaaatggcgccggccgtgtcggttttggctatagccgtttggccaatggtcggtctgaacgactgcgtggacgtggtgcccgacggg
atgcgcatgattccgttcaaactttgcaaaataatacccgactcaaacgacaccgtgttggcattatagttgataattctgccgcgccgattccaaatgcgccgcgccatggtccacaggggcgcttgaatg
ttgttgtgggatcggattcgtagtaagccaacctattcgttacgcccacaaccgaaccgtaatatgtttttgaaagtttggtcaacactttggagtaatcggccgaatggacggccaacggatacgtaataa
agttgcctatcacgttagagtacaacgtgccgtttcgagacatgacgcctggcaccacaactccctcgggactgcccacgttttcgatggctctcgtcaaacccaccgtgttgattacctcgtctccaaaat
agtacgtgtaataggcaaacgtaaagtatgaatttatcaaatagccgtaagcgcgcacgtcaatgtgatcgatgtatatcgaatcgacgtgcaagccgttgccttgagtcacgtacggaaacgcgatcgtg
tttaggatttcttgtattccctgctcttgcctaatttgcgccaatgaatatccgcgcaacatttgactgtacgtgtagggcacacccatgcgcattgagttgcctgccgtccggtgccagcccatcgagttgacg
gccgtgggcagatacaagccgagccagtaacgcgtgagggacgcagcttcgtcgtaatgctgcgtttcgtttagcacaatggtaatgttcataaacacctcgggcattgtgattgtgaaatggtaccaatcg
gcgaccgggccccagggcgcttgttggtgcggcggcggatcgggcagtttgctgcaaatgatgcgcagcccgttaatgagattgtaagccaattcagggttctggtataacgtgtcgctggtgttgttgtagc
gtacgcaataccctatgagcgtgtggcacacggtaccaaagtccggaacgcttgtccacggactcaaattgtcaaataccgtgttgtcgttccaatttctattgggatttaccaccttttcagcttttggctaa
acgaatatgagagattatttttgaaccatcgttcaaagttttgtagttcttgaatgttgttttgccgaaaggccactatgctgttgttgttgttggtagtggtggtagcaacggcagtggtagtagtagcggttgtagc
attgagcggcaagggattgagatcaataaaagcattgtttttattggcatcatttttattattgctatttgataggtacaaaaaacatattaaaaatattacaactatgacaataatcaatacgatagacatgtcg
aatgttgcttattattgatacaaatttttcttaattgtaaacacacagtttttatatttttgtccaaccactcgcatttgccgaacgagtgttttaaacaaatagacgccaacaccgattcagttaacacaaacata
attttgctatttacataaaaaccaaattctggactatacattttaacaagttgctgtagtgtaatgcgagtgtcaacctcgcgcagcaattgcttcatataactcttgtttacaaagcacattaaaacgtcgcgtg
taaccacataccgaccatcaaacgtcgctctaatacactttcagttttattataaacatttattatttttttagttaaaaaattaatctttatgaactgtttgtagtaaaaatagggggtcgtcgtcgttgtcgtcacg
cgctcgccgcctgatgcgttcgttaacttttgcagagttgctggcaaacagcacactttttggtatgcaaaaggaaggtttaaatttagaatctgcgccggcgacaagtaaaccaacacggcaatcaccat
ttcgtagggcaacatgttacaaagttttgtattttaatttatatctaatgccgcttttacatttaggacacgtgttgtccacgtttaccgtttcgtcgcagtcggtgcacagacacgaatgtccgcaatgcagatag
gtaacggcggggtaaactgcgaaactttttctgcaacgcatgcatgcactgtcgacaactttcacattgcattctgtgttgagcggcaagtttctatccgttttaatgatttccagcgttttgttgcgaatgcccg
caaagtcgattacgcgtctgttgtagcgccgcaccgtgaacaccttgcagccgtgtcgatcgacggtcgagtcgtacagcggaggtagcgagcgggccaaaactcgaagcgcttgcgtcttaacgcgac
cgtcgtgcatgttgttgagcgcggtgggggtaaattggaatcgaaacgccgttgccaatttttagttccgacgaacaacgtgggcacagcacgataccgtttttaacaatggccatgcgccgttggtaaaac
gcagcggccgaaggggacatgttactaaacgttttcttaaacaagtacaccaactcgtcgcaaacatagcaagtagcatatcgggtgttcatgacgacttaccagctacaagttgaaaaatttttttaatgtc
ttcgtaattgagcagagtctcttcaaagaccgtgttgacaaacgcaatctttttgtcttggcgttgtttaatgtattttcgatgctgttgtcttccacatttagcatcttgtacacgtacgtgtttttttgtttgtcccata
cgactgattcggtcttgcgcctgcaattcaatttgcgggttccaatgaggctccaacattacaatgtggtttccgcctattaagtttaacccgacgccgccgcacttaatggaaagcagcaaaattcgatgtt
gagtattggcagcatcgttgaatgtcgtctcggccaaaatcctgtcttccacttttaattggcccgtgtacattaacgtagcaatgttttttttgtttaaaaaagttttcaaatattttaaatattccacccattgcga
aaccaatattatcttgtcgtcgcttgtgttcaaaatgtcgtctaccaaatcgagcactcgtttgcacttgctgctcatataatccatttaaaacagtcgtttgtttccaaaatatttttaccgtgcatggccaaata
cgggtggcagcagatttgcctcaatttcagtattagccataacacgtgctgcatttgctgcaatcggcttagtgcgttttcgctttcacgcgctgccacagccttcacatacgcctcttccgattcacactttaa
tttatcgtacaacgtttttttcttcttcattaaaattaacatgtacatactcaaccgtgtgtttaggaatgttagaagaaatttcagatttgtcgcgttttaaaacaattttttttaataatacttttttatgcgatttgtagag
tcgttgtttttatttaacatttttccacactcttggattgttaaaaggacgacattgcaaaaaaattaatcatcgagtacatgtcccaatgcttgttgtggatcggtgtgccggtaatgcaccatcggtttgttgcggtc
aaagcgcacgcggcgttgtgcacgcccgtcttgcagtttttgataatatgcgcttcatctagaacaactcgatgccagcgggttgaaaacagacttgactgtttattttgtttgatcaatttgaaatgtgccaat
aaaacgtcgtacgtggtcaccacaatgtggtaatgctcaaacgtgtcggcatccaaagatttgtaatactttaaaatgttaaaatttaaattatgcttcttgttttcggttacccaatgattgattaaagacaaag
gacacactattagagttttcaattgtagagagttgtttttttgcgattaacattagcacagagagcgttttgccgagtcccatgtcgtcggcaagcacgccgccgttgggccggccgtttttttctctgttaatcatc
cactgtatgcctttttttctgatgcgccaatagattgggcgtttgatgttcaaagttgggatcgtcgttgtcgggcgcttggtcaaaaaaattcttgcaattgcaatttgtaattatccattgtggcgacggctattctct
tcgattaaactatggttattgcgcgctcgtaatctaagatatcacatttaacacggacattatctaatcaacgtgattgacgacgagttcatttattgtatcttcaataatttgttcgtaaactggtaatgttgttact
tgtatcgcaccaagttttctaacatacacacgaacatgtttgtagcttgtttctgaaaaatttttgcacaaaatacggcaacccgtccgcgtacatgttagcttgtagcgtgcgcgtgtacacgaccaatttgag
ggccacgacgtttttaaaaaagcgtgtaattcatgtacgtgtaacgcgggcccacgtgttggcgttgcgccagactttcttcaaatttaaaatattccaacgccggatttggatgctcaaactggtcgatgtgca
cgcggaacgatacaaaattgcaggtcggcaaacggccgcataaccaggccgccgtgtcgtctgacacgcacaaagaatccgcatcctgttgcgtgcaatagccctgctgtcgtaaaacgttcaatgttc
gtttcatcattgtggcgatcgtgcccacgtactctagtcgtctgttgaattccgcgtcattttccatagtcatgtttgtcacccggttggagcacaccacacaatttttgccgcaagcgtgcaaactggaagatg
aagtgctggttttgtacgtaatctatttaaacgggttcgattacacgctacctcgttttgtaaagcgcgaagtgcaagtgaacgccgatggatttgttagattcgactacaacataaaaatatttgattttgcacg
ttttacagacatgacgcaaaccacgcccgaggatattgacgactatataaatctaaccagaattaatagcttgagcgatcacgatttgaaaatgctacagcttgtgtgccgagatcggtggtacaagggcg
atgtggcgcgtttgcggcgaattttgcaacagaaagacgttgatgatctcgtcaaatttgtgtgtaacgtgatgtgggagcgcgcctacgaagatcactacacgttgggtcaacagttgagcatacgaatca
caaccaagttaattcagagtggtttggatttcaaacatcagcccgacacgacagcgcccgtgtcggtgcgtggttgggaagacgcgacgtttgaaaagtatttacaatcgatcacgtccatcagcgaggtt
atcaaacgacacgtgttttccaagaaatacatctgtctagaggtggcagcgtcatattggtcggacgctgtggaatcgttgcaacaggaaaattttagaataattttaaattccaaaacgccgtacgtgttgtt
gatcgagatggacgacgacaaaaattcaatggtgtacttaagaaaattggcacacttgttggagaacaaaatagtaaacctttgttcgtcacggatgtagaatttttactttaaaaatggcaattttatgttttat
ttatacaattcgttaaaattttactactattgcctaaaaaacaagtttgcgtttgaaaacgtcgacaaagaaatattctttttgctgtacacaatcgtggctctggaatggttcaacggtggccacttaaactcgt
ttactctggaaaaatctgcgctttacaatcctttggagttgtcgacgcgtcgcctaaattcaataaaacgcgccgcccaacacaatcgtgtaatcaattgcgacagcgaaataaacatggattacatacgc
ggcaaacgtgtgcgaaccggtgcacattacggcaaacgtgttgtcaattttgatttgaatcaatcgttacactgagccaatcatggacaaacgtgccaattctcgcaaaccgtttctgttctacaacgaagat
tattattgtgagaagcccaaacgttattttcatacaaacaaagtaatttttgaaaaactcgattcgtacgcaacgaatattaatagatgcagagaattgcttacggattttttgattattgcctgccaaaatact
accgacgcaaaaacaaatttacgctgctgtttcgtcttttagaacccgtaatcaaacaggcaggagcgtcggcgttaacgacggtatcaaacgaccaatgccggtggctcgagatcaatcaattttccggt
tgggaacaacgagacaatcaatatgcgcacaaatggttaattaaagttgtgggcgcagacatgggtcagcaaattttatttattataaaacaagtaacaaagaagtttaaaacctgcaatttgggatttcata
attattacaaattgtttcgccgatgtttgtccatattattgtttaaacacaaagaggtattcgtacaatgcttgcaagtaattttaaaagccgcaatgcccgtcaagaaaaagggcgtcgtcaaatcaaactac
gcatttgccgtaacaaatgctttgcactattacattgtcgacaacccacatttgctgtgcaaagacataaacgtggctataaaagtcaggcgattgttaatcaagcatgaaatgctgccaaccgagaaaca
aattaaattatcgtttgaaaaatgttccaagggaattgaagttccgctatacgaaaagcttttgttaaatcatatgatgcgtataaatgacgacaatttacaatggccatcgttgataaataataaaaaaaattat
ggaatggaacgccaaccgctgttttgacgagagtaacaaaattttgcatgtttacattgggcaatattatacaagcagttgtcgccgcgtgaaaaaaagtttttcaaatacaacgggtggaacagacaagg
gcggttttgcagaaccgaaaaattttgtcatttagtcaacttgcaactgaacaaggacggatctaaaaaattaaaacgggttcaacgtaaactgataaattgtaaatagttgcaattaggtatcaaacgatgg
tgctcgtgtgcgcaatttttaagtgtcaaatatgtaaaatgcgaaacgctgcaggattgtttatgataatcgagccggacaaggctgtcctgctgtgtgcccgtcgcgcgtaccggagcgcgaacgcgccc
gccgccgacatgaacgacactttttagaaaaaatttccataccacgaggtcatcgcgattgttgcgacgcaaaagtttacgagacagccgtgcgcgagtttgtggaagaaactggccggtttttgacagc
gcgttcatctacaagttgccatttacgttacaatggaaagatgacggcgtcacctacaagtatttgatatacgtaggcgtcgtgcgcggcaacttgattaacgtgaacgccaaacccaacacgtacaccgt
gaagttgttgccgggcacgtttggcaatgactaccgtataatgttaaaaccgcgacgcttcaattgcgaaatagcgcgcagcctggccatcgtgccgctcaacaaatactttaattatatgaacgacaaac
aactgatcacgtacgattacagcaattacattgaatttttgattttgtgcgcagcgtcaaggcgcgttttgataataggcaattgcaggactttttctacgccactctgaaaaagatagacaacgatgccccc
caaaaattgcacgcacttaggcgggtgtgattcggactgcttgactcgcagcgaaatacaagcgctgttcagggaagccatcaacacgctcaagcacacgatgaacacagaaaacgtctgcgcgcac
atgttggacatcgtgtcgtttgagcgtataaaagaatatataagagctaatttaggccatttcacagtaatcaccgacaaatgttcgaagcgtaaggtgtgtcttcatcacaaacgaattgccaggttgttggg
cattaaaaaaatatatcatcaagaatacaaacgggttgtttcaaaggtttacaagaagcaaacatggtaaacgtgccggagcaacaatctccggagactgcggccgtgtgcaaaaatgaaaagctgttg
aataaactggaatcgagctcttacaacaaatccaacatggaccagctggccgttattgtaaatttcttggaaagaaagaacattaactatatccttaacgtggtgcctgtcatgcaagacgaacgcaaaat
gtccaaacgcaagaagaaagtgattaacaacaataaatacattttgtttaacagttggtacactaagatcaagcagcccgagtggcctagtagcccggccatgtgggatttggtgaaaaacaagcccga
attggcagattttgtgttcatttttgatcacactgaaaagctgggcaaaaaaaatggccgaccgatcgacatcgtcgtcttcaagcgagaacgcagcaattccagcgagcaaaaaaagacagactgtggtg
cttaccaacgcaaatttggcagaattgaaggagagctgcgaaatgcgcgataagctttattcggagttttacagcttgttaaacgagacatttaatcacaatgttgcgcctctgctaagtaacatttacgacg
aagtgttaacgcgcgatttcatcactaaaagtatggccaaatttaagacggtagcgctcaagttgcccgttgcaccgtctactacagagtacgtgcccactcccattagcggcagcaggaaaaggaaaag
ctcagttccagccaaacaacgatcttctataaagactcgccgcaacactgtggctccagcacttttaatggttagcgataatacacaagatacaaatatgtcagattaaataaaaaacttttatgtatattta

atggtttttattattgttattattaataccctcctcgtaatcgtaacaccatgtgaagagtagattccttctgaatattgtaatcggccatagttttggaatcttccagttgtttgcccgcaaagataagtcttgttga
tctacgggcacaccttctttatcggcaatttttgcttaagatcggccaccgtctctgcgggttccgtttcggcggtaatggtttgcccgtcaatgttttgatgaatatttgcatttttacactattacttataaatga
caacggttgctgtgaatgcgcccttgcctccaccgctggtggaattgtgcaacaggcggccgatacccacaccacgaattatttcccttcaacggcaattgatttccacgcccgtggtcaaaaattatcag
gccgacgtgcaagaagcgatagatgctttcaaacgacttaatataacaccgggccacttgggcgaggtgatcgatacgatgggccaacaaggcaagctgttgcccgagattatagaggctgacgacga
ttttaaggtgaatcaaacgcgcaatctcagctgcaaaactgtcgagtacttgaatttttggaaaacgataaattgtttcgctgtcgactttgttacacgcacgccgactggttgtggtgtgattttcaccgaaat
cacgcgtatcgcggcacacgcgacataacatgcaacaactacgtcgagcacttaaacagtgatatgggtgtagtcatgcttatcgaggaatatttttactgtctgtcgtcgtgtaatttttaaacaagatgcca
aacgtgcgctacaaacgcttacaaaatttgaatcgctatccgatctaatggcaagctacaatttttcaacgcccgacctggacacaaacgcgtacgaattgatggactttgagtaagatgcattcaccatgt
ggacgttacagcagcccgatttgtatgcgtacgcggttcatgacggcgctaaaagaaccaaaattgcggcttttgatttggacggcacgctaattagctcaaagacacggtccaaatttccaaaaaaccc
agacgactggcagctccttccgtgcgcacacaagctaaagcgtttgtacgaattaggctacgatttagttgtttttacgaaccaagcccatttaggttccggcaaaatcaaagcgagcgatttattatataaa
ttagagaatattaaaaaagctactggcgtgcccatatcgttttatgtgtcgcccaacaaggatgagcatcgaaagcccgacacgggaatgtggcgcgagatggccaaacagtttacgcatattgacaagg
aacaatctttttatgtgggcgacgcggctggccgaattaatctaaccaccggtcagaaagatttttagcgattcagaccgagtgtttgcaaaaaacttaagtctgcagttttacacgccagagcagtttataca
actcgaccttgatttataatgcgcaataccctaaaaagttatctaatcacacaaattattaactaatctcaatattatctaatcttcaaaaattatctaatcgccaaaaatggcccaaatttttatataaaagcg
ccagtttgctgtggggaattaaagttcatcatgtatcgcttgctggcacttgtagctctagcctctatggctgactgctcagctctgctgacacacatcacgggcacgtctattcccggccagctgtttattaac
cggcagtttttggcagtaaatccagacggagctgtttacggaaccattgaatcggacaacgtcgacaccatattcaagcgcgttgccgttgacagaaatcgcatcgtcattcaaaacgccatcacgtgtgt
gtacctgtgcatggaccggtgtggccaactgtacggatcaaaaactttatccaaggattgctttatgagggaatttctggaaaaaaaataactacaacacgtattataaagtgtacgatcgcaagttgacgtat
gtggcactgaagaacgacggtactccccgaaaactacagatttcaaagagccgcaagctgggcaagcttagcgtgtacgccatggctttattgaaacgcttaagctttcctatatacacgtcgtgtcctaa
tataaaaagtgaaattattgttcgccatcgtaaatgtcacgtgtaatcaatgtgtatttacaataaaaacatttttatttaaataattattttattttagtaaatgacatcatcccttgattgtgttttacacgtagaat
tctactcgtaaagcgagtttagttttgaaaaacaaatgacatcatttttaaaatgacatcatcccttgattgtgttttacaagtagaattctactcgtaaagcgagttcagttttgaaaacaaatgagtcattatttg
attgtgtttataatcgtgtgcaaaagatgacatcagctttgcacaaataaatgacatcatcccttgatcgtgcgttacaagtagaattctactcgtaaagcgagtttagttttgaaaaacaaatgagtcatacat
aaacacgtttataatcatgtgtaaaggatgacatcatccctcgattgtgttttacaagtagaattctacccgtaaagcgagtttagttttgaaaaacaaatgacatcatttgtataatgacatcatcccttgattg
tgttttacaagtagaattctactcgtaaagcgagtttagttttgaaaaacaaatgactcatccaatgacatcatcccttgatcatgcgttacaagtagaattttactcgtaaagcgagtttaattagagccgtgtg
caaaatatgacatcatctctcgattgtgttttacaagtagaattctactcgtaaagccagttcaattttaaaaaacaaatgacatcatcccttgatcatgcgttacaagtagaattctactcgtaaagccagttc
agttttgaaaaacaaatgacatcatttcttgagttcagttttgaaaaacaaatgacatcatctcttgattatgttttacacgtagaattctactcgtaaagcgagttagaacgatgacatcatttacatatggcaa
ttattccgcatcccaaacggccgccagaattgccggttgttttgctcaacggatgatcggtaaggcccaaatcgtctttgtccgtgtgcacgaccaaacttcttccgataatattatgcgggccatatagaga
cataacgttgtccatcatgtttacttcggtcagtgaattgtagccagccgattttatgttgcccaagtcgccaacatgcctaatttcagcatcgggagcgccgtggtcctcattggtgggattaaagtgctcacc
ggccgacgtgcaaccgttgctcgtgtcgccatattcgtgcacgtgaaagccgtgcaaacctcgaggcaaatttaacaaatagccgctaattttaagcggttgattcgctgattcttgttgaaaataaatttttc
catgaacatcgccgctaatgatgcagatggctttcatgttgtttcttaattaattacgccatttttgatgcaaagcaaacatgtaatcgacgtttgaacatgccgaccaaaacggtcaccttgcttcaaacagcg
tggagcttcagtcgacatggctccaacccaagtgctacaacgcgctttaaagcaaaaagaatatataaaaataattgaattggccgtcaaatctcccaacaatcgtgtctatttgcttaatttgcaagatga
cacgttttggaagcgcatatcgcgcgaatgttacggacgcgtcgatttttattcacgtgtttcgcaacaagttggattggaaaattatttcaatttcgccgttgtcgataacaatcgccaacaggtttaaatcaca
tttgatctggtcgttggtgtctgagcaaaaatttttgacgcaagatttttatactcgcgtttggggacttgttggacatggaagaaatatccaaaaattacaataacctctcgttgtcggtgcagcaaaaatacgc
acacaaattaaattggaaacgcatagtggccagtcacatactactaaaagaatggtttcaagaacccattaaacagtacatcaactacgattatgtgtccaaatacaaacacttgaacacggcgctaata
aacaatgtgagctgtatggaaaacgttaatttgtccgcgtacatgcaagactgtgtcaagatcagcgacgcgctcattttgtattgcttacgcgaaggccgcgtgcaagaactaaaattgatcgcccacag
cataccgtggtcggatcacatgctcgtgtttgacgaatatcccggtttggtgaacacgttacattcggactggaagtgtattgacaaatggactgcgttcaacgccccgcccgcgtactacattagacacat
gtttgcgcatccagagtttagaaatgagtttgaagaaaatttcaatgcccagtattggcacaagctgatcaattacacagtgattacaaacgtgcgcgccagcgcggcgttcaacctgatgctgtttcaaaat
tacaaaaatcgcgtcgattggacagagctgcaacaatgcaatcagtttttgaatttaggcgttttgtacagagcgccgtttccccgcgtcgacttgcaagcagtgccgcggttcgacgacgaaaaactttgg
aaagcatacggacgccatttaaagcttaattccgacgaatgcgatgacccgcaagttattccgttgcacatgaatgtcaaaaccgcatatcacaagatgatactggtcgaaccgtgcgccactcaacag
gaacacgatagcgtagaaacgtgtagcgccattttcaagttaaacgggggcgaagagggattgttgaactggaatcttttatcggccacgcagcccatttgccccttttaacctacgacatttgcaaaatgtc
aatgccaacacttatcgcaataaaaataaccattttatggaagatgtgtacgagcaaatggtcgcgatacaaatgaataataattaactcctaatttggataaaagatgctgttttacctcatctgtatgttttct
caagatgttctcaaggaattacaacgcatctcaatctcagcgagatttaaaatctcaattggaagaaattaatagacaaaagcaaaaaataactattgacagccagcattttgagaaaattaaaagcgtc
acaaaaaacgtaaacgagctgcagaatatggagaagagagttatgaagtcaagacaaaactttctcaactacggaatagacaatttttaaataaaatacattatatattacaaaatatagttttattattatt
atttttaatacatgttttattgtttttctaatacattcaagtcgtctagatgattatttaaactttttaaagagtcacgttcgtcgatttcgccatcttctaaattgtcgtcgtagtcgtcgtagccgtggctctgggcgta
cagctttagttcttcattccagtcgctgtgtggtatgggcgacgatgccaaaacatatccgtccaagcgaggagaatctaaagagcggcgcggtcgtttgtcgaccacgtggctttttaatcgatgcgagctg
tcggtccgtttgactgttgcgctgtgcgtgcgaatagatttgttgtacggatgagagacaaactttgtgccatcgggccagtagtatctgccgttgcagtgcaacacctgcctgtaggcgcgcttgtctagaac
gcgcagctgtttgcgcgtgctgccgttccagtcgacgctgtttttcaaatcaggcgcaatatgaattaaaaactccttggagaattttttttccacataatagccgttgtagtacacgttgaacaccattttatta
caccacaaatatttttataaaatcggtgacgtcctgacgacacgtcggacacttttgatctaatccaaaataacactgaacgcaaacgcaaaaatgccgacacggcataagcacggcgtcgcgttggcg
ttcgagacagactttgcattcatatttttcctcaatgtcgtcggcgttttcgtttaaattgtcgtcgtcacgtttatcaacgtgagtgacagtaattgagttttgacaaaattctttacctttcaccgacaatacaaaa
taacattgcgggttttcggcggcgtgtcgctgccaggtgtcttcattagtatgccaatcacgtacgcaacagtcgcaaaagaaacacacagtttcgtctccgcgacccgtgtaaaaaagtcccgcttccgc
aatgtttgtaatcatgtcacgcaatgcggcaggccaaaagttaacaaacgtatccatacgcgactgtaaattggacatgcatctgtacacacacttgggtttgccttctttcactagtacagcgttgatggtaa
tgttgtcgccaaacgattcacgctcggcgatcttgttagcatacgcgcaatacggcgacaaggttacgtgtgcatattcaatacactcgtcttcggaccaatttttttatttctgcttcgcaatactcgcacacaa
cgtgatcgtcaacttgattgtatttaaacccgttaacgatcaagctgttaataaacgccgtgttttcaatgggataattttcaaacgaactatgtctttctattaacatgtcgaatacgtgttcggcggtgttgtcg
cgaaagttgtcacacacgctgataaaataaaacgggggcgtgtcctcgttcattttagctcgttaaagttacggtcaaaatgagcacgtttgcgtcgttttggtttagcgacacgtttatatggcccagtttggtt
tttgtttcggcgttaatgacgtgcactgtggacaaatcgtgttctaaaactacaaactcgtactcgaaaatgtttgatatgtagttggttagccgatctatcttaaaattaaacttttgcaactcgctgatagagca
cacgtccacatacttgtcgataaacccgttgctcaaccgcttcaaaacggtgtaattttgtagcttgaaaggggcgcatttggaatgactaaaaggaatatttttcaataaatcgtcagtagtgtacgcaaac
gcgttgtctacgcacatgctggcaacagagtcgtccatatttattatatatcttatattctgtgaaacacttcaattagacttgaaccacagcagacagcgcacgtcggtagcatggcaactaaacgcaaga
ttggcgatggatattcgagcagcgacgacaaccagccgaagcgggaaagatcggaaggtggtgaggaccagcagctggtgccgtacaacagcggcgcttttaacgtgaaacacgacgaaacgggcg
taatgtgctattttacgccgtcgtccattcaattggagccacacgagttgacgaagatgttgtggcaggagcagatggcgattaatgtgaaacgaggcaacttttcaatactgaattgtagttgtttcgaaggc
agatttttgaaaaacgaattttgtcgactgtctaatttgaacagtttgcacgaatgggaagacaagttgtatcccgagcccgataaaaatatcgtggtgttggagccggccaacggcaagactacgtacacc
atcggtccaagagtgcagggcaagccctgcggattttggttttccgactttggaaccataaaacgtgccaaaagcaactttggccagtttttttagcattcaatacggagacattcacaaacacaacaatata
tttggcaacattttgcaaaaacacttgcaaagcgattttcctttgaagatggaaccaaacgtgtgcattcatttgccggataagaacaaaactagcgagcgcgacatgttgattcgtcggttttacataatcaa
ccgtgacaacaacggatcaatttacgccaccggaaagatcagaaacgttcctcttgatatacagagaatgagcgttgaagattttgatagattgttcgaaatggacaaaatagacggaccttcggaggag
attaaaatgtacatgatgggcaccattgacggcgtcaagtacggcaaagagatgcagatgacggacatgaataataagaaaattaccgaaaagccctattcgttagcttttaagcctggaatatttgttatt
attgaacaataaatgaatatataaaagttttgtttttttattataagtcaaaatgtcttgcatattaactgcgctttgtaagaaaggtcaagcgaatttaaatagtttgattaaattgcaaaataaaaaagttaaaaa
ttattatgttaaaaacaatgaaactgcaatcgacaaaatgttgtgcatagcggccgatatcaaaggccaagtggagcagctcgaactcgtaaaccagtatctgggcgcgcccgaaagtgaaaagctaga
ttttgtatacgattgttccgacttggatataaatgaaaaagatttaaaatctctgtgtttgactaaaaatattgcgtactttacccaaaaatacaatgcgccgaccgtgctgaaggcccaggcagcggtttacg
attcgttcataaaacacagcgaattatttataaacgccatatgtcaaatggatgaaaaacagcaagtgaataatttttgtttagacgaattagtaaaattaaaactgatagccattaaacatttgtgtgcattgg
aatatgtaatagaaaatagtatataaataaaaaacattttattaaataatatatttatttcaattatacatgttttattttattctttctataatcatagggtacaacacaggtttatcatcttttaacaaatcatcatta
aagccaagtaaagggtacaatggttttggcatatcatcgctatcgtattcattaggtaaattaatgtatgctcgattgtcttctaaattattgcaaattgtgacaaattcattgttatttttgtattgcacaaccacgt

tatttgattgatgacaccgttttttatacagcaaaacaaaaataacaataatccacagaataggatcattgcagcgacaatagcgataattaccacggtggtattactgtcgtcttgcgagtcttcgtcgggc
gccgcgctgttggtgtaaattttagactgatggaaaaaatgatggggcgtcgatgtgtaatctttaatgtgaaatttatgaggaagttgcaataaaacgggccgggtaaaattgtcaacagtcggtataaaact
ggtggtcgctttttcaagcaacatttggtcgatgagtagttgttaaaattaaatttcttttaggccagtatgtagttgaatggctgacataaaatcgcgaatctacagttatggtcacatcatacgttattctaatcga
acacgtttgcgccatagtcgaaaacataacaccaacgccgggcgaaacttttgtatcaacttctacaagcgcgtcgtgacaataatagtggacggacacggtcgtgtttggaaaaatgtacaaccattttc
ggtagtctattaaagtgttcacgtaagctttttgggattataattggccaccctaatgtcgcacatgttgtccacgtcgtcgctatatcgacccataaacatttcaatctcgcacactttcgagttcatagtggc
aatcggctcagtttcgggacacaaaaactcattgtagcccgtaaactcttggcaatcttgtctaaagttatcagatacgggcacataatagttttgctgtctgtaatcccaaagtattcattgcccgaaatgat
aagcagacacattttgcccctgcaaaaaggcaccgtcatgacgcgatacaaattaaacgaaacggcggtggattttaacaatggcatggtaataaacataatgaccgtgtttaaatcaacatataaatgc
aatttatacgcttgcgacaaatcgaaataactattcatgtccttgttaaaatccactacccacacgcggttttggttgtataaagtttgtgttaattttttcatttccgttagcagcgtgtcgtcgtctataaccgtgt
tggagatgttgagtttgttctttcgcgtttgttttactattcgctccagcatcttggccaaagctgctgcgtcgtccaatttgtcgttgagcacgttcatcatgttggctaaaaacatgcactcgttttcaacggtacg
gttgctcacaatgcagtttatgtattcttcgtatttgctgctcacgtcgctggtggtaaccgtgtccaaaaacattactctgtctttgggcgggtttcggagcaaagcttgcgcgtcggatacagtcatgttgctcc
aatgcgaaggctgttcgtattcgtcgtaaacgtcaaactcaccgctgtcgtcatattccagatcgatgttgtccagcacgtcgttgcttaaaatgctgctttccacagagtatgtgatgagatttttgtttagttgc
gtcaaaaactgcagctgattttcaatgcgagctttgatgatactcgatctgttctgttttatttgtctgacgtttgtgcaagttgaaacttgcgcttcatataattcaattttagatttgatattattgagctcttgcagc
accgaaacaaaattgaatttttcacaaacacaaatctgtcgtctgtatgatgtcgcaagccgtacatgcgttcaaacatgaggcctgcgctatcttctatcggcgtaaacgtgatacggtcacgcccgaca
actgtactcgtcgtcgataacaaaattattagcgcgcacgtccctagcagctttacaccgtccatgataaaagcctgatattgactgaatttgcaagccaacatgtacggcttatatactaactactgacgta
acatttaaatgacttattgtttgtgttatcacaattaagataactgccctctgacgtcattgtaattaactcgagatctggtatcaagataaaagcgataagttggctcgttgatgccgttgcgcgataaagggta
ggagccgggcacttcaaaggcaaaagctgcgtccgtgttactatcaaaagtgtcactgggacataacaatttgttattagaataatttagaatagaggtgtacagattaacgcactctacgcgatatctatat
gatgccgtacttttatccaggccatctataatgctggcacacatagaggtcctgtcccccggcacaatgtgcgtaacacgcgtgaccgcttcgtcgccgcagtcacactctcccgtttcaaatacgggtcta
acactggtgtgcacgttgctaacgttagtgcacacgttcgggagacactcgaggggattaagcggattaacaaacatgagattattgttgttatctcgggcgttgcagcgcatttcgaaccgcctagtgccgt
cctccagcaactcgtcccagctgcggcgaaacgtgttagtggtcacgttcacctctcggcctagtaatcggtcaaacaggacgttcctatcactctgtccgggcataatgcggtcaaagtgttgtctgcctg
cgagttgcgtcatgttatctgtgcccgcatagtaacgcgggtcctcggctatgcacgtccactgattgagactcattaaaatgacgctcgtctcgcgattacagttgcggggcaaactattagtggagcagta
tccgcccaccgtcagtctgtggccgttgaccacaaacacatccgtttcgttgacaaagaaatacgccgcagacgggttatcgcatatctgactacaatcaaacatcggcaggtcaatgtcactaacaaa
caaaggcaccgacatgcatttaatcgtgtcaccggccaacgtgccgagattagtatcaaacctgattgcgggcaacgcttcaatcggcacgtagtgccttctacgcattacggattcaatgtaatccatcc
tatcgtccaacgtgtcattgtaatcttgttgggcatgtccgatatgcaagtatgcctgataaaagggctggtacactatataaagaaacacgaacaggaaaaacacgatcagtactcgatacatattgccat
tacttagttgttgaattaaatattcaaaatgttaaataaaatcactgcagttttacaattcgggcttaatgcctcattattattagcatatctgatagtgtttgtgttatctataatgggagtgatagacagcagatac
gcgtttttgctggaaattgaaggaaagcaatcggttatcaatctatctataccgataatgctggcgttcggcatgtggatattattttacatgttttattttatatggaagattatcgtgtggacgaaaaatcgcag
cagctcttcaaacacaaatgtaaatttcaatgcgaaagaaaacttttacgtcgctatcacgtgtattatggtaaacgtgatcactgggctgtgttggatgctgtttgccgcgtttcaaatctacgtgttcaaaaa
tggtcatctaccggcgctagatgttttatatcgtcactacgatttggaatcaatgtgctggaactctatcgtgtatctggaaattgactacgcaaataccgaaaccttgagccaaaactgtgtgtatgtgaacat
ttacaaaaaatgcatcatgtgccgagcgattgtaagcgatcacgagccgaccgtgtttaatcaaaattatcccgtaatcataatgggcgtgcttacgatactggctgtgcaatgctggaatttgtacgtgcag
ttaaaggagatgcgtcacaacatatacatgaaaaggcgcgccgaagccgaaaaggcatcttatgaacattattgcgacattgactatcgccgcgaagaacgagagagtaactctagattgctggaagtg
gtctccgaaggcagaaatagcagcagcgtggttgccgtaacgcacccgccttctaccacttctaccaccacttcggtttcagaaacactttcatcttttatcgcgccttcggacttgtcttcacagccatcgc
cgtcttcacacccatcgtctccgttcggcaaccacaacgagttctattttggcaacacatatagcgtgcccaagccggtgtacagcgtgcccaagaaagttttgtgtacctcaccggacgcgtgcacaatg
tgcgagcctccgttgccggatgaagaaatcattagtcggcgcacgccgacgctttcgcccaagctgttccgcaagtccagagaactgtcgcccatcaaacccgtaaggactcctactccacccgctccc
actccaccgcccatgtgttttagcgaagagttgcagcgtaaatttcaagagaaaaaactattacccgtttatgatgattagaaagtattatataaagtattatttttgattatataaagtattaatttagtattagc
gtaatttagtattaatttagtactaatgtaaataaagtattatagtgtaaataaagtattgtaaataattgtattaataaataccaaagtattaatgtataaataaaattgtattattttataaaaaataaaacaata
attagaactgaacaagtttatttatttatttttaataatatttttccaacaaaaaacttaacacgttggttaacattttatgttgaatcgtaaccccgtcgctgttagcaatgcagaaattaacaaatccatttagca
ataaacttttaccaaacatgttgagttcaaagttgatcaacgagttgatcaaattgcacgcggtcattaaatttaaattatttttaaaaagcaatgtgcttacaaacgcgatggcgtactcgtaaaaactactgt
gattacgattggcaaataaagttgcaagggtttgttttgcgtgcgcatttgaacagtgtagtcgtgttgcgtcgccgctattcattatggaacgtttgttgaatcaactaaatcttggcgtcctaccttatattacga
caaaggatattgaagatcgcctgcgcgataaaattgtggctaaagcaaagttggcatttatcaaagattgtttcgaagccgtagtttgcgaaaacggcggtttattcgtgttaactggaggcgcagccgtgac
atgccatattgacgacgacagaagcgcattaaaatgcatagattttgactattacggcttttgcgcaaaaatgttttgcaacctgcaaactaatttacaaaaatgcgtcgaccaacattacgctgaattggac
gtgctaacgcgtcaagtttacatgtcggatccgttggtagtgctgaaatgttatcaaaacggagcctatagattaaacggccaaatcaacttgcacttaaaccgccatataaaatgcataaagacacagta
caacgatgagtttgatttggttagatttgcgctacaaattgacataacaagcgcagacggcgtcgacgagtataccgacaacggcgttaaaataaccacagccccgctgtctttcaacgtgtttttgtcaac
gtgcgcattatgaaacgacccttcaacgcggaccgctgcatcaaaaatttttcgctttttggaaacgaatatcatgtgttggtgtcgtcgttgcagcgcgttttgaacgaccagttaatgtgcttgctgaaagac
attttcaccaacaagtttgattataaaattgaacgcagactcaagcatctcaagcgcttgtttgccaacctgccggcggaatcgtacaactcgtgcgtcaacgatcacaccgacatgtgtttgtacaaggaa
caaaacgaaaccattaccaattttgtgaaaaaaatacttgacataagcggccccgctttgggttgtcgtaaattaatgcacatctacttgaccaccgacacgtttagcggacaactgcccgcctatttaact
cattacgtaaactatccacacaagagtttgtgtgaccaaaattggaagcggtttatgtcatgtattttttcgttatactaataaaacaacacaattttctaacaatacatttatttaaatttaaaaataaataaaat
agttaatagctgtctacccgtaaaataatttaacgcaggaacccatttagtaacattattttttttattaatattataatttttatctacctttataaattttactacattatttactaaaacagattcattgttattattat
aaattaaacctttaagcatattccatttgttatacacattagcctcgttaataaatacttttaacattgtctgttggttttgaattgaatccaccacgtccttcactttacggtcggcgacagcagcgtcgacagaa
tataatggtccaatgacataaggcattgttgtaccttcatcgtccggccttacaaacccgtcaaacgttacatgcttttcgttagatttgttaacaattgttacactatcgctaatgctacaaaaggtagcctga
atataagcgctatcaaagccatctacaataatcaatactttgaatttggcaaaaggagaacggcatctgacatttaaacagatctcagtatctattaaaatacatggcattatggtaatgcgctcaacgcaat
tgggctcttcctcacctttcaaattaataggcgttaatatcactttgagattcatcttgcaaatgtttaattattaaacgtttaccactagcaatctcgtccggatttataaactcgaaccggcgactattttcataa
actatgactcgcgaaccaaacgccgccgtcagcagactcgctagtgcctgtcgatctttgtctccatcggcgcgtgcttggcaaaacacgtactcggcgttgtacaatgtgcacagctcggccaaacgtct
ccttacaaacttttcgttgccaataattgtgcgaattgtgttgtcacatctgactactccaaagtatttccaatgagctggtttactaatatttctgtgcgttgtcttcttctgtgcactttcaagttttctaaccaactc
ggcaataatttcgttctttttcttagcaaattttgcaaatgccgcagtttctttttatgtacagactgtatcttattgaaattcaagtgcgtgttacgcaacgaagatattatgtggcatattttctgacgtttttcttga
taattttttggaatcacgctggtgacggtggtggaggtgatggtgtgtgtcaaactgttcacagaaccgaccggcacggcaagctggcgctttttgaacggagcaaattgattgctactggcacataagcgc
gtttgaacagactccatttcgatgtttcgcctttgaacgtcctggcacaataatgacctgctttggttggtgagcgtttataaaggctggaagtgcgcagtcgcacaaagcccctctcaaaaaaaccaggtaa
ataatggtgaataacatttattgacgtaattaaaagtgtgttaagtgattcattacaatagaaaacacagatttatacatgtagtaatttgaatacggtacattggcagctgttgtttttaatgaatacaattgacg
actgatgtcatgtctatatcgaattacgcgttctgtgaatttgatggctttatctagaggcggaaacgttgctttatcatgattgatgagcgcatataactcggccatgtgtttttgtaggtaggcgcgttaaacaa
cacgtcgtttatcgccactagtagttgatcgctggaaacggtaacagtgtccaaggcgcgggctacgccgagttgctgtaatttgtgcgcatggtaaaactggtcgcccatcatgggcagacacaccatgg
gtatcccggcttccaaggcctcgtcgctcgattgtagtccgccttgcgtaataaacgccgccatttttttatgacgcagcacggcgcgttgattaaaccaattttgcgtgattacgttggcgggcaacgttatgt
tttttactacttcgtcgtcaattttccataatatggtgtaattatccaacgttttgaacgtattgattaacatgtaaagaaactcgtttgcaaacgatttggtgtcaatgctcgacccaaaacttacgtaaatcgttc
cgctttttgacttgttcatttgcgcgttgatgaccggacttaatttggtcaacggcgcgctctttacaagatggattcctccgccaagatactgcacgctgggcggcacgggtcggttgttgtcaaatatgggat
gcaggtttagcaaaagcaattgcaccttgttgcgtagtttttcaattgtcggtgtgttgggtccaaactgttgtttgagcaacgcgttggacatgttggccaaaattttaaattctttatacaaacgcatttccgtca
tcacgtttgcctccgtgtcgtcgaaattgctgcgccaaatgttaggatggtggacggggtgccgcgccacggcgccgaccgtgtcaaagttttccgccaaaccgtagccaggcgcgatttgaattacgggc
gccggatcgtacaagtgaccaaacaccaacgcataatcggcaaacgcttccacgacgactaaatcaaacgtctggttgttggcaatgagattgcgcacgttgatattgtcaaactggtctttgaacatttc
aatcaagcctaggtagttagcggcggttaccgtgtctgtatcggacaccactccgcgctttctaaacattgccgaattcgccactagtttttttgtattgctcaacagacatgtcggcattaatttccgtgatatta
ccgcaataagttttagttgaatacgcaaacagtttgggcttgacgaccgtaacgttgtgacatttttcggcaagggcttcaatatacactttgtacactatatggtggctgtaagctggcgtaggaaacacggc
caatatattggccgcatttacagcagtaagcgtagacagcagtgcaagccagcagagaatagtcattttgcttcaaaccgaataactgtactggtaatttgattatgtaacaaattttaaagggtgaccttttt

atagccaaaaattaacacaatcgaggtcaagttcaaagggcaacatgttagtgtgcaattatacgcagaagcgcgtcgacatgatgtgggacgccattgcgtacaacgacagccgcaagtacgcgttca
tgacggtcaacgcgcgctggattcacgccgacagatattttgatacctccgcacaattgtatagttatattgtgcaaaacaaagtgtccgacgtgcacgtcaaaccgttggacgacggcggcggcaggga
atgggtcgtagacgccgattacaaaaattatgttgacgaacacgatttaatgctgaaaatttacattggcgccaccggcgtttctgttgttttacacggaagagaacgtgtcaagagtcatgtataccggcaac
cgtggatttcatttgtggttaaaattcaccgacaagtttaaaatcacgtccgctcaaaatgttcgcgtgcatcggtacaaggctttcgagaaacctgcaaagttggacagtgattacattcagccgggcagttt
tgcgcattgtgtgagagaggccgtacgtttgtacgttccgcatatgcaagattcaaacttggacgcgctcacgttgcagtattggccggacgtggacagggatattttttgtaacgttaacaaacaaatacgc
gcaccgtacagctataattataagggaacgaaatttctcgttgtataacaaaggaattgctagacaagctaaagcaatgctatcctggttatggaactggtggatgtggtccggtgacaacgacaacgac
gccatcgccgccgaagatcggttcaatgcagacgactacaaaaagtaccacataaacgcccaacaatggtcgcacatcgttaaatgggactcattcaaatgcaacacgcacagtttcaagtacagata
cgtgcacagcgacacgaacgcaaaatgctacaatgtgatagatttttgcaaaggtcttgaaattgcgcacgacgacatacttgattgcaattgggacggcgatcaagtttaccatttaaacgaaattgtttt
cacaaacagagatccaaacgcgatctcaactcgttgggcgcattgttcgcgaccaagcacgggttgttggaaattttgatgcggttaaattttgccaacaaaagcaacgcgttgctgcacattcaaactga
aggcgagcgtgatgatttgcgcgacaagattgaatctgttttaaaacatgtaaagaaactgaatgcaaacagcgaaaaatttatggtcacccacgagacgttcaagaacgaggtgggcaacaggtttgag
cagtttgaattgcgtttgcacgaactcgacgccaaacttaacatgctgcagtcggccgaaaaattgaaaaccgccgtcgtggcggaaagcaaaaatggcacggtgacgtttccgcgcgacatcaccaa
acaccagcatttggccgtattttcggaacgcatcgacgaccgcatcaaactcgcttttgttttgggccaagagcggcattttcgcaagcgaaaaatgcgctttgaagacgacatggaagtactctatgacgg
tgtgcacccaaatcccttgttggctattcaatgtattaacgaaaaactctacgataaacattacaaaattagaaaaatagctaaacgtgtaattgacgtggattgtactcctaatgtagttagagaggttatac
aggaagtattgtaaaataaaaacatgtataatagttattattatgtatttatttcgtataaataatatatataaagtttgattgtacattttgtttttctaaatatttacacaacgaaacaataaatgaattacgaggaca
ttggttttttgtttgcatattaaacaatcgcgcgttggctcaaatgggtaaaatccgttgcaatcttcactttccacaattgctatattatacaaatgatatttcatttctacttgtttatttgttacactaaaaattgtc
gctctctggtgaccgcaaatactcgcaatccatgatgatttggcatagaattcgtcttgtctcagttcgttgcaaatacagttgatgtattcgttgtagcttgttatttgcacgtctgtctttttgtacacgcctttcaa
atgttggtgattagtactatcaaatatcacgttatatttactagaaggtcttttacagacaagcaccaaattgcagtaatcagataaagataaatactcgacaatcttgtcaaataattcttgaggcagagaatt
taatttacaatgacgacgcttcatggttgaggctgccaagcttgtgataaaaacagacggccattctttaataatagcaaccgataacagctcccctccttgcgaaattatatataaatagcacgcgtgcct
gtacatcagtattattgagcctagcaccgctttactgacaacatgtctctcgctgcaaagttaataatttacaattattatgctaaatacaatgaggtgcacgacgtgtatggcgagtcgtatcaccattaccgt
atagttcaggaatacctatcagaatcgtacgtaaacggtatgtcgtgcatagagcgggacgtgaccgctatgcggcgtttgaagagcggcagttgcacctttgacgaagcggtcaagatgatagacgcgg
gcgattccatcaagagcctttcccattggttcagcaccagcgaaacgatgggcatcgacgataacgtccgtgaagtgttggagcaaatcgatgcggtggttcccgtcagcgtgcgcgttcaaaacggatg
gcagatattctctttgaacaattttgaacgagaaatttctcaagacatgttggattgcttgcaaattattttgggccgctttgaatatttcatgagaaacgggaaactgctgcgcatcgccaacgtgttcaaccct
aataacgacgttgtcggatggtggtacaataaattttgtgtggtaacttatgtgcacagaataatgtatcgaagcgtgcccgccgaattggtgcctcgtttgtcggaagcggtgaaaaaatttattcgtttgcgc
aagagtgattatgatgatcgtttacacgtagacgaatcgtacaattgtccgcgtgtaatcgccgaaatgtatggccggttttgcggtataggcaaggagcattttagtaaacacaagttgtcgtgcatgcacat
tttatttcaatatctgcgcggcaaaaccacgcaagaagaaaagtcttttccttgttatagggttattaaagattttggtagacaatgtaaagatgtgtataaaaatttaaaagatgtatttgatttgttacatgcac
atagtatgtcggataaagacaaaaactcattgatggacctgttgtgtgttatggattgtgaagaaatagatgtagattgtttttattatatatttgaatcgtttttaaataataaacatttacaataatttttatgatttt
attaacaccactcacgatcatgctttacacgtagaattctactcgtaaaacaagtctgttttaaaacaaatgagtcattgataaaaaaatgacatcatcttttagatcgtgtgttacacgtagaattctactcgta
aaacaagtggcatttttacgacaaaaactcatgttttataatttgtttgtaatttgtgagtctacaatctatgttgtatcttgttaggcagtacacaaagtcacgagcgttaacgttacgaacgtgttttaaaacgc
caatgtcattgtattgctgacgacgcttcatgctattcaagtccaacatttcgtcttcgctttttttcaaatgggtgtcggccgccggttagcaacttgtatgttaacacgccgacggcgtaccagtcaaacgaac
gtgcatagttgtggtgtcgaatttttccggactaaaatactccaacgtgccgtcgtgcacgctaagtgagttttcgtgtttgcacaatccgtaatcgcaaacatacacgcgatcaagtgcttcgaaatataag
acattttcgagttttatgtcgttgtgtatgaaattgtgcttgtgcaaatcgttgagcgcttcacacagctgtctaataatattgctaacaagttggtacgaaagctcgcctttaatttgtagtgtttcaaataaatcgg
gacagtcaatgtagtccatcacgatcacttggttgttgatggaaccgtgattaaaatagatttttataaaatttgggtggtcgctcattaactggtgaacttttatttcgtcagcgttaaaattgtgcgccgaaatt
gttttctgcaaatacaatttgctagtgggcttgtgcgataaaatagatattttgccaaagcgcccgttgattattttataccgagttttgacgtttttgcagttttcgtaaaattgtaccagcgtttgcagcgtagcat
ttgtggtggccatgacgaatcgtagatatgaatctgtacaatcttatttattcaataataggaacaacaaaattgatgcgcatcaatttttttgaacgtgtcgatactgcggaagcgcaaattatcaaaaacaac
atttacgacaacacagtggtgctgaacagagatgttcttttaaatattttaaagcttgccaacgacgtttttgacaacaaagcatacatgtacgtcgacgacagcgaagtgtctcgccactacaatgctgtag
taaaaatgaaaaggcttgttatcggcgtgcgggatccgagcttgagacaatctctttacaatacaattgcttacattgagcggttgttaaatatcggcacggtaaacgatagcgaaattacaatgcttatagc
agatttttacgatttgtattccaattataatattgaattgccaccaccccaagcgctgcctcgaagcagacgaccttccgttgtgcagccagcggcgcctgcgccggtgcccacaatcgtgcgcgaacaaa
ctaaaccagaacaaattataccggcggcaccgccgccaccaccttctcccgtgcctaacattccagcgcctccaccaccaccaccaccatcgatgtctgaattgccgcccgctccaccaatgccgac
ggaacctcaacccgctgcacctttagacgacagacaacaattgttggaagctattagaaacgaaaaaaatcgcactcgtctcagaccggtcaaaccaaaaacggcgcccgaaaccagtacaatagtt
gaggtgccgactgtgttgcctaaagagacatttgagcctaaaccgccgtctgcatcaccgccaccacctccgcctccgcctccgccgccagccccgcctgcgcctccaccgatggtagatttatcatca
gctccaccaccgccgccattagtagatttgccgtctgaaatgttaccaccgcctgcaccatcgctttctaacgtgttgtctgaattaaaatcgggcacagttagattgaaacccgcccaaaaacgcccgca
atcagaaataattccaaaaagctcaactacaaatttgatcgcggacgtgttagccgacacaattaataggcgtcgtgtggctatggcaaaatcgtcttcggaagcaacttctaacgacgagggttgggacg
acgacgataatcggcctaataaagctaacacgcccgatgttaaatatgtccaagctttatttaacgtgtttacgtcgagtcaattgtacactaacgacagtgatgaaagaaatacaaaagcgcataatatttt
gaacgacgtcgaacctttattacaaaacaaaacacaaacgaatatcgacaaagctagattgctgctacaagatttggcaagttttgtggcgttgagcgaaaatccattagatagtccagccatcggttcgg
aaaaacaacccttgtttgaaactaatcgaaacctattttacaaatctattgaggatttaatatttaaattcagatataaagacgctgaaaatcatttgattttcgctctaacataccaccctaaagattataaatt
taatgaattattaaaatacgtacaacaattgtctgtaaatcaacaacgcacagaatctagcgcttaataaatgtactaataacaatgtatcgtgttttaatacgccggaccagtgaacagaggtgcgtctggt
gcaaactcctttactttgaacaccagggaaacttcaaggagaatttcctcctcttcagcagagtcggtaccgatgtaaacgatgggcttgtagaagttctcccagatgacacgatcgatgaactgttcgaac
gagttggtgtactcagagtgaaggttcattattgggcagccgccgcccttcttagccaggctgatgcggtactcgttgttgctgcccacccatgaaggctcgacgatcctaatcacgtcatgaggtacatagt
cggggtcgcaacgcagagcgtgttgggccaggaatttgtaacaacggttgggtctagtgggacgcatgttgacaacaaggaaaacatccatcacttcttggtcgttaacaatggggaagctgtcttccatg
aagcgggtccaagtttccctgtagaactcttttcctttccatccaacgacaagcttcatcgtgtcgggtttaacattacggatttccttgaagagagtgagtttttggttcttgccgggtcccaggaaaggatca
gatccatacgaaacgggaatgcggtaattacgctttgtttttataagtcagattttaatttttattggttaacataacgaaaggtaaaatacataaggcttactaaaagccagataacagtatgcgtatttgcgc
gctgatttttgcggtataagaatatatactgatatgtatacccgaagtatgtcaaaaagaggtgtgctatgaagcagcgtattacagtgacagttgacagcgacagctatcagttgctcaaggcatatgatgtc
aatatctccggtctggtaagcacaaccatgcagaatgaagcccgtcgtctgcgtgccgaacgctggaaagcggaaaatcaggaagggatggctgaggtcgcccggtttattgaaatgaacggctcttttg
ctgacgagaacagggactggtgaaatgcagtttaaggtttacacctataaaagagagagccgttatcgtctgtttgtggatgtacagagtgatattattgacacgcccgggcgacggatggtgatccccctg
gccagtgcacgtctgctgtcagataaagtctcccgtgaactttacccggtggtgcatatcggggatgaaagctggcgcatgatgaccaccgatatggccagtgtgccggtctccgttatcggggaagaagt
ggctgatctcagccaccgcgaaaatgacatcaaaaacgccattaacctgatgttctggggaatataaatgtcaggctccgttatacacagccagtctgcagtcatggtaccggcagtgtattctgccgggc
agcctgcatcgctgcctgttgccattgattatccggcagctctggcacttcgccagatgtcgatggtccatgatgaactgccaaaatatctgctggctccggaagtgagtgccctgctccattacgtcccgga
tctgcaccgcaagatgctgctggccacactgtggaacaccggagcacgcattaatgaagcactggcgctgacgcgggggatttttcgcttgcgcctccgtatccgtttgtgcagcttgcgaccctgaaac
aacggaccgaaaaagctgccaggacggcggggagaatgcccgccggtcagcagactcaccggctggttccgctctctgactcctggtacgtcagccagctgcagacgatggtggcaacactgaaaat
acccatggagcggcgtaaccgtcgcacaggaaggacagagaaagcgcggatctgggaagtgacggacagaacggtcaggacctggattggggaggcggttgccgccgctgctgctgacggtgtgac
gttctctgttccggtcacaccacatacgttccgccattcctatgcgatgcacatgctgtatgccggtataccgctgaaagttctgcaaagcctgatgggacataagtccatcagttcaacggaagtctacac
gaaggtttttgcgctggatgtggctgcccggcaccgggtgcagtttgcgatgccggagtctgatgcggttgcgatgctgaaacaattatcctgagaataaatgccttggcctttatatggaaatgtggaactga
gtggatatgctgtttttgtctgttaaacagagaagctggctgttatccactgagaagcgaacgaaacagtcgggaaaatctcccattatcgtagagatccgcattattaatctcaggagcctgtgtagcgttta
taggaagtagtgttctgtcatgatgcctgcaagcggtaacgaaaacgatttgaatatgccttcaggaacaatagaaatcttcgtgcggtgttacgttgaagtggagcggattatgtcagcaatggacagaac
aacctaatgaacacagaaccatgatgtggtctgtccttttacagccagtagtgctcgccgcagtcgagcgacagggcgaagccctcgagtgagcgaggaagcaccagggaacagcacttatatattctg
cttacacacgatgcctgaaaaaacttcccttggggttatccacttatccacggggatatttttataattatttttttatagtttttagatcttctttttagagcgccttgtaggcctttatccatgctggttctagaga
aggtgttgtgacaaattgccctttcagtgtgacaaatcaccctcaaatgacagtcctgtctgtgacaaattgcccttaaccctgtgacaaattgccctcagaagaagctgtttttcacaaagttatccctgctt

attgactctttttatttagtgtgacaatctaaaaacttgtcacacttcacatggatctgtcatggcggaaacagcggttatcaatcacaagaaacgtaaaaatagcccgcgaatcgtccagtcaaacgacct
cactgaggcggcatatagtctctcccgggatcaaaaacgtatgctgtatctgttcgttgaccagatcagaaaatctgatggcaccctacaggaacatgacggtatctgcgagatccatgttgctaaatatgc
tgaaatattcggattgacctctgcggaagccagtaaggatatacggcaggcattgaagagtttcgcggggaaggaagtggtttttatcgccctgaagaggatgccggcgatgaaaaaggctatgaatcttt
tccttggtttatcaaacgtgcgcacagtccatccagagggctttacagtgtacatatcaacccatatctcattcccttctttatcgggttacagaaccggtttacgcagtttcggcttagtgaaacaaaagaaa
tcaccaatccgtatgccatgcgtttatacgaatccctgtgtcagtatcgtaagccggatggctcaggcatcgtctctctgaaaatcgactggatcatagagcgttaccagctgcctcaaagttaccagcgtat
gcctgacttccgccgccgcttcctgcaggtctgtgttaatgagatcaacagcagaactccaatgcgcctctcatacattgagaaaaagaaaggccgccagacgactcatatcgtattttccttccgcgata
tcacttccatgacgacaggatagtctgagggttatctgtcacagatttgagggtggttcgtcacatttgttctgacctactgagggtaatttgtcacagttttgctgtttccttcagcctgcatggattttctcatact
ttttgaactgtaatttttaaggaagccaaatttgagggcagtttgtcacagttgatttccttctctttcccttcgtcatgtgacctgatatcgggggttagttcgtcatcattgatgagggttgattatcacagtttatta
ctctgaattggctatccgcgtgtgtacctctacctggagtttttcccacggtggatatttcttcttgcgctgagcgtaagagctatctgacagaacagttcttctttgcttcctcgccagttcgctcgctatgctcgg
ttacacggctgcggcgagcgctagtgataataagtgactgaggtatgtgctcttcttatctccttttgtagtgttgctcttattttaaacaactttgcggtttttttgatgactttgcgattttgttgttgctttgcagtaaa
ttgcaagatttaataaaaaaacgcaaagcaatgattaaaggatgttcagaatgaaactcatggaaacacttaaccagtgcataaacgctggtcatgaaatgacgaaggctatcgccattgcacagtttaat
gatgacagcccggaagcgaggaaaataacccggcgctggagaataggtgaagcagcggatttagttggggtttcttctcaggctatcagagatgccgagaaagcagggcgactaccgcacccggatat
ggaaattcgaggacgggttgagcaacgtgttggttatacaattgaacaaattaatcatatgcgtgatgtgtttggtacgcgattgcgacgtgctgaagacgtatttccaccggtgatcggggttgctgcccata
aaggtggcgtttacaaaacctcagtttctgttcatcttgctcaggatctggctctgaaggggctacgtgttttgctcgtggaaggtaacgacccccagggaacagcctcaatgtatcacggatgggtaccag
atcttcatattcatgcagaagacactctcctgcctttctatcttggggaaaaggacgatgtcacttatgcaataaagcccacttgctggccggggcttgacattattccttcctgtctggctctgcaccgtattg
aaactgagttaatgggcaaatttgatgaaggtaaactgcccaccgatccacacctgatgctccgactggccattgaaactgttgctcatgactatgatgtcatagttattgacagcgcgcctaacctgggtat
cggcacgattaatgtcgtatgtgctgctgatgtgctgattgttcccacgcctgctgagttgtttgactacacctccgcactgcagtttttcgatatgcttcgtgatctgctcaagaacgttgatcttaaagggttcg
agcctgatgtacgtattttgcttaccaaatacagcaatagtaatggctctcagtccccgtggatggaggagcaaattcgggatgcctggggaagcatggttctaaaaaatgttgtacgtgaaacggatgaag
ttggtaaaggtcagatccggatgagaactgtttttgaacaggccattgatcaacgctcttcaactggtgcctggagaaatgctctttctatttgggaacctgtctgcaatgaaattttcgatcgtctgattaaacc
acgctgggagattagataatgaagcgtgcgcctgttattccaaaacatacgctcaatactcaaccggttgaagatacttcgttatcgacaccagctgccccgatggtggattcgttaattgcgcgcgtagga
gtaatggctcgcggtaatgccattactttgcctgtatgtggtcgggatgtgaagtttactcttgaagtgctccggggtgatagtgttgagaagacctctcgggtatggtcaggtaatgaacgtgaccaggagct
gcttactgaggacgcactggatgatctcatcccttcttttctactgactggtcaacagacaccggcgttcggtcgaagagtatctggtgtcatagaaattgccgatgggagtcgccgtcgtaaagctgctgca
cttaccgaaagtgattatcgtgttctggttggcgagctggatgatgagcagatggctgcattatccagattgggtaacgattatcgcccaacaagtgcttatgaacgtggtcagcgttatgcaagccgattgca
gaatgaatttgctggaaatatttctgcgctggctgatgcggaaaatatttcacgtaagattattacccgctgtatcaacaccgccaaattgcctaaatcagttgttgctctttttctcaccccggtgaactatct
gcccggtcaggtgatgcacttcaaaaagcctttacagataaagaggaattacttaagcagcaggcatctaaccttcatgagcagaaaaaagctggggtgatatttgaagctgaagaagttatcactctttta
acttctgtgcttaaaacgtcatctgcatcaagaactagtttaagctcacgacatcagtttgctcctggagcgacagtattgtataagggcgataaaatggtgcttaacctggacaggtctcgtgttccaactga
gtgtatagagaaaattgaggccattcttaaggaacttgaaaagccagcaccctgatgcgaccacgttttagtctacgtttatctgtctttacttaatgtcctttgttacaggccagaaagcataactggcctga
atattctctctgggcccactgttccacttgtatcgtcggtctgataatcagactgggaccacggtcccactcgtatcgtcggtctgattattagtctgggaccacggtcccactcgtatcgtcggtctgattatta
gtctgggaccacggtcccactcgtatcgtcggtctgataatcagactgggaccacggtcccactcgtatcgtcggtctgattattagtctgggaccacggtcccactcgtatcgtcggtctgattattagtctg
ggaccacggtcccac

8.2.5.2 SynBac1

atgacatcatatttgagaataacaaatgacattatccctcgattgtgttttacaagtagaattctacccgtaaagcgagtttagttttgaaaaacaaatgacatcatttgtataatgacatcatcccctgattgtg
ttttacaagtagaattctatccgtaaagcgagttcagttttgaaaacaaatgagtcatacctaaacacgttaataatcttctgatatcagcttatgactcaagttatgagccgtgtgcaaaacatgagataagtt
tatgacatcatccactgatcgtgcgttacaagtagaattctactcgtaaagccagttcggttatgagccgtgtgcaaaacatgacatcagcttatgactcatacttgattgtgttttacgcgtagaattctactcg
taaagcgagttcggttatgagccgtgtgcaaaacatgacatcagcttatgagtcataattaatcgtgcgttacaagtagaattctactcgtaaagcgagttgaaggatcatatttagttgcgtttatgagataag
attgaaagcacgtgtaaaatgtttcccgcgcgttggcacaactatttacaatgcggccaagttataaaagattctaatctgatatgttttaaaacacctttgcggcccgagttgtttgcgtacgtgactagcga
agaagatgtgtggaccgcagaacagatagtaaaacaaaaccctagtattggagcaataatcgatttaaccaacacgtctaaatattatgatggtgtgcatttttttgcgggcgggcctgttatacaaaaaaat
tcaagtacctggccagactttgccgcctgaaagcatagttcaagaatttattgacacggtaaaagaatttacagaaaagtgtcccggcatgttggtgggcgtgcactgcacacacggtattaatcgcaccg
gttacatggtgtgcagatatttaatgcacaccctgggtattgcgccgcaggaagccatagatagattcgaaaaagccagaggtcacaaaattgaaagacaaaattacgttcaagatttattaatttaattaat
attatttgcattctttaacaaatactttatcctattttcaaattgttgcgcttcttccagcgaaccaaaactatgcttcgcttgctccgtttagcttgtagccgatcagtggcgttgttccaatcgacggtaggattag
gccggatattctccaccacaatgttggcaacgttgatgttacgtttatgcttttggttttccacgtacgtcttttggccggtaatagccgtaaacgtagtgccgtcgcgcgtcacgcacaacaccggatgtttgc
gcttgtccgcggggtattgaaccgcgcgatccgacaaatccaccactttggcaactaaatcggtgacctgcgcgtctttttctgcattatttcgtctttcttttgcatggtttcctggaagccggtgtacatgcg
gtttagatcagtcatgacgcgcgtgacctgcaaatctttggcctcgatctgcttgtccttgatggcaacgatgcgttcaataaactcttgtttttttaacaagttcctcggtttttttgcgccaccaccgcttgcagcg
cgtttgtgtgctcggtgaatgtcgcaatcagcttagtcaccaactgtttgctctcctcctcccgttgtttgatcgcgggatcgtacttgccggtgcagagcacttgaggaattacttcttctaaaagccattcttgt
aattctatggcgtaaggcaatttggacttcataatcagctgaatcacgccggatttagtaatgagcactgtatgcggctgcaaatacagcgggtcgccccttttcacgacgctgttagaggtagggcccca
ttttggatggtctgctcaaataacgatttgtatttattgtctacatgaacacgtatagctttatcacaaactgtatattttaaactgttagcgacgtccttggccacgaaccggacctgttggtcgcgctctagcac
gtaccgcaggttgaacgtatcttctccaaatttaaattctccaattttaacgcgagccattttgatacacgtgtgtcgattttgcaacaactattgtttttttaacgcaaactaaacttattgtggtaagcaataatta
aatatgggggaacatgcgccgctacaacactcgtcgttatgaacgcagacggcgccggtctcggcgcaagcggctaaaacgtgttgcgcgttcaacgcggcaaacatcgcaaaagccaatagtacag
ttttgatttgcatattaacggcgatttttttaaattatcttatttaataaatagttatgacgcctacaactccccgcccgcgttgactcgctgcacctcgagcagttcgttgacgccttcctccgtgtggccgaaca
cgtcgagcgggtggtcgatgaccagcggcgtgccgcacgcgacgcacaagtatctgtacaccgaatgatcgtcgggcgaaggcacgtcggcctccaagtggcaatattggcaaattcgaaaatatata
cagttgggttgtttgcgcatatctatcgtggcgttgggcatgtacgtccgaacgttgatttgcatgcaagccgaaattaaatcattgcgattagtgcgattaaaacgttgtacatcctcgcttttaatcatgccgtc
gattaaatcgcgcaatcgagtcaagtgatcaaagtgtggaataatgttttctttgtattcccgagtcaagcgcagcgcgtattttaacaaactagccatcttgtaagttagtttcatttaatgcaactttatccaat
aatatattatgtatcgcacgtcaagaattaacaatgcgcccgttgtcgcatctcaacacgactatgatagagatcaaataaagcgcgaattaaatagcttgcgacgcaacgtgcacgatctgtgcacgcgt
tccggcacgagctttgattgtaataagtttttacgaagcgatgacatgacccccgtagtgacaacgatcacgcccaaaagaactgccgactacaaaattaccgagtatgtcggtgacgttaaaactattaa
gccatccaatcgaccgttagtcgaatcaggaccgctggtgcgagaagccgcgaagtatggcgaatgcatcgtataacgtgtggagtccgctcattagagcgtcatgtttagacaagaaagctacatattta
attgatcccgatgattttattgataaattgaccctaactccatacacggtattctacaatggcggggttttggtcaaaatttccggactgcgattgtacatgctgttaacggctccgcccactattaatgaaatta
aaaattccaattttaaaaaacgcagcaagagaaacatttgtatgaaagaatgcgtagaaggaaagaaaaatgtcgtcgacatgctgaacaacaagattaatatgcctccgtgtataaaaaaaatattgaa
cgatttgaaagaaaacaatgtaccgcgcggcggtatgtacaggaagaggtttatactaaactgttacattgcaaacgtggtttcgtgtgccaagtgtgaaaaccgatgtttaatcaaggctctgacgcatttc
tacaaccacgactccaagtgtgtgggtgaagtcatgcatcttttaatcaaatcccaagatgtgtataaaccaccaaactgccaaaaaatgaaaactgtcgacaagctctgtccgtttgctggcaactgcaa
gggtctcaatcctatttgtaattattgaataataaaacaattataaatgctaaatttgtttttttattaacgatacaaaccaaacgcaacaagaacatttgtagtattatctataattgaaaacgcgtagttataatc
gctgaggtaatatttaaaatcattttcaaatgattcacagttaatttgcgacaatataattttattttcacataaactagacgccttgtcgtcttcttcttcgtattccttctctttttcatttttctcctcataaaaatta

acatagttattatcgtatccatatatgtatctatcgtatagagtaaattttttgttgtcataaatatatatgtctttttaatggggtgtatagtaccgctgcgcatagtttttctgtaatttacaacagtgctattttctgg
tagttcttcggagtgtgttgctttaattattaaatttatataatcaatgaatttgggatcgtcggttttgtacaatatgttgccggcatagtacgcagcttcttctagttcaattacaccatttttttagcagcaccggatt
aacataactttccaaaatgttgtacgaaccgttaaacaaaaacagttcacctcccttttctatactattgtctgcgagcagttgtttgttgttaaaaataacagccattgtaatgagacgcacaaactaatatc
acaaactggaaatgtctatcaatatatagttgctgatgggtaccttctagaattccggagcggccgctgcagatcccccgggctgcaggggggggggggcgctgaggtctgcctcgtgaagaaggtgttgc
tgactcataccaggcctgaatcgccccatcatccagccagaaagtgagggagccacggttgatgagagctttgttgtaggtggaccagttggtgattttgaacttttgctttgccacggaacggtctgcgttgt
cgggaagatgcgtgatctgatccttcaactcagcaaaagttcgatttattcaacaaagccgccgtcccgtcaagtcagcgtaatgctctgccagtgttacaaccaattaaccaattctgattagaaaaactc
atcgagcatcaaatgaaactgcaatttattcatatcaggattatcaataccatatttttgaaaaagccgtttctgtaatgaaggagaaaactcaccgaggcagttccataggatggcaagatcctggtatcgg
tctgcgattccgactcgtccaacatcaatacaacctattaatttcccctcgtcaaaaataaggttatcaagtgagaaatcaccatgagtgacgactgaatccggtgagaatggcaaaagcttatgcatttctt
tccagacttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcatcaaccaaaccgttattcattcgtgattgcgcctgagcgagacgaaatacgcgatcgctgttaaaaggacaattacaa
acaggaatcgaatgcaaccggcgcaggaacactgccagcgcatcaacaatattttcacctgaatcaggatattcttctaatacctggaatgctgttttcccggggatcgcagtggtgagtaaccatgcatc
atcaggagtacggataaaatgcttgatggtcggaagaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaaacaactctggcgca
tcgggcttcccatacaatcgatagattgtcgcacctgattgcccgacattatcgcgagcccatttatacccatataaatcagcatccatgttggaatttaatcgcggcctcgagcaagacgtttcccgttgaa
tatggctcataacaccccttgtattactgtttatgtaagcagacagttttattgttcatgatgatatatttttatcttgtgcaatgtaacatcagagattttgagacacaacgtggctttccccccccccccctgcag
gaattaagtcttcgaaccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttag
ctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctc
actaaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatcccccgggctgcaggaattcacataacaggaagaaaaatgccccgcttacgcagggcatccatttatt
actcaaccgtaaccgattttgccaggttacgcggctgtcgacctcgagggggggcccggtacccaattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactggg
aaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggac
gcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcttcgaagacgctcgacagcgacacacttgcatcggatgcagcccggttaacgtgccggcacggcctgggtaaccaggtatt
ttgtccacataaccgtgcgcaaaatgttgtggataagcaggacacagcagcaatccacagcaggcatacaaccgcacaccgaggttactccgttctacaggttacgacgacatgtcaatacttgcccttg
acaggcattgatggaatcgtagtctcacgctgatagtctgatcgacaatacaagtgggaccgtggtcccagaccgataatcagaccgacaacacgagtgggatcgtggtcccagactaataatcagacc
gacgatacgagtgggaccgtggtcccagactaataatcagaccgacgatacgagtgggaccgtggttccagactaataatcagaccgacgatacgagtgggaccgtggtcccagactaataatcagac
cgacgatacgagtgggaccatggtcccagactaataatcagaccgacgatacgagtgggaccgtggtcccagtctgattatcagaccgacgatacgagtgggaccgtggtcccagactaataatcaga
ccgacgatacgagtgggaccgtggtcccagactaataatcagaccgacgatacgagtgggaccgtggtcccagtctgattatcagaccgacgatacaagtggaacagtgggcccagagagaatattca
ggccagttatgctttctggcctgtaacaaaggacattaagtaaagacagataaacgtagactaaaacgtggtcgcatcagggtgctggcttttcaagttccttaagaatggcctcaattttctctatacactca
gttggaacacgagacctgtccaggttaagcaccattttatcgcccttatacaatactgtcgctccaggagcaaactgatgtcgtgagcttaaactagttcttgatgcagatgacgttttaagcacagaagtta
aaagagtgataacttcttcagcttcaaatatcaccccagctttttttctgctcatgaaggttagatgcctgctgcttaagtaattcctctttatctgtaaaggctttttgaagtgcatcacctgaccgggcagatagt
tcaccggggtgagaaaaaagagcaacaactgatttaggcaatttggcggtgttgatacagcgggtaataatcttacgtgaaatattttccgcatcagccagcgcagaaatatttccagcaaattcattctgc
aatcggcttgcataacgctgaccacgttcataagcacttgttgggcgataatcgttacccaatctggataatgcagccatctgctcatcatccagctcgccaaccagaacacgataatcactttcggtaagt
gcagcagctttacgacggcgactcccatcggcaatttctatgacaccagatactcttcgaccgaacgccggtgtctgttgaccagtcagtagaaaagaagggatgagatcatccagtgcgtcctcagtaa
gcagctcctggtcacgttcattacctgaccatacccgagaggtcttctcaacactatcaccccggagcacttcaagagtaaacttcacatcccgaccacatacaggcaaagtaatggcattaccgcgag
ccattactcctacgcgcgcaattaacgaatccaccatcggggcagctggtgtcgataacgaagtatcttcaaccggttgagtattgagcgtatgttttggaataacaggcgcacgcttcattatctaatctcc
cagcgtggtttaatcagacgatcgaaaatttcattgcagacaggttcccaaatagaaagagcatttctccaggcaccagttgaagagcgttgatcaatggcctgttcaaaaacagttctcatccggatctga
cctttaccaacttcatccgtttcacgtacaacatttttagaaccatgcttccccaggcatcccgaatttgctcctccatccacggggactgagagccattactattgctgtatttggtaagcaaaatacgtac
atcaggctcgaaccctttaagatcaacgttcttgagcagatcacgaagcatatcgaaaaactgcagtgcggaggtgtagtcaaacaactcagcaggcgtgggaacaatcagcacatcagcagcacata
cgacattaatcgtgccgatacccaggttaggcgcgctgtcaataactatgacatcatagtcatgagcaacagtttcaatggccagtcggagcatcaggtgtggatcggtgggcagtttaccttcatcaaattt
gcccattaactcagtttcaatacggtgcagagccagacaggaaggaataatgtcaagccccggccagcaagtgggctttattgcataagtgacatcgtccttttccccaagatagaaaggcaggagagtg
tcttctgcatgaatatgaagatctggtacccatccgtgatacattgaggctgttccctgggggtcgttaccttccacgagcaaaacacgtagccccttcagagccagatcctgagcaagatgaacagaaac
tgaggttttgtaaacgccacctttatgggcagcaaccccgatcaccggtggaaatacgtcttcagcacgtcgcaatcgcgtaccaaacacatcacgcatatgattaatttgttcaattgtataaccaacacg
ttgctcaacccgtcctcgaatttccatatccgggtgcggtagtcgccctgctttctcggcatctctgatagcctgagaagaaaccccaactaaatccgctgcttcacctattctccagcgccgggttattttcc
tcgcttccgggctgtcatcattaaactgtgcaatggcgatagccttcgtcatttcatgaccagcgtttatgcactggttaagtgtttccatgagtttcattctgaacatcctttaatcattgctttgcgtttttttattaa
atcttgcaatttactgcaaagcaacaacaaaatcgcaaagtcatcaaaaaaccgcaaagttgtttaaaataagagcaacactacaaaaggagataagaagagcacatacctcagtcacttattatcact
agcgctcgccgcagccgtgtaaccgagcatagcgagcgaactggcgaggaagcaaagaagaactgttctgtcagatagctcttacgctcagcgcaagaagaaatatccaccgtgggaaaaactccag
gtagaggtacacacgcggatagccaattcagagtaataaactgtgataatcaaccctcatcaatgatgacgaactaacccccgatatcaggtcacatgacgaagggaaagagaaggaaatcaactgtg
acaaactgccctcaaatttggcttccttaaaaattacagttcaaaaagtatgagaaaatccatgcaggctgaaggaaacagcaaaactgtgacaaattaccctcagtaggtcagaacaaatgtgacgaa
ccaccctcaaatctgtgacagataaccctcagactatcctgtcgtcatggaagtgatatcgcggaaggaaaatacgatatgagtcgtctggcggcctttctttttctcaatgtatgagaggcgcattggagtt
ctgctgttgatctcattaacacagacctgcaggaagcggcggcggaagtcaggcatacgctggtaactttgaggcagctggtaacgctctatgatccagtcgattttcagagagacgatgcctgagccatc
cggcttacgatactgacacagggattcgtataaacgcatggcatacggattggtgatttcttttgtttcactaagccgaaactgcgtaaaccggttctgtaacccgataaagaagggaatgagatatgggttg
atatgtacactgtaaagccctctggatggactgtgcgcacgtttgataaaccaaggaaaagattcatagcctttttcatcgccggcatcctcttcagggcgataaaaaaccacttccttccccgcgaaactc
ttcaatgcctgccgtatatccttactggcttccgcagaggtcaatccgaatatttcagcatatttagcaacatggatctcgcagataccgtcatgttcctgtagggtgccatcagattttctgatctggtcaacg
aacagatacagcatacgtttttgatcccgggagagactatatgccgcctcagtgaggtcgtttgactggacgattcgcgggctatttttacgtttcttgtgattgataaccgctgtttccgccatgacagatcca
tgtgaagtgtgacaagtttttagattgtcacactaaataaaaaagagtcaataagcagggataactttgtgaaaaaacagcttcttctgagggcaatttgtcacagggttaagggcaatttgtcacagacagg
actgtcatttgagggtgatttgtcacactgaaagggcaatttgtcacaacaccttctctagaaccagcatggataaaggcctacaaggcgctctaaaaaagaagatctaaaaactataaaaaaaataatta
taaaaatatccccgtggataagtggataaccccaagggaagtttttttcaggcatcgtgtgtaagcagaatatataagtgctgttccctggtgcttcctcgctcactcgagggcttcgccgtcgctcgactgcg
gcgagcctactggctgtaaaaggacagaccacatcatggttctgtgttcattaggttgttctgtccattgctgacataatccgctccacttcaacgtaacaccgcacgaagatttctattgttcctgaaggcat
attcaaatcgttttcgttaccgcttgcaggcatcatgacagaacactacttcctataaacgctacacaggctcctgagattaataatgcggatctctacgataatgggagattttcccgactgtttcgttcgctt
ctcagtggataacagccagcttctctgtttaacagacaaaaacagcatatccactcagttccacatttccatataaaggccaaggcatttattctcaggataattgtttcagcatcgcaaccgcatcagact
ccggcatcgcaaactgcacccggtgccgggcagccacatccagcgcaaaaaccttcgtgtagacttccgttgaactgatggacttatgtcccatcaggctttgcagaactttcagcggtataccggcata
cagcatgtgcatcgcataggaatggcggaacgtatgtggtgtgaccggaacagagaacgtcacaccgtcagcagcagcggcggcaaccgcctccccaatccaggtcctgaccgttctgtccgtcactt
cccagatccgcgctttctctgtccttcctgtgcgacggttacgccgctccatgggtattttcagttgtgccaccatcgtctgcagctggctgacgtaccaggagtcagagagcggaaccagccggtgagtct
gctgaccggcgggcattctccccgccgtcctggcagctttttcggtccgttgtttcagggtcgcaagctgcacaaacggatacggaggcgcaagcgaaaaatccccccgcgtcagcgccagtgcttcatt
aatgcgtgctccggtgttccacagtgtggccagcagcatcttgcggtgcagatccgggacgtaatggagcagggcactcacttccggagccagcagatattttggcagttcatcatggaccatcgacatct
ggcgaagtgccagagctgccggataatcaatggcaacaggcagcgatgcaggctgcccggcagaatacactgccggtaccatgactgcagactggctgtgtataacggagcctgacatttatattcccc
agaacatcaggttaatggcgtttttgatgtcattttcgcggtggctgagatcagccacttcttccccgataacggagaccggcacactggccatatcggtggtcatcatgcgccagctttcatccccgatatg
caccaccgggtaaagttcacgggagactttatctgacagcagacgtgcactggccaggggggatcaccatccgtcgcccgggcgtgtcaataatatcactctgtacatccacaaacagacgataacggct
ctctcttttataggtgtaaaccttaaactgcatttcaccagtccctgttctcgtcagcaaaagagccgttcatttcaataaaccgggcgacctcagccatcccttcctgattttccgctttccagcgttcggcac
gcagacgacgggcttcattctgcatggttgtgcttaccagaccggagatattgacatcatatgccttgagcaactgatagctgtcgctgtcaactgtcactgtaatacgctgcttcatagcacacctctttttga

catacttcgggtatacatatcagtatatattcttataccgcaaaaatcagcgcgcaaatacgcatactgttatctggcttttagtaagccttatgtattttacctttcgttatgttaaccaataaaaattaaaatctg
acttataaaaacaaagcgtaattaccgcattcccgtttcgtatggatctgatcctttcctgggacccggcaagaaccaaaaactcactctcttcaaggaaatccgtaatgttaaacccgacacgatgaagc
ttgtcgttggatggaaaggaaaagagttctacagggaaacttggacccgcttcatggaagacagcttccccattgttaacgaccaagaagtgatggatgttttccttgttgtcaacatgcgtcccactagacc
caaccgttgttacaaattcctggcccaacacgctctgcgttgcgaccccgactatgtacctcatgacgtgattaggatcgtcgagccttcatgggtgggcagcaacaacgagtaccgcatcagcctggcta
agaagggcggcggctgcccaataatgaaccttcactctgagtacaccaactcgttcgaacagttcatcgatcgtgtcatctgggagaacttctacaagcccatcgtttacatcggtaccgactctgctgaa
gaggaggaaattctccttgaagtttccctggtgttcaaagtaaaggagtttgcaccagacgcacctctgttcactggtccggcgtattaaaacacgatacattgttattagtacatttattaagcgctagattct
gtgcgttgttgatttacagacaattgttgtacgtattttaataattcattaaatttataatctttagggtggtatgttagagcgaaaatcaaatgattttcagcgtctttatatctgaatttaaatattaaatcctcaata
gatttgtaaaataggtttcgattagtttcaaacaagggttgtttttccgaaccgatggctggactatctaatggattttcgctcaacgccacaaaacttgccaaatcttgtagcagcaatctagctttgtcgatatt
cgtttgtgttttgttttgtaataaaggttcgacgtcgttcaaaatattatgcgcttttgtatttctttcatcactgtcgttagtgtacaattgactcgacgtaaacacgttaaataaagcttggacatatttaacatcgg
gcgtgttagctttattaggccgattatcgtcgtcgtcccaaccctcgtcgttagaagttgcttccgaagacgattttgccatagccacacgacgcctattaattgtgtcggctaacacgtccgcgatcaaatttg
tagttgagctttttggaattatttctgattgcgggcgtttttgggcgggtttcaatctaactgtgcccgattttaattcagacaacacgttagaaagcgatggtgcaggcggtggtaacatttcagacggcaaatct
actaatggcggcggtggtggagctgatgataaatctaccatcggtggaggcgcaggcggggctggcggcggaggcggaggcggaggtggtggcggtgatgcagacggcggtttaggctcaaatgtctct
ttaggcaacacagtcggcacctcaactattgtactggtttcgggcgccgtttttggtttgaccggtctgagacgagtgcgatttttttcgtttctaatagcttccaacaattgttgtctgtcgtctaaaggtgcagc
gggttgaggttccgtcggcattggtggagcgggcggcaattcagacatcgatggtggtggtggtggtggaggcgctggaatgttaggcacgggagaaggtggtggcggcggtgccgccggtataatttgttc
tggtttagtttgttcgcgcacgattgtgggcaccggcgcaggcgccgctggctgcacaacggaaggtcgtctgcttcgaggcagcgcttggggtggtggcaattcaatattataattggaatacaaatcgtaa
aaatctgctataagcattgtaatttcgctatcgtttaccgtgccgatatttaacaaccgctcaatgtaagcaattgtattgtaaagagattgtctcaagctcggatcccgcacgccgataacaagccttttcatt
tttactacagcattgtagtggcgagacacttcgctgtcgtcgacgtacatgtatgctttgttgtcaaaaacgtcgttggcaagctttaaaatatttaaaagaacatctctgttcagcaccactgtgttgtcgtaaa
tgttgtttttgataatttgcgcttccgcagtatcgacacgttcaaaaaattgatgcgcatcaattttgttgttcctattattgaataaataagattgtacagattcatatctacgattcgtcatggccaccacaaatg
ctacgctgcaaacgctggtacaattttacgaaaactgcaaaaacgtcaaaactcggtataaaataatcaacgggcgctttggcaaaatatctattttatcgcacaagcccactagcaaattgtatttgcag
aaaacaatttcggcgcacaattttaacgctgacgaaataaaagttcaccagttaatgagcgaccacccaaattttataaaaatctattttaatcacggttccatcaacaaccaagtgatcgtgatggactac
attgactgtcccgatttatttgaaacactacaaattaaaggcgagctttcgtaccaacttgttagcaatattattagacagctgtgtgaagcgctcaacgatttgcacaagcacaatttcatacacaacgaca
taaaactcgaaaatgtcttatatttcgaagcacttgatcgcgtgtatgtttgcgattacggattgtgcaaacacgaaaactcacttagcgtgcacgacggcacgttggagtattttagtccggaaaaaattcga
cacacaactatgcacgtttcgtttgactggtacgccgtcggcgtgttaacatacaagttgctaaccggcggccgacacccatttgaaaaaagcgaagacgaaatgttggacttgaatagcatgaagcgtc
gtcagcaatacaatgacattggcgttttaaaacacgttcgtaacgttaacgctcgtgactttgtgtactgcctaacaagatacaacatagattgtagactcacaaattacaaacaaattataaaacatgagtt
tttgtcgtaagttttacgagtagaattctacgtgtaacacacgatctaaaagatgatgtcattttttatcaatgactcatttgttttaaaacagacttgttttacgagtagaattctacgtgtaaagcatgatcgtga
gtggtgttaataaaatcataaaaattattgtaaatgtttattatttaaaaacgattcaaatatataataaaaacaatctacatctatttcttcacaatccataacacacaacaggtccatcaatgagtttttgtcttt
atccgacatactatgtgcatgtaacaaatcaaatacatcttttaaatttttatacacatctttacattgtctaccaaaatctttaataaccctataacaaggaaaagacttttcttcttgcgtggttttgccgcgca
gatattgaaataaaatgtgcatgcacgacaacttgtgtttactaaaatgctccttgcctataccgcaaaaccggccatacatttcggcgattacacgcggacaattgtacgattcgtctacgtgtaaacgatc
atcataatcactcttgcgcaaacgaataaatttttcaccgcttccgacaaacgaggcaccaattcggcgggcacgcttcgatacattattctgtgcacataagttaccacacaaaatttattgtaccaccat
ccgacaacgtcgttattagggttgaacacgttggcgatgcgcagcagtttcccgtttctcatgaaatattcaaagcggcccaaaataatttgcaagcaatccaacatgtcttgagaaatttctcgttcaaaatt
gttcaaagagaatatctgccatccgttttgaacgcgcacgctgacgggaaccaccgcatcgatttgctccaacacttcacggacgttatcgtcgatgcccatcgtttcgctggtgctgaaccaatgggaaa
ggctcttgatggaatcgcccgcgtctatcatcttgaccgcttcgtcaaaggtgcaactgccgctcttcaaacgccgcatagcggtcacgtcccgctctatgcacgacataccgtttacgtacgattctgata
ggtattcctgaactatacggtaatggtgatacgactcgccatacacgtcgtgcacctcattgtatttagcataataattgtaaattattaactttgcagcgagagacatgttgtcagtaaagcggtgctaggctc
aataatactgatgtacaggcacgcgtgctatttatatataatttcgcaaggaggggagctgttatcggttgctattattaaagaatggccgtctgtttttatcacaagcttggcagcctcaaccatgaagcgtcg
tcattgtaaattaaattctctgcctcaagaattatttgacaagattgtcgagtatttatctttatctgattactgcaatttggtgcttgtctgtaaaagaccttctagtaaatataacgtgatatttgatagtactaatc
accaacatttgaaaggcgtgtacaaaaagacagacgtgcaaataacaagctacaacgaatacatcaactgtatttgcaacgaactgagacaagacgaattctatgccaaatcatcatggattgcgagta
tttgcggtcaccagagagcgacaatttttagtgtaacaaataaacaagtagaaatgaaatatcatttgtataatatagcaattgtggaaagtgaagattgcaacggattttacccatttgagccaacgcgcga
ttgtttaatatgcaaacaaaaaaaccaatgtcctcgtaattcatttattgtttcgttgtgtaaatatttagaaaaacaaaatgtacaatcaaactttatatattatttatacgaaataaatacataataataactatt
atacatgtttttattttacaatacttcctgtataacctctctaactacattaggagtacaatccacgtcaattacacgtttagctatttttctaattttgtaatgtttatcgtagagtttttcgttaatacattgaatagcc
aacaagggatttgggtgcacaccgtcatagagtacttccatgtcgtcttcaaagcgcatttttcgcttgcgaaaatgccgctcttggcccaaaacaaaagcgagtttgatgcggtcgtcgatgcgttccgaa
aatacggccaaatgctggtgtttggtgatgtcgcgcggaaacgtcaccgtgccatttttgctttccgccacgacggcggttttcaatttttcggccgactgcagcatgttaagtttggcgtcgagttcgtgcaaa
cgcaattcaaactgctcaaacctgttgcccacctcgttcttgaacgtctcgtgggtgaccataaatttttcgctgtttgcattcagtttctttacatgttttaaaacagattcaatcttgtcgcgcaaatcatcacg
ctcgccttcagtttgaatgtgcagcaacgcgttgcttttgttggcaaaatttaaccgcatcaaaatttccaacaacccgtgcttggtcgcgaacaatgcgcccaacgagttgagatcgcgtttggatctctgttt
gtgaaaaacaatttcgtttaaatggtaaacttgatcgccgtcccaattgcaatcaagtatgtcgtcgtgcgcaatttcaagacctttgcaaaaatctatcacattgtagcattttgcgttcgtgtcgctgtgcacg
tatctgtacttgaaactgtgcgtgttgcatttgaatgagtcccatttaacgatgtgcgaccattgttgggcgtttatgtggtactttttgtagtcgtctgcattgaaccgatcttcggcggcgatggcgtcgttgtcgt
tgtcaccggaccacatccaccagttccataaccaggatagcattgctttagcttgtctagcaattcctttgttatacaacgagaaaatttcgttcccttataattatagctgtacggtgcgcgtatttgtttgttaa
cgttacaaaaaatatccctgtccacgtccggccaatactgcaacgtgagcgcgtccaagtttgaatcttgcatatgcggaacgtacaaacgtacggcctctctcacacaatgcgcaaaactgcccggctg
aatgtaatcactgtccaactttgcaggtttctcgaaagccttgtaccgatgcacgcgaacattttgagcggacgtgattttaaacttgtcggtgaattttaaccacaaatgaaatccacggttgccggtataca
tgactcttgacacgttctcttccgtgtaaaacaacagaaacgccgtggcgccaatgtaaattttcagcattaaatcgtgttcgtcaacataatttttgtaatcggcgtctacgacccattccctgccgccgccg
tcgtccaacggtttgacgtgcacgtcggacactttgttttgcacaatataactatacaattgtgcggaggtatcaaaatatctgtcggcgtgaatccagcgcgcgttgaccgtcatgaacgcgtacttgcggct
gtcgttgtacgcaatggcgtcccacatcatgtcgacgcgcttctgcgtataattgcacactaacatgttgccctttgaacttgacctcgattgtgttaatttttggctataaaaaggtcaccctttaaaatttgtta
cataatcaaattaccagtacagttattcggtttgaagcaaaatgactattctctgctggcttgcactgctgtctacgcttactgctgtaaatgcggccaatatattggccgtgtttcctacgccagcttacagcc
accatatagtgtacaaagtgtatattgaagcccttgccgaaaaatgtcacaacgttacggtcgtcaagcccaaactgtttgcgtattcaactaaaacttattgcggtaatatcacggaaattaatgccgacat
gtctgttgagcaatacaaaaaactagtggcgaattcggcaatgtttagaaagcgcggagtggtgtccgatacagacacggtaaccgccgctaactacctaggcttgattgaaatgttcaaagaccagtttg
acaatatcaacgtgcgcaatctcattgccaacaaccagacgtttgatttagtcgtcgtggaagcgtttgccgattatgcgttggtgtttggtcacttgtacgatccggcgcccgtaattcaaatcgcgcctggc
tacggtttggcggaaaactttgacacggtcggcgccgtggcgcggcaccccgtccaccatcctaacatttggcgcagcaatttcgacgacacggaggcaaacgtgatgacggaaatgcgtttgtataaa
gaatttaaaattttggccaacatgtccaacgcgttgctcaaacaacagtttggacccaacacaccgacaattgaaaaactacgcaacaaggtgcaattgcttttgctaaacctgcatcccatatttgacaa
caaccgacccgtgccgcccagcgtgcagtatcttggcggaggaatccatcttgtaaagagcgcgccgttgaccaaattaagtccggtcatcaacgcgcaaatgaacaagtcaaaaagcggaacgattt
acgtaagttttgggtcgagcattgacaccaaatcgtttgcaaacgagtttctttacatgttaatcaatacgttcaaaacgttggataattacaccatattatggaaaattgacgacgaagtagtaaaaaacata
acgttgcccgccaacgtaatcacgcaaaattggtttaatcaacgcgccgtgctgcgtcataaaaaaatggcggcgtttattacgcaaggcggactacaatcgagcgacgaggccttggaagccgggata
cccatggtgtgtctgcccatgatgggcgaccagttttaccatgcgcacaaattacagcaactcggcgtagcccgcgccttggacactgttaccgtttccagcgatcaactactagtggcgataaacgacgt
gttgtttaacgcgcctacctacaaaaaacacatggccgagttatatgcgctcatcaatcatgataaagcaacgtttccgcctctagataaagccatcaaattcacagaacgcgtaattcgatatagacatg
acatcagtcgtcaattgtattcattaaaaacaacagctgccaatgtaccgtattcaaattactacatgtataaatctgtgttttctattgtaatgaatcacttaacacacttttaattacgtcaataaatgttattca
ccattatttacctggtttttttgagaggggctttgtgcgactgcgcacttccagcctttataaacgctcaccaaccaaagcaggtcattattgtgccaggacgttcaaaggcgaaacatcgaaatggagtctgt
tcaaacgcgcttatgtgccagtagcaatcaatttgctccgttcaaaaagcgccagcttgccgtgccggtcggttctgtgaacagtttgacacacaccatcacctccaccaccgtcaccagcgtgattccaa
aaaattatcaagaaaaacgtcagaaaatatgccacataatatcttcgttgcgtaacacgcacttgaatttcaataagatacagtctgtacataaaaagaaactgcggcatttgcaaaatttgctaagaaaa
aagaacgaaattattgccgagttggttagaaaacttgaaagtgcacagaagaagacaacgcacagaaatattagtaaaccagctcattggaaatactttggagtagtcagatgtgacaacacaattcgca

caattattggcaacgaaaagtttgtaaggagacgtttggccgagctgtgcacattgtacaacgccgagtacgtgttttgccaagcacgcgccgatggagacaaagatcgacaggcactagcgagtctgct
gacggcggcgtttggttcgcgagtcatagtttatgaaaatagtcgccggttcgagtttataaatccggacgagattgctagtggtaaacgtttaataattaaacatttgcaagatgaatctcaaagtgatattaa
cgcctattaatttgaaaggtgaggaagagcccaattgcgttgagcgcattaccataatgccatgtattttaatagatactgagatctgtttaaatgtcagatgccgttctccttttgccaaattcaaagtattgatt
attgtagatggctttgatagcgcttatattcaggctaccttttgtagcattagcgatagtgtaacaattgttaacaaatctaacgaaaagcatgtaacgtttgacgggtttgtaaggccggacgatgaaggtaca
acaatgccttatgtcattggaccattatattctgtcgacgctgctgtcgccgaccgtaaagtgaaggacgtggtggattcaattcaaaaccaacagacaatgttaaaagtatttattaacgaggctaatgtgta
taacaaatggaatatgcttaaaggtttaatttataataataacaatgaatctgttttagtaaaataatgtagtaaaatttataaaggtagataaaaattataatattaataaaaaaaataatgttactaaatgggtt
cctgcgttaaattattttacgggtagacagctattaactattttatttatttttaaatttaaataaatgtattgttagaaaattgtgttgttttattagtataacgaaaaaatacatgacataaaccgcttccaattttgg
tcacacaaactcttgtgtggatagtttacgtaatgagttaaataggcgggcagttgtccgctaaacgtgtcggtggtcaagtagatgtgcattaatttacgacaacccaaagcggggccgcttatgtcaagta
tttttttcacaaaattggtaatggtttcgttttgttccttgtacaaacacatgtcggtgtgatcgttgacgcacgagttgtacgattccgccggcaggttggcaaacaagcgcttgagatgcttgagtctgcgttca
attttataatcaaacttgttggtgaaaatgtctttcagcaagcacattaactggtcgttcaaaacgcgctgcaacgacgacaccaacacatgatattcgtttccaaaaagcgaaaaatttttgatgcagcggt
ccgcgttgaagggtcgtttcataatgcgcacgttgacaaaaaacacgttgaaagacagcggggctgtggttattttaacgccgttgtcggtatactcgtcgacgccgtctgcgcttgttatgtcaatttgtagc
gcaaatctaaccaaatcaaactcatcgttgtactgtgtctttatgcattttatatggcggtttaagtgcaagttgatttggccgtttaatctataggctccgttttgataacatttcagcactaccaacggatccga
catgtaaacttgacgcgttagcacgtccaattcagcgtaatgttggtcgacgcatttttgtaaattagtttgcaggttgcaaaacatttttgcgcaaaagccgtaatagtcaaaatctatgcattttaatgcgctt
ctgtcgtcgtcaatatggcatgtcacggctgcgcctccagttaacacgaataaaccgccgttttcgcaaactacggcttcgaaacaatctttgataaatgccaactttgctttagccacaattttatcgcgca
ggcgatcttcaatatcctttgtcgtaatataaggtaggacgccaagatttagttgattcaacaaacgttccataatgaatagcggcgacgcaacacgactacactgttcaaatgcgcacgcaaaacaaac
ccttgcaactttatttgccaatcgtaatcacagtagtttttacgagtacgccatcgcgtttgtaagcacattgctttttaaaaataatttaaatttaatgaccgcgtgcaatttgatcaactcgttgatcaactttga
actcaacatgtttggtaaaagtttattgctaaatggatttgttaatttctgcattgctaacagcgacggggttacgattcaacataaaatgttaaccaacgtgttaagtttttttgttggaaaaatattattaaaaata
aataaataaacttgttcagttctaattattgttttatttttttataaaataatacaattttatttatacattaatactttggtatttattaatacaattatttacaatactttatttacactataatactttatttacattagtact
aaattaatactaaattacgctaatactaaattaatactttatataatcaaaaataatactttatataatactttctaatcatcataaacgggtaatagttttttctcttgaaatttacgctgcaactcttcgctaaaa
cacatgggcggtggagtgggagcgggtggagtaggagtccttacgggtttgatgggcgacagttctctggacttgcggaacagcttgggcgaaaacgtcggcgtgcgccgactaatgatttcttcatcgca
cgaggcgtcgcacattgtgcacgcgtccggtgaggtacacaaaactttcttgggcacgctgtacaccggcttgggcacgctatatgtgttgccaaaatagaactcgttgtggttgccgaacggagacgatg
ggtgtgaagacggcgatggctgtgaagacaagtccgaaggcgcgataaaagatgaaagtgtttctgaaaccgaagtggtggtagaagtggtagaaggcgggtgcgttacggcaaccacgctgctgctat
ttctgccttcggagaccacttccagcaatctagagttactctctcgttcttcgcggcgatagtcaatgtcgcaataatgttcataagatgccttttcggcttcggcgcgccttttcatgtatatgttgtgacgcatc
tcctttaactgcacgtacaaattccagcattgcacagccagtatcgtaagcacgcccattatgattacgggataattttgattaaacacggtcggctcgtgatcgcttacaatcgctcggcacatgatgcattt
tttgtaaatgttcacatacacacagttttggctcaaggtttcggtatttgcgtagtcaatttccagatacacgatagagttccagcacattgattccaaatcgtagtgacgatataaaacatctagcgccggtag
atgaccatttttgaacacgtagatttgaaacgcggcaaacagcatccaacacagcccagtgatcacgtttaccataatacacgtgatagcgacgtaaaagttttctttcgcattgaaatttacatttgtgtttg
aagagctgctgcgatttttcgtccacacgataatcttccatataaaataaaacatgtaaaataatatccacatgccgaacgccagcattatcggtatagatagattgataaccgattgctttccttcaatttcc
agcaaaaacgcgtatctgctgtctatcactcccattatagataacacaaacactatcagatatgctaataataatgaggcattaagcccgaattgtaaaactgcagtgattttatttaacattttgaatatttaa
ttcaacaactaagtaatggcaatatgtatcgagtactgatcgtgtttttcctgttcgtgtttctttatatagtgtaccagcccttttatcaggcatacttgcatatcggacatgcccaacaagattacaatgacacg
ttggacgataggatggattacattgaatccgtaatgcgtagaaggcactacgtgccgattgaagcgttgcccgcaatcaggtttgatactaatctcggcacgttggccggtgacacgattaaatgcatgtcg
gtgcctttgtttgttagtgacattgacctgccgatgtttgattgtagtcagatatgcgataacccgtctgcggcgtatttctttgtcaacgaaacggatgtgtttgtggtcaacggccacagactgacggtgggcg
gatactgctccactaatagtttgccccgcaactgtaatcgcgagacgagcgtcattttaatgagtctcaatcagtggacgtgcatagccgaggacccgcgttactatgcgggcacagataacatgacgca
actcgcaggcagacaacactttgaccgcattatgcccggacagagtgataggaacgtcctgtttgaccgattactaggccgagaggtgaacgtgaccactaacacgtttcgccgcagctgggacgagtt
gctggaggacggcactaggcggttcgaaatgcgctgcaacgcccgagataacaacaataatctcatgtttgttaatccgcttaatcccctcgagtgtctcccgaacgtgtgcactaacgttagcaacgtgc
acaccagtgttagacccgtatttgaaacgggagagtgtgactgcggcgacgaagcggtcacgcgtgttacgcacattgtgccgggggacaggacctctatgtgtgccagcattatagatggcctggataa
aagtacggcatcatatagatatcgcgtagagtgcgttaatctgtacacctctattctaaattattctaataacaaattgttatgtcccagtgacacttttgatagtaacacggacgcagcttttgcctttgaagtg
cccggctcctaccctttatcgcgcaacggcatcaacgagccaacttatcgcttttatcttgataccagatctcgagttaattacaatgacgtcagagggcagttatcttaattgtgataacacaaacaataag
tcatttaaatgttacgtcagtagttagtatataagccgtacatgttggcttgcaaattcagtcaatatcaggcttttatcatggacggtgtaaagctgctagggacgtgcgcgctaataattttgttatcgacgac
gagtacagttgtcgggcgtgaccgtatcacgtttacgccgatagaagatagcgcaggcctcatgtttgaacgcatgtacggcttgcgacatcatacagacgacagatttgtgtttgtgaaaaaattcaatttt
gtttcggtgctgcaagagctcaataatatcaaatctaaaattgaattatatgaagcgcaagtttcaacttgcacaaacgtcagacaaataaaacagaacagatcgagtatcatcaaagctcgcattgaaaa
tcagctgcagtttttgacgcaactaaacaaaaatctcatcacatactctgtggaaagcagcattttaagcaacgacgtgctggacaacatcgatctggaatatgacgacagcggtgagtttgacgtttacga
cgaatacgaacagccttcgcattggagcaacatgactgtatccgacgcgcaagctttgctccgaaacccgcccaaagacagagtaatgtttttggacacggttaccaccagcgacgtgagcagcaaat
acgaagaatacataaactgcattgtgagcaaccgtaccgttgaaaacgagtgcatgtttttagccaacatgatgaacgtgctcaacgacaaattggacgacgcagcagctttggccaagatgctggagc
gaatagtaaaacaaacgcgaaagaacaaactcaacatctccaacacggttatagacgacgacacgctgctaacggaaatgaaaaaattaacacaaactttatacaaccaaaaccgcgtgtgggtagt
ggattttaacaaggacatgaatagttatttcgatttgtcgcaagcgtataaattgcatttatatgttgatttaaacacggtcattatgtttattaccatgccattgttaaaatccaccgccgtttcgtttaatttgtatc
gcgtcatgacggtgcctttttgcaggggcaaaatgtgtctgcttatcatttcgggcaatgaatactttgggattacagacagcaaaaactattatgtgcccgtatctgataactttagacaagattgccaagag
tttacgggctacaatgagtttttgtgtcccgaaactgagccgattgccactatgaactcgaaagtgtgcgagattgaaatgtttatgggtcgatatagcgacgacgtggacaacatgtgcgacattagggtgg
ccaattataatcccaaaaaagcttacgtgaacactttaatagactaccgaaaatggttgtacatttttccaaacacgaccgtgtccgtccactattattgtcacgacgcgcttgtagaagttgatacaaaagt
ttcgcccggcgttggtgttatgttttcgactatggcgcaaacgtgttcgattagaataacgtatgatgtgaccataactgtagattcgcgattttatgtcagccattcaactacatactggcctaaaaagaaattt
aattttaacaactacatcgaccaaatgttgcttgaaaaagcgaccaccagttttataccgactgttgacaattttacccggcccgttttattgcaacttcctcataaatttcacattaaagattacacatcgacg
ccccatcattttttccatcagtctaaaatttacaccaacagcgcggcgcccgacgaagactcgcaagacgacagtaataccaccgtggttattatcgctattgtcgctgcaatgatcctattctgtggattatt
gttatttttgttttgctgtataaaaaaacggtgtcatcaatcaaataacgtggttgtgcaatacaaaaataacaatgaatttgtcacaatttgcaataatttagaagacaatcgagcatacattaatttacctaat
gaatacgatagcgatgatatgccaaaaccattgtaccctttacttggctttaatgatgatttgttaaaagatgataaacctgtgttgtaccctatgattatagaaagaataaaataaaacatgtataattgaaat
aaatatattatttaataaaatgtttttatttatatactattttctattacatattccaatgcacacaaatgtttaatggctatcagttttaattttactaattcgtctaaacaaaaattattcacttgctgtttttcatccat
ttgacatatggcgtttataaataattcgctgtgttttatgaacgaatcgtaaaccgctgcctgggccttcagcacggtcggcgcattgtatttttgggtaaagtacgcaatatttttagtcaaacacagagattta
aatctttttcatttatatccaagtcggaacaatcgtatacaaaatctagcttttcactttcgggcgcgcccagatactggtttacgagttcgagctgctccacttggcctttgatatcggccgctatgcacaacat
tttgtcgattgcagtttcattgtttttaacataataatttttaactttttattttgcaatttaatcaaactatttaaattcgcttgacctttcttacaaagcgcagttaatatgcaagacattttgacttataataaaaaa
caaaacttttatatattcatttattgttcaataataacaaatattccaggcttaaaagctaacgaatagggcttttcggtaattttcttattattcatgtccgtcatctgcatctctttgccgtacttgacgccgtcaat
ggtgcccatcatgtacattttaatctcctccgaaggtccgtctattttgtccatttcgaacaatctatcaaaatcttcaacgctcattctctgcatatcaagaggaacgtttctgatctttccggtggcgtaaattg
atccgttgttgtcacggttgattatgtaaaaccgacgaatcaacatgtcgcgctcgctagttttgttcttatccggcaaatgaatgcacacgtttggttccatcttcaaaggaaaatcgctttgcaagtgtttttgc
aaaatgttgccaaatatattgttgtgtttgtgaatgtctccgtattgaatgctaaaaaactggccaaagttgcttttggcacgttttatggttccaaagtcggaaaaccaaaatccgcagggcttgccctgcact
cttggaccgatggtgtacgtagtcttgccgttggccggctccaacaccacgatatttttatcgggctcgggatacaacttgtcttcccattcgtgcaaactgttcaaattagacagtcgacaaaattcgtttttc
aaaaatctgccttcgaaacaactacaattcagtattgaaaagttgcctcgtttcacattaatcgccatctgctcctgccacaacatcttcgtcaactcgtgtggctccaattgaatggacgacggcgtaaaat
agcacattacgcccgtttcgtcgtgtttcacgttaaaagcgccgctgttgtacggcaccagctgctggtcctcaccaccttccgatctttcccgcttcggctggttgtcgtcgctgctcgaatatccatcgcca
atcttgcgtttagttgccatgctaccgacgtgcgctgtctgctgtggttcaagtctaattgaagtgtttcacagaatataagatatataataaatatggacgactctgttgccagcatgtgcgtagacaacgcgtt
tgcgtacactactgacgatttattgaaaaatattccttttagtcattccaaatgcgcccctttcaagctacaaaattacaccgttttgaagcggttgagcaacgggtttatcgacaagtatgtggacgtgtgctct
atcagcgagttgcaaaagtttaattttaagatagatcggctaaccaactacatatcaaacattttcgagtacgagtttgtagttttagaacacgatttgtccacagtgcacgtcattaacgccgaaacaaaaa

ccaaactgggccatataaacgtgtcgctaaaccaaaacgacgcaaacgtgctcatttgaccgtaactttaacgagctaaaatgaacgaggacacgcccccgttttattttatcagcgtgtgtgacaacttt
cgcgacaacaccgccgaacacgtattcgacatgttaatagaaagacatagttcgtttgaaaattatcccattgaaaacacggcgtttattaacagcttgatcgttaacgggtttaaatacaatcaagttgac
gatcacgttgtgtgcgagtattgcgaagcagaaataaaaaaattggtccgaagacgagtgtattgaatatgcacacgtaaccttgtcgccgtattgcgcgtatgctaacaagatcgccgagcgtgaatcgttt
ggcgacaacattaccatcaacgctgtactagtgaaagaaggcaaacccaagtgtgtgtacagatgcatgtccaatttacagtcgcgtatggatacgtttgttaacttttggcctgccgcattgcgtgacatga
ttacaaacattgcggaagcgggactttttacacgggtcgcggagacgaaactgtgtgtttctttgcgactgttgcgtacgtgattggcatactaatgaagacacctggcagcgacacgccgccgaaaacc
cgcaatgttattttgtattgtcggtgaaaggtaaagaattttgtcaaaactcaattactgtcactcacgttgataaacgtgacgacgacaatttaaacgaaaacgccgacgacattgaggaaaaatatgaatg
caaagtctgtctcgaacgccaacgcgacgccgtgcttatgccgtgtcggcatttttgcgtttgcgttcagtgttattttggattagatcaaaagtgtccgacgtgtcgtcaggacgtcaccgattttataaaaat
atttgtggtgtaataaaatggtgttcaacgtgtactacaacggctattatgtggaaaaaaaattctccaaggagtttttaattcatattgcgcctgatttgaaaaacagcgtcgactggaacggcagcacgcgc
aaacagctgcgcgttctagacaagcgcgcctacaggcaggtgttgcactgcaacggcagatactactggcccgatggcacaaagtttgtctctcatccgtacaacaaatctattcgcacgcacagcgca
acagtcaaacggaccgacagctcgcatcgattaaaaagccacgtggtcgacaaacgaccgcgccgctctttagattctcctcgcttggacggatatgttttggcatcgtcgcccataccacacagcgact
ggaatgaagaactaaagctgtacgcccagagccacggctacgacgactacgacgacaatttagaagatggcgaaatcgacgaacgtgactctttaaaaagtttaaataatcatctagacgacttgaatg
tattagaaaaacaataaaacatgtattaaaaataataataataaaactatatttgtaatatataatgtattttatttaaaaattgtctattccgtagttgagaaagttttgtcttgacttcataactctcttctccata
ttctgcagctcgtttacgtttttgtgacgcttttaattttctcaaaatgctggctgtcaatagttattttttgcttttgtctattaatttcttccaattgagattttaaatctcgctgagattgagatgcgttgtaattccttg
agaacatcttgagaaaacatacagatgaggtaaaacagcatcttttatccaaattaggagttaattattattcatttgtatcgcgaccatttgctcgtacacatcttccataaaatggttatttttattgcgataag
tgttggcattgacattttgcaaatgtcgtaggttaaaggggcaaatgggctgcgtggccgataaaagattccagttcaacaatccctcttcgcccccgtttaacttgaaaatggcgctacacgtttctacgcta
tcgtgttcctgttgagtggcgcacggttcgaccagtatcatcttgtgatatgcggttttgacattcatgtgcaacggaataacttgcgggtcatcgcattcgtcggaattaagctttaaatggcgtccgtatgcttt
ccaaagtttttcgtcgtcgaaccgcggcactgcttgcaagtcgacgcggggaaacggcgctctgtacaaaacgcctaaattcaaaaactgattgcattgttgcagctctgtccaatcgacgcgatttttgta
attttgaaacagcatcaggttgaacgccgcgctggcgcgcacgtttgtaatcactgtgtaattgatcagcttgtgccaatactgggcattgaaattttcttcaaactcatttctaaactctggatgcgcaaacat
gtgtctaatgtagtacgcgggcggggcgttgaacgcagtccatttgtcaatacacttccagtctgaatgtaacgtgttcaccaaaccgggatattcgtcaaacacgagcatgtgatccgaccacggtatgct
gtgggcgatcaattttagttcttgcacgcggccttcgcgtaagcaatacaaaatgagcgcgtcgctgatcttgacacagtcttgcatgtacgcggacaaattaacgttttccatacagctcacattgtttattag
cgccgtgttcaagtgtttgtatttggacacataatcgtagttgatgtactgtttaatgggttcttgaaaccattcttttagtagtatgtgactggccactatgcgtttccaatttaatttgtgtgcgtatttttgctgcac
cgacaacgagaggttattgtaatttttggatatttcttccatgtccaacaagtccccaaacgcgagtataaaatcttgcgtcaaaaatttttgctcagacaccaacgaccagatcaaatgtgatttaaacctgt
tggcgattgttatcgacaacggcgaaattgaaataattttccaatccaacttgttgcgaaacacgtgaataaaatcgacgcgtccgtaacattcgcgcgatatgcgcttccaaaacgtgtcatcttgcaaatt
aagcaaatagacacgattgttgggagatttgacggccaattcaattatttttatatattctttttgctttaaagcgcgttgtagcacttgggttggagccatgtcgactgaagctccacgctgtttgaagcaaggt
gaccgttttggtcggcatgttcaaacgtcgattacatgtttgctttgcatcaaaatggcgtaattaattaagaaacaacatgaaagccatctgcatcattagcggcgatgttcatggaaaaatttattttcaaca
agaatcagcgaatcaaccgcttaaaattagcggctatttgttaaatttgcctcgaggtttgcacggctttcacgtgcacgaatatggcgacacgagcaacggttgcacgtcggccggtgagcactttaatcc
caccaatgaggaccacggcgctcccgatgctgaaattaggcatgttggcgacttgggcaacataaaatcggctggctacaattcactgaccgaagtaaacatgatggacaacgttatgtctctatatggc
ccgcataatattatcggaagaagtttggtcgtgcacacggacaaagacgatttgggccttaccgatcatccgttgagcaaaacaaccggcaattctggcggccgtttgggatgcggaataattgccatatgt
aaatgatgtcatcgttctaactcgctttacgagtagaattctacgtgtaaaacataatcaagagatgatgtcatttgtttttcaaaactgaactcaagaaatgatgtcatttgtttttcaaaactgaactggcttta
cgagtagaattctacttgtaacgcatgatcaagggatgatgtcatttgtttttcaaaaccgaactcgctttacgagtagaattctacttgtaaaacataatcgaaagatgatgtcatttgtttttttaaaattgaact
ggctttacgagtagaattctacttgtaaaacacaatcgagagatgatgtcatattttgcacacggctctaattaaactcgctttacgagtaaaattctacttgtaacgcatgatcaagggatgatgtattggatg
agtcatttgtttttcaaaactaaactcgctttacgagtagaattctacttgtaacgcacgcccaagggatgatgtcatttatttgtgcaaagctgatgtcatcttttgcacacgattataaacacaatcaaataat
gactcatttgtttttcaaaactgaactcgctttacgagtagaattctacttgtaaaacacaatcaagcgatgatgtcattttaaaaatgatgtcatttgtttttcaaaactaaactcgctttacgagtagaattctac
gtgtaaaacacaatcaagggatgatgtcatttactaaaataaaataattatttaaataaaaatgtttttattgtaaaatacacattgattacacgtgacatttacgatggcgaacaataatttcactttttatatta
ggacacgacgtgtatataggaaagcttaagcgtttcaataaagccatggcgtacacgctaagcttgcccagcttgcggctctttgaaatctgtagttttcggggagtaccgtcgttcttcagtgccacatacgt
caacttgcgatcgtacactttataatacgtgttgtagttatttttttccagaaattccctcataaagcaatccttggataaagtttttgatccgtacagttggccacaccggtccatgcacaggtacacacacgtg
atggcgttttgaatgacgatgcgatttctgtcaacggcaacgcgcttgaatatggtgtcgacgttgtccgattcaatggttccgtaaacagctccgtctggatttactgccaaaaactgccggttaataaacag
ctggccgggaatagacgtgcccgtgatgtgtgtcagcagagctgagcagtcagccatagaggctagagctacaagtgccagcaagcgatacatgatgaactttaagtccccacagcaaactggcgcttt
tatataaaaatttgggccatttttggcgattagataatttttgaagattagataatattgagattagttaataatttgtgtgattagataactttttagggtattgcgcattataaatcaaggtcgagttgtataaactg
ctctggcgtgtaaaactgcagacttaagttttttgcaaacactcggtctgaatcgctaaaatctttctgaccggtggttagattaattcggccagccgcgtcgcccacataaaaagattgttccttgtcaatatg
cgtaaactgtttggccatctcgcgccacattcccgtgtcgggctttcgatgctcatccttgttgggcgacacataaaacgatatgggcacgccagtagctttttaatattctctaatttatataataaatcgctc
gctttgattttgccggaacctaaatgggcttggttcgtaaaaacaactaaatcgtagcctaattcgtacaaacgctttagcttgtgtgcgcacggaaggagctgccagtcgtctgggtttttggaaatttggac
cgtgtctttgagctaattagcgtgccgtccaaatcaaaagccgcaattttggttcttttagcgccgtcatgaaccgcgtacgcatacaaatcgggctgctgtaacgtccacatggtgaatgcatcttactcaa
agtccatcaattcgtacgcgtttgtgtccaggtcgggcgttgaaaaattgtagcttgccattagatcggatagcgattcaaattttgtaagcgtttgtagcgcacgtttggcatcttgtttaaaattacacgacga
cagacagtaaaaatattcctcgataagcatgactacacccatatcactgtttaagtgctcgacgtagttgttgcatgttatgtcgcgtgtgccgcgatacgcgtgatttcggtgaaaatcacaccacaaccag
tcggcgtgcgtgtaacaaagtcgacagcgaaacaatttatcgttttccaaaaaatttaaatactcgacagttttgcagcttagattccgcgtttgattcaccttaaaatcgtcgtcagcctctataatctcgggc
aacagcttgccttgttgccccatcgtatcgatcacctcccccaagtggcccggtgttatattaagtcgtttaaaatcatttattgcttcctgcacgtcggcctggtaatttttgaccacgggcgtggaaatcaatt
gccgttgaagggaaataattcgtggtgtgggtatcggccgcctgttgcacaattccaccagcggtggaggcaagggcgcattcacagcaaccgttgtcatttataagtaatagtgtaaaaatgcaaatattc
atcaaaacattgacgggcaaaaccattaccgccgaaacggaacccgcagagacggtggccgatcttaagcaaaaaattgccgataaagaaggtgtgcccgtagatcaacaaagacttatctttgcgg
gcaaacaactggaagattccaaaactatggccgattacaatattcagaaggaatctactcttcacatggtgttacgattacgaggagggtattaataataacaataataaaaaccattaaatatacataaa
agtttttatttaatctgacatatttgtatcttgtgtattatcgctaaccattaaaagtgctggagccacagtgttgcggcgagtctttatagaagatcgttgtttggctggaactgagcttttcctttcctgctgccg
ctaatgggagtgggcacgtactctgtagtagacggtgcaacgggcaacttgagcgctaccgtcttaaatttggccatacttttagtgatgaaatcgcgcgttaacacttcgtcgtaaatgttacttagcagag
gcgcaacattgtgattaaatgtctcgtttaacaagctgtaaaactccgaataaagcttatcgcgcatttcgcagctctccttcaattctgccaaatttgcgttggtaagcaccacagtctgtctttttttgctcgct
ggaattgctgcgttctcgcttgaagacgacgatgtcgatcggtcggccatttttttgcccagcttttcagtgtgatcaaaaatgaacacaaaatctgccaattcgggcttgtttttcaccaaatcccacatggcc
gggctactaggccactcgggctgcttgatcttagtgtaccaactgttaaacaaaatgtatttattgttgttaatcactttcttcttgcgtttggacattttgcgttcgtcttgcatgacaggcaccacgttaaggata
tagttaatgttctttctttccaagaaatttacaataacggccagctggtccatgttggatttgttgtaagagctcgattccagtttattcaacagcttttcatttttgcacacggccgcagtctccggagattgttgct
ccggcacgtttaccatgtttgcttcttgtaaacctttgaaacaacccgtttgtattcttgatgatatattttttaatgcccaacaacctggcaattcgtttgtgatgaagacacaccttacgcttcgaacatttgtc
ggtgattactgtgaaatggcctaaattagctcttatatattcttttatacgctcaaacgacacgatgtccaacatgtgcgcgcagacgttttctgtgttcatcgtgtgcttgagcgtgttgatggcttccctgaaca
gcgcttgtatttcgctgcgagtcaagcagtccgaatcacacccgcctaagtgcgtgcaatttttggggggcatcgttgtctatctttttcagagtggcgtagaaaaagtcctgcaattgcctattatcaaaacg
cgccttgacgctgcgcacaaaatcaaaaaattcaatgtaattgctgtaatcgtacgtgatcagttgtttgtcgttcatataattaaagtatttgttgagcggcacgatggccaggctgcgcgctatttcgcaatt
gaagcgtcgcggttttaacattatacggtagtcattgccaaacgtgcccggcaacaacttcacggtgtacgtgttgggtttggcgttcacgttaatcaagttgccgcgcacgacgcctacgtatatcaaatac
ttgtaggtgacgccgtcatctttccattgtaacgtaaatggcaacttgtagatgaacgcgctgtcaaaaaaccggccagtttcttccacaaactcgcgcacggctgtctcgtaaacttttgcgtcgcaacaat
cgcgatgacctcgtggtatggaaatttttctaaaaaagtgtcgttcatgtcggcggcgggcgcgttcgcgctccggtacgcgcgacgggcacacagcaggacagccttgtccggctcgattatcataaac
aatcctgcagcgtttcgcattttacatatttgacacttaaaaaattgcgcacacgagcaccatcgtttgatacctaattgcaactatttacaatttatcagtttacgttgaacccgttttaatttttagatccgtcc
ttgttcagttgcaagttgactaaatgacaaaatttttcggttctgcaaaaccgcccttgtctgttccacccgttgtatttgaaaaaactttttttcacgcggcgacaactgcttgtataatattgcccaatgtaaac
atgcaaaattttgttactctcgtcaaaacagcggttggcgttccattccataatttttttattatttatcaacgatggccattgtaaattgtcgtcatttatacgcatcatatgatttaacaaaagcttttcgtatagcg
gaacttcaattcccttggaacatttttcaaacgataatttaatttgtttctcggttggcagcatttcatgcttgattaacaatcgcctgacttttatagccacgtttatgtctttgcacagcaaatgtgggttgtcgac

aatgtaatagtgcaaagcatttgttacggcaaatgcgtagtttgatttgacgacgcccttttcttgacgggcattgcggcttttaaaattacttgcaagcattgtacgaatacctcttgtgtttaaacaataatat
ggacaaacatcggcgaaacaatttgtaataattatgaaatcccaaattgcaggttttaaacttctttgttacttgttttataataaataaaatttgctgacccatgtctgcgcccacaactttaattaaccatttgt
gcgcatattgattgtctcgttgttcccaaccggaaaattgattgatctcgagccaccggcattggtcgtttgataccgtcgttaacgccgacgctcctgcctgtttgattacgggttctaaaagacgaaacagc
agcgtaaatttgtttttgcgtcggtagtattttggcaggcaataatcaaaaaaatccgtaagcaattctctgcatctattaatattcgttgcgtacgaatcgagtttttcaaaaattacttgtttgtatgaaaataa
cgtttgggcttctcacaataataatcttcgttgtagaacagaaacggtttgcgagaattggcacgtttgtccatgattggctcagtgtaacgattgattcaaatcaaaattgacaacacgtttgccgtaatgtgc
accggttcgcacacgtttgccgcgtatgtaatccatgtttatttcgctgtcgcaattgattacacgattgtgttgggcggcgcgttttattgaatttaggcgacgcgtcgacaactccaaaggattgtaaagcgc
agatttttccagagtaaacgagtttaagtggccaccgttgaaccattccagagccacgattgtgtacagcaaaaagaatatttcttgtcgacgttttcaaacgcaaacttgtttttaggcaatagtagtaaaa
ttttaacgaattgtataaataaaacataaaattgccatttttaaagtaaaattctacatccgtgacgaacaaaaggtttactattttgttctccaacaagtgtgccaattttcttaagtacaccattgaatttttgtc
gtcgtccatctcgatcaacaacacgtacggcgttttggaatttaaaattattctaaaattttcctgttgcaacgattccacagcgtccgaccaatatgacgctgccacctctagacagatgtatttcttggaaa
acacgtgtcgtttgataacctcgctgatggacgtgatcgattgtaaatacttttcaaacgtcgcgtcttcccaaccacgcaccgaaacgggcgctgtcgtgtcgggctgatgtttgaaatccaaaccactctg
aattaacttggttgtgattcgtatgctcaactgttgacccaacgtgtagtgatcttcgtaggcgcgctcccacatcacgttacacacaaatttgacgagatcatcaacgtctttctgttgcaaaattcgccgca
aacgcgccacatcgcccttgtaccaccgatctcggcacacaagctgtagcatttttaaatcgtgatcgctcaagctattaattctggttagatttatatagtcgtcaatatcctcgggcgtggtttgcgtcatgt
ctgtaaaacgtgcaaaatcaaacatttttatgttgtagtcgaatctaacaaatccatcggcgttcacttgcacttcgcgctttacaaaacgaggtagcgtgtaatcgaacccgtttaaatagattgcgtacaa
aaccagcacttcatcttccagtttgcacgcttgcggcaaaaattgtgtggtgtgctccaaccgggtgacaaacatgactatggaaaataacgcggaattcaacagacgactagagtacgtgggcacgatc
gccacaatgatgaaacgaacattgaacgttttacgacagcagggctattgcacgcaacaggatgcggattctttgtgcgtgtcagacgacacggcggcctggttatgcggccgtttgccgacctgcaatttt
gtatcgttccgcgtgcacatcgaccagtttgagcatccaaatccggcgttggaatattttaaatttgaagaaagtctggcgcaacgccaacacgtgggcccgcgttacacgtacatgaattacacgcttttta
aaaacgtcgtggccctcaaattggtcgtgtacacgcgcacgctacaagctaacatgtacgcggacgggttgccgtattttgtgcaaaatttttcagaaacaagctacaaacatgttcgtgtgtatgttagaaa
acttggtgcgatacaagtagcgacattatcagtttacgaacaaattattgaagatacaataaatgaactcgtcgtcaatcacgttgattagataatgtccgtgttaaatgtgatatcttagattacgagcgcgc
aataaccatagtttaatcgaagagaatagccgtcgccacaatggataattacaaattgcaattgcaagaatttttttgaccaagcgcccgacaacgacgatcccaactttgaacatcaaacgcccaatcta
ttggcgcatcagaaaaaaaggcatacagtggatgattaacagagaaaaaaacggccggcccaacggcggcgtgcttgccgacgacatgggactcggcaaaacgctctctgtgctaatgttaatcgcaaa
aaacaactctctacaattgaaaactctaatagtgtgtcctttgtctttaatcaatcattgggtaaccgaaaacaagaagcatgatttaaattttaacattttaaagtattacaaatctttggatgccgacacggtt
gagcattaccacattgtggtgaccacgtacgacgtttttattggcacatttcaaattgatcaaacaaaataaacagtcaagtctgttttcaacccgctggcatcgagttgttctagatgaagcgcatattatcaa
aaactgcaagacgggcgtgcacaacgccgcgtgcgctttgaccgcaacaaaccgatggtgcattaccggcacaccgatccacaacaagcattgggacatgtactcgatgattaatttttttgcaatgtcgt
ccttttaacaatccaagagtgtggaaaatgttaaataaaaacaacgactctacaaatcgcataaaaagtattattaaaaaaaattgttttaaaacgcgacaaatctgaaatttcttctaacattcctaaacaca
cggttgagtatgtacatgttaattttaatgaagaagaaaaaacgttgtacgataaattaaagtgtgaatcggaagaggcgtatgtgaaggctgtggcagcgcgtgaaaacgaaaacgcactaagccgattg
cagcaaatgcagcacgtgttatggctaatactgaaattgaggcaaatctgctgccacccgtatttggccatgcacggtaaaaatattttggaaacaaacgactgttttaaaatggattatatgagcagcaag
tgcaaacgagtgctcgacttggtagacgacattttgaacacaagcaacgacaagataatattggtttcgcaatgggtggaatatttaaaaatatttgaaaactttttaaacaaaaaaacattgctacgttaat
gtacacgggccaattaaaagtggaagacaggattttggccgagacgacattcaatgatgctgccaatactcaacatcgaattttgctgctttccattaagtgcggcggcgtcgggttaaacttaataggcgg
aaaccacattgtaatgttggagcctcattggaacccgcaaattgaattgcaggcgcaagaccgaatcagtcgtatgggacaaacaaaaaacacgtacgtgtacaagatgctaaatgtggaagacaaca
gcatcgaaaaatacattaaacaacgccaagacaaaaagattgcgtttgtcaacacggtctttgaagagactctgctcaattacgaagacattaaaaaatttttcaacttgtagctggtaagtcgtcatgaac
acccgatatgctacttgctatgtttgcgacgagttggtgtacttgtttaagaaaacgtttagtaacatgtccccttcggccgctgcgttttaccaacggcgcatggccattgttaaaaacggtatcgtgctgtgc
ccacgttgttcgtcggaactaaaaattggcaacggcgtttcgattccaatttacccccaccgcgctcaacaacatgcacgacggtcgcgttaagacgcaagcgcttcgagttttggcccgctcgctacctc
cgctgtacgactcgaccgtcgatcgacacggctgcaaggtgttcacggtgcggcgctacaacagacgcgtaatcgactttgcgggcattcgcaacaaaacgctggaaatcattaaaacggatagaaac
ttgccgctcaacacagaatgcaatgtgaaagttgtcgacagtgcatgcatgcgttgcagaaaaagtttcgcagtttaccccgccgttacctatctgcattgcggacattcgtgtctgtgcaccgactgcgac
gaaacggtaaacgtggacaacacgtgtcctaaatgtaaaagcggcattagatataaattaaaatacaaaactttgtaacatgttgccctacgaaatggtgattgccgtgttggtttacttgtcgccggcgca
gattctaaatttaaaccttccttttgcataccaaaaaagtgtgctgtttgccagcaactctgcaaaagttaacgaacgcatcaggcggcgagcgcgtgacgacaacgacgacgacccctatttttactaca
aacagttcataaagattaatttttaactaaaaaaataataaatgtttataataaaactgaaaagtgtattagagcgacgtttgatggtcggtatgtggttacacgcgacgttttaatgtgctttgtaaacaagag
ttatatgaagcaattgctgcgcgaggttgacactcgcattacactacagcaacttgttaaaatgtatagtccagaatttggtttttatgtaaatagcaaaattatgtttgtgttaactgaatcggtgttggcgtctat
ttgtttaaaacactcgttcggcaaatgcgagtggttggacaaaaatataaaaactgtgtgtttacaattaagaaaaatttgtattaataataagcaacattcgacatgtctatcgtattgattattgtcatagttgt
aatatttttaatatgtttttgtacctatcaaatagcaataataaaaatgatgccaataaaaacaatgcttttattgatctcaatcccttgccgctcaatgctacaaccgctactactaccactgccgttgctacc
accactaccaacaacaacaacagcatagtggcctttcggcaaaacaacattcaagaactacaaaactttgaacgatggttcaaaaataatctctcatattcgtttagccaaaaagctgaaaaggtggta
aatcccaatagaaattggaacgacaacacggtatttgacaatttgagtccgtggacaagcgttccggactttggtaccgtgtgccacacgctcatagggtattgcgtacgctacaacaacaccagcgaca
cgttataccagaaccctgaattggcttacaatctcattaacgggctgcgcatcatttgcagcaaactgcccgatccgccgccgcaccaacaagcgccctggggcccggtcgccgattggtaccatttcac
aatcacaatgcccgaggtgtttatgaacattaccattgtgctaaacgaaacgcagcattacgacgaagctgcgtccctcacgcgttactggctcggcttgtatctgcccacggccgtcaactcgatgggct
ggcaccggacggcaggcaactcaatgcgcatgggtgtgccctacacgtacagtcaaatcttgcgcggatattcattggcgcaaattaggcaagagcagggaatacaagaaatcctaaacacgatcgcg
tttccgtacgtgactcaaggcaacggcttgcacgtcgattcgatatacatcgatcacattgacgtgcgcgcttacggctatttgataaattcatactttacgtttgcctattacacgtactattttggagacgagg
taatcaacacggtgggtttgacgagagccatcgaaaacgtgggcagtcccgagggagttgtggtgccaggcgtcatgtctcgaaacggcacgttgtactctaacgtgataggcaactttattacgtatccgt
tggccgtccattcggccgattactccaaagtgttgaccaaactttcaaaaacatattacggttcggttgtgggcgtaacgaataggttggcttactacgaatccgatcccacaaacaacattcaagcgccc
ctgtggaccatggcgcggcgcatttggaatcggcgcggcagaattatcaactataatgccaacacggtgtcgtttgagtcgggtattattttgcaaagtttgaacggaatcatgcgcatcccgtcgggcacc
acgtccacgcagtcgttcagaccgaccattggccaaacggctatagccaaaaccgacacggccggcgccattttggtgtacgccaagtttgcggaaatgaacaatttgcaatttaaatcgtgcacgttgtt
ctacgatcacggcatgttccagctatattacaacattggcgtggaaccaaactcgctcaacaacacaaacgggcgggtgattgtgctaagcagagacacgtcggtcaacaccaacgatttgtcatttgaa
gcgcaaagaattaacaacaacaactcgtcggaaggcaccacgttcaacggtgtggtctgtcatcgcgttcctatcacaaacatcaacgtgccttctctgaccgttcgaagtcccaattctagcgtcgaact
agtcgagcagataattagttttcaaacaatgtacacggccacggcttcggcctgttacaaattaaacgtcgaaggtcattcggattccctgagagcttttagagttaattccgacgaaaacatttatgtaaac
gtgggcaacggcgttaaagccctgtttaattatccctgggtaatggtcaaagaaaataacaaagtgtctttcatgtcggctaacgaagacactactataccatttagcgttataatgaattccttcacctctat
cggcgaaccagctttgcaatactctccatcaaattgctttgtgtatggaaacggtttcaaattgaacaacagcacgtttgatttacaatttatttttgaaattgtgtaattatatttagggagaatgtgatattcaaa
agactgactgttaacacaaaagactgatattgttgttgttacaaaatagataataaaacaaaaaataaattaaatattatttatttattaaactgtttaattttaatgctaacgcgtacaaatcacgctgttccga
cgtggacatggaattgcgcagaaaagtcttgatagtgtcgatttcttcgccgtcatccacttccatatatttgatttcttcctcgatttgcatttccaagtttgcgtattcttgcaaataataatctagtcgttgggcg
acctcgccaattttaaataatacattatccgacaccaaatgccagcgagtgactgtgcgctccatcatcctggcactttttaatgtgaatattaaaaggttgttgcatatatatcgttaaacgtttatgtttacttt
cacgttagctcgtttcattgatgtaaacatttagttttataacagcgtcggtaattttatttttaaagtaaacagaccaaaatcaaaggtgtcttcgacaggtacgattattttcccattgacactgttttcgtgca
cagatataattttatcaccgtttattatttgcccaaacacacgtactcgtttcttctcaagccaactatttctaaacaattcacttttctattatcgtgtacgcaattaaaagtaaacgaagcgctacaattgtcgt
attctattacaattctgcggcatttataaaatttattaatgttgacgcaaattccatgcagcgcatccatttcgtactgcaaatgcggcgcaattaaaaaatttcctcgtcgttgttaacaatcttgggcgctaaa
aagcacgccaacacgcccacgtctttaatgcaatattccaatttgaacggcagttcctcggacatgtatattgtcacggtgggcgccaaaggagcggctttagcaaaatgacacaagtaatcgcccgcaa
aagtgtgcgttacggtttgctttgctttgagaacggaaaagttttcgttgtccgcgctcatctgcacgtccgccgagccaatgtcgccatttgctctaaactgcagacccttcttggaacacgacacaataata
tcgtggtcgaattgcgtcatgtctttgcacacctgcgcaaactcgacgctcgacatgtggacgacgcaatcgtaatcgctatccggaattcccaaatgttccacgtcgatgcacatcaacttgagcgtgtac
gtgcagattctattgtcgttgttgaacacgaacgccatcacatcgccctgatcttccgctttcatcagtacagagctgcgctcgttaacgcatttgacaattttacttaaactgtttatggacacgttgagcggc
acgttgcggtcacatctatatttttgaaaccctcggcgtgtagttgcaacgacacgagcgcgacatgcgaggtgtccataacctgcatgcttacgcctcgattatcacaatcaaaagtagcgtgcggcagc
agatccttaaaagtttccaccagcctcttcaaaactgcgccggttttaaattccgcttcgaacatttttagcagtgattctaattgcagctgctctttgatacaactaattttacgacgacgatgcgagcttttatt

caaccgagcgtgcatgtttgcaatcgtgcaagcgttatcaatttttcattatcgtattgttgcacatcaacaggctggacaccacgttgaactcgccgcagttttgcggcaagttggacccgccgcgcatcca
atgcaaactttccgacattctgttgcctacgaacgattgattcttgtccattgatcgaagcgagtgccttcgactttttcgtgtccagtgtggcttgttttaataaattctttgaaaatattgtcgggtgtattattaa
atagcatgtatggtatgttgaagatgggataacgcttggcgtgcgggtcgtcatgatttccaccgcgcaccacatatttgcgctcaattttatcaaaattggactggcgagacaaaaacgagacgggcgaca
ggcatatttgggcgtgcgtaccatcttcggccatccactcggtcaggtcttcgctgcggttaaacacacctttctgaccgtgaatgccacatatttttattccttccaaatcgttggtggacgtgactatgactat
tttaagcataacgttgtcgccgttaaccaccatgctggcgtcgagtttttcaatttttttgattttaatttgtctaaagtaaacgtacactttgtaaacgttaaaattgccgttggtgcacgtttcaattttgtaccgtc
ggccgtcgtacacccaattaatctttgcgttgctcaccaacacaccggccatgtacagcacaagtccgtcgtctagcgcaacgtaatttttgtcgctactattcgtaaactttactaaacacgactgcttggg
gccgaccacaagcttgcccttcaatttgttcactttgttgttgtataaacaaatgggcagcgcaatgtgcggaatgtacggatcttcggcggtcatgagtttattgtctcgcaccaacgtccacaatttaaacat
tttattgttgagcaaaatggacttgtttaccgccacagagtagccatttggtaaacccgatacgcaattttcctctttgtactcaaacacgggcatggcattctttagattggttagggacacaatcaatttgggt
acgggcgtggtatgaaataaatgtataaaattacgataataatactgctccaacttggacatgagcgatttgacgtcatcgttttctacgatcgtacactgaataatgggattatagtatatagaatgtttatagt
ggtattcgtagggtgtcaacaatacgttaatgtcggcttcgttgttcacccgcaacttttttttgatgcatatcattccttcgtgatgattaacgtaaagtattctgtctgtaatcttcaattcgatgggcgccatgttt
cttttcatagtgtacacgataaacgacgtgtttgattttaaacattttaaatttgtgggtctatcattaaacgcgatcagcaacgagtcgtcttgaacgtcgttgaggtcgtccacgaacgcgaccagattgtgtt
ttagcaaatattgaaatttttgcgcaaccatttcgtagtccacgttgggcaaacatgcgttgcggcaaaggaaaaacttttttgcccgcccacggtcatttcgccgtgaaaaaaaactgccaataaatttcacaa
aatcctttttttgcttcaacattttctggcgcatgctgtcgttggtgattcgcgccacctcgttgccgacgcgatatttttaacacgggcaacgaaatttcaatattgttattgctgctgttgtcctgttgattgggaaa
gactttgcgttgcttgctaaaagttttcgatacgcaatatatgagacgcccgttgactatacaatcgacaatctttttcgactctttgttgtacaagacgctttgaatttttacgacgcttgttcgccaccgtgtacg
cgtcgtcgtcggccgtcttgtcgagaactcgttgatagttttgcaaaattgtcgaagttaataacagttcttatcaaataggcgtgcttgtatacaattttgttggccaaactgtctatagaatagtttatgtcgtgat
tcataataatttttatgtgttccacgagttgttgcttgtgaagcgtgttgtattcgaagagaaaatcgagcggtttccatttgccgctgttggccagatatgtttccagcacagaatttaaatcttccgtcactacgt
aatcgctagcgtacacgtctcgagcaaacaggacgtcgtcttgtttgtcgtaaactagttggattgcgcgattgatgtgcttctcttgatccacgttgccgtacaaaaacatgcgtttgcaatgtttggcgtata
gcttgtcgtagaaattgtgcaccaaaacgttgttgttcatcattatgttgggaaaactcaaaaatctgccgtccagcataaaagttccgttaatattgttgtttgcgtcgacatcgtccgttctctaaattgcttgt
ctaagcgcgtgccgaatataacgggcacacatttatgcattacgcaactgagctgttcattaagagcgcaacacaaataagacttgcgttcttgaatagcgcaaaaaagcatacgttcattgctgtttgtag
cgcaatcaaaagtatattttaatttgtatttattttcaattctatcgtacaactcgttgaaatcttgaaccacgtccgtcatcgtgaagcgattactgcgcactaattatgtctaaacgtgttcgtgaacggtcggtt
gtttcggatgaaacggccaaacgcattcgacaaaacgaacactgtcatgccaaaaatgaatctttttggggttttgcaacttggaagaaattgattattatcaatgtttaaaaatgcaatacgttccggacc
aaaagtttgacaacgattttattttaacagtgtacagaatggccaacgtggtgacgaaacaagttagaccgtataacagtatcgacgaaaagcaccattacaacacggtgcgtaacgtgttgattttaataa
aaaatgcgcgtttagtgcttagtaatagtgtcaaaaagcaatactatgacgatgtgttaaaattgaaaaaaaatacagacttggaatcgtacgatccattgattacggtctttttacaaattggcgaatctgta
aatgaagaaatacaaaaactcagaaaagctttggtcaatatttttactaataaacccgacaagtcggatataaacaacccagatgtagtttcgtatcaatttatttttggcagagtacaaaaattgtataaca
gggcaattaaacaaaaaactaaaactataattgtaaaacgtcctacaactatgaacagaattcaaatagattggaaaactctttccgaagacgaacaaaaaatgactagacaagaaattgccgaaaaa
attgtaaagccttgttttgagcaatttggcactatattacacatatacgtatgtcctttaaaacacaaccgaattattgtcgagtatgcaaactcagagtcggtacaaaaagccatgactgtaaatgacgaca
ctcgatttacagttacagagttttccgtggttcagtactacaacgtggccaaaacagaaatggtgaaccagcgaattgacataataagcaaggacattgaggatttaagaaacgctttaaaatcttacacat
aaattaaaatatcgaacaaaggaaaaaaacaattgtaacaaaaataatttacattaaaatttacaagttttttctagtgtcgtactttttacaatgcgtctgttgtccgtcgagcattgcaaacatattgtgga
cggcgcaaaatagcaaacaaaaggcacgtccgcgctctcccacgctattctaaaacgatgaatccatattaatttttcattgtcgccaaacgtcgctccgctgcctccttccaataacaaatactcagaa
acacaaacatgtacaattgctgtcgcggcgttaattgtcgctgtttttccaaatagtctattatgggaaacaaacacttgtcacaacacaaatactcgttaattgtcacaaccgacaagcacatttggcaaaa
tgcgtcgcaatttttgtacggacgagattctatgcgaagttcgttgtccatgacgtcttgggtccactttttcaacaagacactttatatttgtgatttgtacaactttggtacgtgttagagtgtttttgataagcttt
gataagtttaaaactgttggagtaaggccacgtcattatgttctgcaccttttgtttaaaagacagaaattactatatgttcaaactatttaaagattattggccaacgtgcacgacagaatgccagatatgtctt
gagaaaattgacgataacgggggcatagtggcaatgcccgacactggcatgttaaacttggaaaagatgtttcacgaacaatgtattcagcgttggcgtcgcgaacatactcgagatccctttaatcgtgtt
ataaaatattattttaactttcccccaaaaacactagaggagtgcaacgtgatgcttcgagaaactaaagggtctataggcgatcacgaaattgatcgcgtttacaaacgcgtttatcaacgcgttacacag
gaagacgccctggacattgaactcgattttaggcatttttttaaaatgcaatcatgacgaacgtatggttcgcgacggacgtcaacctgatcaattgtgtactgaaagataatttatttttgatagataataatta
cattattttaaatgtgttcgaccaagaaaccgatcaagttagacctctgtgcctcggtgaaattaacgcccttcaaaccgatgcggccgcccaagccgatgcaatgctggatacatcctcgacgagcgaat
tgcaaagtaacgcgtccacgtaacaattattcagatcccgataacgaaaacgacatgttgcacatgaccgtgttaaacagcgtgtttttgaacgagcacgcgaaattgtattatcggcacttgttgcgcaac
gatcaagccgaggcgagaaaaacaattctcaacgccgacagcgtgtacgagtgcatgttaattagaccaattcgtacggaacatttttagaagcgtcgacgaggctggcgaacacaacatgagcgtttta
aagatcatcatcgatgcggtcatcaagtacattggcaaactggccgacgacgagtacattttgatagcggaccgcatgtatgtcgatttaatctattccgaatttagggccattattttgcctcaaagcgcgta
cattatcaaaggagattacgcagaaagcgatagtgaaagcgggcaaagtgtcgacgtttgtaatgaactcgaatatccttggaaattaattacggcgaacaattgtattgtttctacggacgagtcacgtca
gtcgcaatacatttatcgcacttttctttgtacaatacagtcttgaccgcaattcttaaacaaaacaatccattcgacgtaattgccgaaaatacttctatttcaattatagtcaggaatttgggcagctgtcca
aacaataaagatcgggtaaagtgctgcgatcttaattacggcggcgtcccgccgggacatgtcatgtgcccgccgcgtgagatcaccaaaaaattttttcattacgcaaagtgggttcgaaatcccaaca
agtacaaacgatacagcgagttaatcgcgcgccaatcagaaaccggcggcggatctgcgagtttacgcgaaaacgtaaacaaccagctacacgctcgagatgtgtctcaattacatttattggattggg
aaaactttatgggtgaattcagcagttattttggtctgcacgcacacaacgtgtagcatcgccagtatttaacagctgacctatttgttaaacaagcattcttatctcaataattggtccgacgtggtgacaattg
tatccacaatcatgaaaaaagtagcgcttggaaaaattatcgaaaacacagtagaaagcaaatataaaagcaacagtgtgtcgtcgtcattgtcaacgggcgccagtgcaaaattgagtttaagcgaata
ttacaaaacttttgaagcaaataaagtgggccagcacactacgtacgacgtggtcggcaagcgagattacacgaaatttgacaaattggtgaaaaaaatattgacatgctgcgatcaatcatgcgacgtttc
aagagtacaaacaatctcagcaaaaaaccctccgattattatgtagtgttatgtccaaagtgttattttgtgacgtcggccgaagtgagcgtggctgaatacatagaaatgcataaaaattttaacacgaaat
tcgccgatcggtgccctaacgatttttattgtgaccaactctaaaagttggaataatcatgaaaattgttctgccctattttaccctctgtgttaataaagtttgttgtttgtattttgtggttttatttatttacgctagat
attgggtttaaggttcttagaaatagagttgtattttccctaccaaaagggatttgagcttcatataaatacaatattcgctcgacaagcggtttatttcactcggaggtattatatcaggcagtcgaacgtgcgc
gatgaaacatcccgtttacgctagatatttggagtttgatgatgtagtgttagatttgactagtttaatatttttagagtttgataacgctcaaaatgaagagtacattattttttatgaatgtaaaaaaggcgtttttac
aaaaactttcacattacttgtgatctgtcgcttgaaacgctgaccgtgttggtgtacgaaaaagctcgcctaattgtgaaacaaatggagtttgagcagccgccaaactttgttaattttatcagtttcaacgcg
accgacaacgacaactccatgataatagacttgtgttccgacgcgcgcataatcgtggccaagaagctgacgcccgacgaaacgtatcatcagcgcgtgtccggattttttggattttcaaaaacgtaact
gcatacctcggcccccaatcgagtcggacccaaaagtgcgagacgccttggatcgtgaactagaaataaaactatacaagtagaaaaaaattaatttattaatagttgtaataattatcttcgtcctcatctt
cgctggtgtcataatgcggtggtgtgtttgtgttttgttttaatcgtttgcgcgtcgacaccacttcgccgataggaaatttttttggatttcgcattaaatgccctcttagcgacgcgccgtttacgactactaaaca
tgttgacgcgctcgtcgtcttcagtgtcataatccgtgctagtgttttcgttgttatttctatgagacgatcgtttgatttagtttcgtagaattgtccgcgttatcgtcgctttcgtcgatgtcgtccctaactatctc
gtaggcggctttgcgcggaatccaagatttttgcaatgtatctattttaacgtactttttcttcgagcgcttttctagctttatgcatagcaatgtcttcgtcgccgccgttcattttatgatactttgtaaacgtctcgac
gaataactttttggcgcgaggaggcatttttttcattgtataacatatcgggaatttgatacattgtaattagaattaagcaagttcgtcttcggttgtactgtattcggtttctgtatctgtagtggaatcctctgtact
agtagtagtgtcgctattgttggcgtcaggccttggctgccatttaccgtctatcaacatgtatttttttcctaacagcacaacatgctagcttggtagctatctgtgtcgacttatattttttgtaaactacgatcgta
gaatttttcaaatatcctcttaccgttatagggaaggttttgataatatttaggcaacatatcaataaaaagacaatataaaaactttgtgtttgtgtttttatttatcacataaaatggacgtctggcaagaatcaca
accaatattagtgtttttttttcttacattacgagattcaacttgatactaaaattaattattaattaaattaaattaaattttgaagcatttttttcgctatcgttttcagactcaaaattatcgacgctatcgctatgaaa
agcgtaatatttgttggctttgagatattctatattttgctcatttttaacaataaacacgcgactcttttcgtcgcgtctcaccataacaccgtttttacaaatggaaatgtatttgtaaaacggcaacagagcgt
cgcgagtttttttaagtaacagcttttgctccgctgtggcggccacaaatatttttacgggcccgtcgtaattaatgtttaaattaaaatttttaagtcgacgctcgcgcgacttggtttgccattctttagcgcgcg
tcgcgtcacacagcttggccacaatgtggtttttgtcaaacgaagattctatgacgtgtttaaagtttaggtcgagtaaagcgcaaatctttttttaaataatagtttctaatttttttattattcagcctgctgtcgtg
aataccgtatatctcaacgctgtctgtgagattgtcgtattctagcctttttagtttttcgctcatcgacttgatattgtccgacacattttcgtcgatttgcgttttgatcaacgacttgagcagagacacgttaatc
aactgttcaaattgatccatattaactatatcaacccgatgcgtatatggtgcgtaaaatatatttttttaaccctcttatactttgcactctgcgttaatacgcgttcgtgtacagacgtaatcatgttttctttttttgg
ataaaactcctactgagtttgacctcatattagaccctcacaagttgcaaaacgtggcatttttttaccaatgaagaatttaaagttattttaaaaaaatttcatcacagatttaaagaagaaccaaaaattaaatt
atttcaacagtttaatcgaccaattaatcaacgtgtacacagacgcgtcggtgaaaaacacgcagcccgacgtgttggctaaaattatcaaatcaacttgtgttatagtcacagatttgccgtccaacgtgtt

tctcaaaaagttgaagaccaacaagtttacagacactattaattatttaattttgccccactttattttgtgggatcacaatttgttatattttaaacaaagctttcaattctaaacatgaaaacgatctggttga
catttcgggcgctctgcagaaaatcaaacttacacacggtgtcatcaaagatcagttgcagagcaaaaacgggtacgcggtccaatacttgtacgcgacgtttctcaacacggcctcgttctacgccaac
gtgcaatgtttaaatggtgtcaacgaaattatgccgccgcggagcagcgtaaagcgctattatggacgtgatgtggacaacgtgcgtgcatggaccacgcgtcatcccaacattagccagctgagtacgc
aagtctcggacgtccacattaacgagtcatctaccgactggaatgtaaaagtgggtctgggaatatttcccggcgctaacacagactgcgacggtgacaaaaaaattattacattttacccaaacctaatt
ccctaatcgactcggaatgccttttgtacggcgaccctcggtttaatttcatttgctttgacaaaaaccgtttgtcgtttgtgtcacaacaaatttattatttgtacaaaaatattgacgcaatggaggcgttgttta
aatctacaccattggtttacgcgctgtggcaaaaacataaacatgagcagtttgcacagaggctagagatgttgttgcgtgatttttgcttaattgccagttcaaacgctagttatttacttttaaacagcttac
acagctcatagctaacgaagaaatggtgtgcggagatgaagaaatattcaatttaggcggccaatttgtagacatgattaaaagcggtgctaaaggcagtcaaaatctgattaaaagcacgcaacaatac
cgacagactttaaatacagatattgaaactgtgtcttcacgagccaccaccagtttaaatagttacatatcttctcacaataaggtaaaagtgtgtggcgccgacatatatcataacacggttgtgttacaga
gcgtgtttattaaaaataactatgttgttacaaaaacgacgaacgtacaatcatgaatatttgcgctttgccctctgagtttctgtttccagaacatttgctcgacatgttcattgaatgataatataaatagagc
gcatttgattgcatgcaatcagtgttttattaattttagagcaacatgtacgataaatttatgatctatcttcacttgaatgggctgcacggagaagcaaaatactacaaatatttaatgtctcaaatggattttga
aaatcaagtagccgatgaaatcaagcggttttgtgaaactcgtctgaaaccggcaatcagttgcaacactttaactgcggaaagtctcaatacgctcgtagacagcgtagtctgcaaaaatggactgttaa
atccttacgccaaagaagtacagtttgctttgcaatatctttttgacgatgacgaaatatccaaacgagatcaagatggctttaaactatttttattacataattatgacaggtgtgaaaatatggaagaatattt
tttaattaacaattttagcatagcagactacgaatttgaagacatgtttgaaattgttcgtattgattgtagagatctgttattacttcttgctaaatataatatgtaattaaaattttgtttgttttattaaaatcctgga
ttaaaaaatgacgaataatttgatttgcgtgcacgccaacaagattcttcgtcattatgatcaatgcgtgcatcaagtttatgcttttgtaattggcttctgaccactttagccatttgagcgtatctgcattcgtcg
tctagagtttcaaacaccagatcggcgcaattataaaatccttcacccacgggatctatgcgctgccaacgcacatacattacaaattgatttgacctgtacggtattactacgggtatagaatagactaga
ctgttgtcacataatgaatcgcccggatttggaattaaatttgaatcgttaccacctatgtattctaattcgttccaagttattggattgcgacgatcccagtttgatttagtaataaacacttcaaaataactggg
ctcgtgtatggctgttggacaaaaatgaacattcatctgataaaccggttgatagcgatttaaatatagcgtatttggcctccagttgttaaaaggttcgtccattccgcttttatcaccaaacacagaattgcg
atcgtttgaaccggcaccgcaaagtgtgtgcggcacaaccctttgtttgattaggtcaaaatcgtcataattaggaccggccacagccgcgtattccatatactgttgaaacatgtattgcgctgtggaagcg
gccgccccggattctaaatcgagagctcgatatttataatagactgatttgtaagcattgcggcacgcggcgtcgggaatgttatcgccattgtcgggccaataaaagtttccatctttaaaacatttatattg
acgggccgtcggcacggacaaatagccgtgagagcgcactgccggcgcgtgaatcgcagcaaacaatgcaattaataatgcaatcattatgattatacttatagaacactaatcggaataataaccgct
gtcgtaatcttggtcaaaaacgttatgttgaaacataataacaccttacagtaacatacaataaaacaacatagtatcgtatataattataaactttattttttcattttatacaaacaaaatttatacgtattgtta
gcacattgagtgtcattttcgctgtctgaactatcacaatcatcgtcatcatcatcatcattgtcatcgtcgtcgtcacgtttgcgtttgacactgcatttttttggttaattttcactaacactggttcttttcgatcg
tacaattgattctgcatgtacttttgcatgatcgcggtaaaacactttgcaattttatccttttgttcgtcgccaaatatttccagcaactcgttcataaatgtgcacaaaatgcccatgtgttttatccagctgattc
gcattttcactggatcgaacaaacgcaaggggtacgctttttctgttaccttgccttcgatgtctatcaaaaggtacgggatacgatctccgttgccgggcacaaaatccgtgcctttgttaaccaaaatttctc
tacaatgcctagccaccgtaatcacgcgtcttttgggtgacggaccctcattatcgtcagttgatttgcgtttttttgcccgggttatcgttataggtcatactaaagctgtagtcggtcaacgattttgatttggca
aactcatcatagtattcataaaaactagtctgtaaactttgcaaacatttgtccatgtccaaatgacgcaatatttgttccactgccgtcctaaacgcgattctcataaaaacgggcatatcctttttaactaac
caacccttgtatacgattttattctcactgttgagatagcaatattttttctttttttaatagtattaaaactttcattaaattttcaaatgccattttgtaaccgtccgtgaatgagttattaacgcgtgtctcaacatgt
gtgcatatttgttttaatgtgtcggtttcgttggatatttcgttatagttaaatgtgggcaaaacaaatgtagaatctgtgtcgccgtacacaactttaaaagtgatgctgcccagattgaatttttctaaaatctcag
ggtcgttgctcaaaccttcaatcagagaaatggccagccgcaactgattgcgaccaactctagtgatgtagtttgcaagcactttgtaaaaaaatgccataataaccgtatatgctattggcggtgcgcttcac
ggaattttgtttttgatcgtacagatcgtacaagaatgccgattcgctttgattgtcgcgattctttttaaatttgcacctttcgcttaacaattttaatagcaatttaacaactattgcacgcgaattgtggttcaaat
acacgttgccgtcttcgcataaaattaaattggacaaacaagcacaaatggctatcattatagtcaagtacaaagaattaaaatcgagagaaaacgcgttcttgtaaatgcctgcacgaggttttaacactt
tgccgcctttgtacttgaccgtttgattggcgggtcccaaattgatggcatctttaggtatgtttttttagaggtatcaattttcttttgagattagaaatacccgctgcggctttgtcggctttgaattggcccgatatt
attgacagatcgtttttgttaaaaaaaatacgggtcaggctcctctttgccggtgctctcgttaatgcgcgtgtttgtgatggctgcgtaaaagcacgccacgctaatcaaatgcgaaatattacatatcacgtcg
tctgtacacaaacgatgcaatatacattgcgaatatacagaatcggccattttcaatttgacaaacaattttatcggcaacatgcaatcctgcacgttgtacttggcaatcacgtccagccgtcgagtgttgta
catcttgaccatttcggtccaaggcaaatcgattttgttttcacccaaatagtaactactgattgtgttcaattgaaagttttcaactttatgctgattagaatcgctgctgaaaaatttatacaaatcaatgtgaat
gtaatagttaaaataatacgtgtccactttgttgcccaacttgtttataaacagctttgtcgtcggcgccgcagccggcaaatcgtaacgctttaatagcattttggttttattcaatcgtccaagtatatagggc
agatcaaatacgtctccgttaaaatccaaaatcacatcgggatttgtaatttttatcatgtcaaaaaacgctgtaatcatgtcgatttcattttgaaacatgaccacatacgtgtcatcgtcataggtctctgga
atctgggtcggcagcttgtgatacataaaacaaaattttgcatactcgtcgtttttgtacaccacaaatcctatagacattatgcaatcaaccgatgctttcgacatgttgtggccgtccgaatgagtctcaatg
tcatagcacgacaaaacgggcatgatgccgctggttaaagtcatttcatcgaccaactcaaagtcttcattaaaatgttgcaaattaaacatgcgcgtcgtcgatccaccgacatagttattttggcagcgtt
gtgttttcttgaatcgcatataggcgccttccacaaacggcgtttgcatgtgtacgcgattaacgttgtgaagaaacttgtccaaacacgccgcgttgtccgatggcgctgctttgtttctttcgtatttaatcacg
tttatcttgttcaaataatttccttccacgcccggcgccacaaacgtggtgtagctgatgcacttgttgcggcaagacggaaatatgtgcttgtcgtagcattgtttgtaagaatacaaatttagttttactttaaa
gtaaaactgcagcactcgttctttgatatttgtattacaaaatgcaaacaagcaaccttgtttttcatcgtaatgcaaacgaatgatacgaaacgtatcggctgaagtaatattgaattctcctggttttgcatat
tctgcaaagcgcgttttgagttcattgtaaggatatattttcatttttaaatatgcagcgatggcccaaatatggaggcacagacgtcaacacgcgcactgtacacgatttgttaaacaccataaacaccatg
agtgctcgaatcaaaactctggagcggtatgagcacgctttgcgagagattcacaaagtcgttgtaattttgaaaccgtccgcgaacacacatagctttgaacccgacgctctgccggcgttgattatgca
atttttatcggatttcgccggccgagatatcaacacgttgacgcacaacatcaactacaagtacgattacaattatccgccggcgcccgtgcccgcgatgcaaccaccgccaccgcctcctcaaccccc
cgcgccacctcaaccaccgtattacaacaattatccgtattatccgccgtatccgttttcgacaccgccgccaacacagccgccagaatcgaacgtcgcgggcgtcggcggctcgcaaagtttgaatca
aatcacgttgactaacgaggaggagtctgaactggcggctttatttaaaaacatgcaaacgaacatgacttgggaacttgttcaaaatttcgttgaagtgttaatcaggatcgtacgcgtgcacgtagtaaa
caacgtgaccatgattaacgttatatcgtctataacttccgttcgaacattaattgattacaattttacagaatttattagatgcgtataccaaaaaacaaacatacgttttgcaatagatcagtatctgtgcact
aacatagttacgtttatagatttttttactagagtcttttatttggtgatgcgaacaaattttcagttcaccacttttgaccaattgacccaatactctaacgaactttacacaagaattcaaacgagcatacttca
aagcgcggctcctctttctcctccgaccgtggaaacggtcaacagcgatatcgtcatttcaaatttgcaagaacaattaaaaagagaacgcgctttgatgcaacaaatcagcgagcaacatagaattgca
aacgaaagagtggaaactctgcaatcgcaatacgacgagttggatttaaagtataaagagatatttgaagacaaaagtgaattcgcacaacaaaaaagtgaaaacgtgcgaaaaattaaacaattaga
gagatccaacaaagaactcaacgacaccgtacagaaattgagagatgaaaatgccgaaagattgtctgaaatacaattgcaaaaaggcgatttggacgaatataaaaacatgaatcgccagttgaacg
aggacatttataaactcaaaagaagaatagaatcgacatttgataaagattacgtcgaaaccttgaacgataaaattgaatcgttggaaaagcaattggatgataaacaaaatttaaaccgggaactaag
aagcagcatttcaaaaatagacgaaactacacagaggtacaaacttgacgccaaagatattatggaactcaaacagtcggtatcgattaaagatcaagaaattgccatgaaaaacgctcaatatttaga
attgagtgctatatatcaacaaactgtaaatgaattaactgcaactaaaaatgaattgtctcaagtcgcgacaaccaatcaaagtttatttgcagaaaatgaagaatctaaagtgcttttagaaggcacgttg
gcgtttatagatagcttttatcaaataattatgcagattgaaaaacctgattacgtgccgatttctaaaccacagcttacagcacaagaaagtatatatcaaacggattatatcaaagattggttgcaaaaatt
gaggtctaaactgtcaaacgccgacgttgccaatttgcaatcagtttccgaattgagtgatttaaaaagtcaaataatttctattgtaccacgaaatattgtaaatcgaattttaaaagaaaattataaagtaa
aagtagaaaatgtcaatgcagaattactggaaagtgttgctgtcacaagtgctgtaagcgctttagtacagcaatatgaacgatcagaaaagcaaaacgttaaacttagacaagaattcgaaataaaatta
aacgatttacaaagattattggagcaaaatcagactgattttgagtcaatatcagagtttatctcacgagatccggctttcaacagaaatttaaatgacgagcgattccaaaacttgaggcaacaatacgac
gaaatgtctagtaaatattcagccttggaaacgactaaaattaaagagatggagtctattgcagatcaggctgtcaaatctgaaatgagtaaattaaacacacaactagatgaattaaactctttatttgttaa
atataatcgtaaagctcaagacatatttgagtggaaaactagcatgcttaaaaggtacgaaacgttggcgcgaacaacagcggccagcgttcaaccaaacgtcgaatagaattacaaaaatttatattca
ttttcatcttcgtcatacttcaacagtcccaacacgttcatgttgtgattctcgccgttctcgacagttacgtaaatagttactttgattaaattatcttccagcagcattgagatttgattgaaatccgcacatagc
ttttgtagcgaatccgcttcggttttttatttgtgttgacgtagaaaacagatttgttccatttgcccaagtcggaagaggtagaacagtcatccgaatcggcaatgttcaactcgtcgcttttaaactgcacaat
aaacttgttatcgcccatgtcattttcttccaattcgctttttaacacatttacattgtacgaagcaacgtgtttgttcgatcgactaatgttgatctttgcgtttgtgcaattttgcaaatttgaatatgcttcgctttct
ttagcctcgcacaattcgatgcgcgtagagttgaccacgttccaattcatgtacacgtttgatccattaaaaatttgttgacactttatactgtaaatggtaaagatttggttttcattgtcttttaaatatttaaaca
cctcattgatgtcgtcagacccctttatattgttcttgaatagatttattagtgttttcgcattgacagaacattccacttgaaccacgtcgggatcgtcgttgagatttttgtacacaacctcaaaaacaactttgt
acaaaccgctgttgattttcttgtagataaatttgtactttacaataatattgacgccatcttcattttcaaaatgtttgttagtcaaatagtcgctcatgggggttgcagtttcaatttccatttcacattctttgtattc

gttgatctgaatcatttgactaaactttgttttcacataatttaaactaatgtcatagcacttgccttcttccatgtctttgaaagattgcgaatcgccgtagtattcttgaattttgttgtcggacattattcgaaaag
tgtaatggtattcattatcgatactcaacgtcattttgctcatcaatttaccactaatcctttgtaatttttctctaatcttcttggggctactggccatagccatgcgttttataagcggctcaccgctactttctcca
gacaaagatcttttggtcgccatattgctgttgtcgatatgtgggaatctatccgatggcaaatactgaatggcgacgaaatcgaagtgtcgccagagcaccgttcgttagcgtggagggagttgattataaa
cgtggccagcaacacgccgctcgacaacacgttcagaacaatgtttcaaaaagccgattttgaaaatttcgactacaacacgccgattgtgtacaatttaaaaacaaaaactttaacaatgtacaacgag
agaataagagcggctctgaacagacccgtccgatttaacgatcaaacggtcaatgttaatattgcgtacgtatttttgttctttatttgtatagttttgctgagcgtgttggccgtctttttcgacacaaacattgcg
accgacacgaagagtaaaaatgttgcagcaaaaattaaataaactcaaagatggtttgaacacgttcagcagcaagtcggtggttgcgctcgctcaaaattatttgacaaacgcccaacgcgcagacc
tagatgttggcgaaaactatcagagatcgacaaaaagtttcacgtttgccgacacgttgacacgtttttggatttgtgcggcggaccgggcgagtttgccaactataccatgtcgttgaacccgctttgcaaa
gcgtatggcgtcacgttgacaaacaactcggtgtgcgtgtacaaaccgacagtgcgcaaacgcaaaaatttcacaaccattacggggcccgacaagtcaggcgacgtgtttgataaaaatgttgtatttg
agattagcatcaagtgtggcaacgcgtgcgatctggtgttggcagatggctcggttgacgttaatggacgcgaaaacgaacaagaacgtctcaactttgatttgatcatgtgcgagacgcagctaattttaat
ttgcctgcgtcccggcggcaattgcgttttaaaagttttcgacgcgtttgaacacgaaacgatccaaatgctaaacaagtttgttaaccatttcgaaaaatgggttttatacaaaccgccttcttctcggcctg
ccaattccgaacgctatttaatttgtttcaataaattagttagaccgtattgtaacaattatgtcaacgagttggaaaaacagtttgaaaaatattatcgcatacaattaaaaaacttaaacaagttgataaact
tgttgaaaatataacgtgtgtataaaaagccagcggcttcaaatcaggcatcattcaacatggattcgctagccaatttgtgcttgaaaaccctgccttacaagttgagccgcctaagtttttacgaacaaa
atattgcgacgcatgtcgctacagatttttaccaaaattttctgatgaaaaaattttgtggacaatgcatatgcaacatatgcaacaatccaaaaaatatagattgtccatcatcatatatatcgaaaattaaac
cgaagaaagaaaacaaagaaatatatattaccagcaacaagtttaataaaacgtgcaaaaacgaatgtaatcaacaatcaaaccggagatgtttaatttcctatttacaaatgaaagttgtaaagagct
caattgttgttggtttaataaaaactgttacatgtgtttggaatataaaaagaatttatacaatgtaaatttgtatacgattgatggtcattgtccttcgtttaaagccgtttgtttttcatgtataaaaagaatcaaaa
cgtgccaagtttgcaatcaacctttattgaaaatgtacaaagagaagcaagaagagcgtttgaagatgcagtcgctgtacgcaacgttggccgatgtagatttaaaaatattagacatttacgatgtcgaca
attattctagaaaaatgatattgtgtgctcaatgtcatatatttgcacgctgtttttgtaccaataccatgcaatgtttttgtcctcgacagggttataagtgtgaatgtatatgccgacgatctaaatattttaaaaa
taatgtattgtgtgttaaaagtaaagcggcttgttttaataaaatgaaaataaaacgtgttccaaaatggaagcatagtgtagattatactttcaaaagtatatacaagttaataaatgtttaattttaaggatatt
gttatggaataaactataaaatgaatttgatgcaatttaatttttttgatactttccacagacggtagattcagaacgatggcaaacatgtcgctagacaatgagtacaaacttgaattggccaaaacggggct
gttttctcacaataacctgattaaatgtataggctgtcgcacgattttggacaagattaacgccaagcaaattaaacgacacacgtattcgaattattgcatatcgtcaaccaacgcgttgatgttcaatgaat
cgatgagaaaaaaaatcatttacgagttttaaaagctctcggcgtcagtttgcatcacaatccgtggtcgttgacatgttggctcgtcgcggcttctattattttggcaaagccggccatttgcgttgttccggatg
ccatatagtttttaaatataaaagcgtagacgacgcccaacgccggcacaaacaaaattgcaagtttctcaacgcaatagaagactattccgtcaatgaacaatttggcaaactcgatgttgcggaaaaa
gaaatactggctgccgatttgattcctccgcggctaagcgttaaaccttcggcgccgcccgccgaaccgctaactcaacaggtctccgaatgcaaagtttgttttgatagagaaaaatcggtgtgtttcatg
ccgtgccgtcacctggctgtgtgcacggaatgttcgcgtcggtgcaagcgttgttgtgtgtgcaacgcaaaaattatgcagcgcatcgaaacattacctcagtaaacattgcaaacgactacgacattcttt
aaaaataagctatatataaatattgcattgtatgacaaaaaaattattaacctactgcaaagtaaaacttgtaaaaggcttttcaaaaaaaatttgcgagtttattttgtcgctgcgtcgtgtcgcatctaagcga
cgaagacgacagcgacggtgatcgctattatcagtataataacaattgtaatttcatatacataaatattgtaaaataaaagacatattattgtacataatgttttattgtaattaaattaatacaccaatttaaa
cacatgttgatgttgttgtgaataatttttaaatttttactttttcgtcaaacactatggcgttgctttcgattagtttttcgttagcatttcatctaaaaaatcaaactgtttgcccggcgcgtttagggattctatggt
gtagtcgggcgtgtcgctgtttagatattggtccacttcgcgcattatgtccaagacgttgttctgcaaatgaatgagctttgtcaccacgtccacggacgtgttcatgtttctttttgaaaactaaattgcaaca
attgtacgtgtccactatacaattcggcttaatatactcgtcggcgcaatcgtatttgcaatccaatttcgtgttcaacaaattggtgatgatatctttgaacgtgcacgttttcaatttgtccttatcggccaacgc
aagtttcaattcgctctgtaaagtttctaaaattttgtctttattgttgtcaaattcgtgcgtgttgcgttccaaccacaatttgaacggctcgtcgacaaaaatgctgcgcaacacctcgtacaactgtctgccta
acgtgtacacttgctcgtattctttcatgctgacctctttgctaacgtacattactaaaaaatctacaagtattttcaaacatttgtaataggcgacgtattttgatttaagttttaaaccgtccaccgtgtattcgtc
cacgttcgcatcgaccacttttcgattattatcgccgcttgttgccggcgcgtcggcctgttcggttttaactatatccggttcaatatttaaagtttcaaaagatttaatggcattcataaaatcatctttttgcttt
ggcgtggtcaatggtaaatctatcgaggagttgtcgtccgtgtgctcttcgggcacgctgttcagacgtaacgtaatctttttgggatcgtcttcatcgggtatcaaatcggctttaattttattagaattgagcaa
cgacatggtggtcgcttgtaaatttaataaattaattaaagactgaaattgtatattgcacaaatttattttcatttttattgatcttactattaatacgctggcagttggtatgcttcatccattttttgtgactagaaaa
tttgctaaaaaactgagctcgtcctgtgttaaaacgttgtcgtccacgaatctatgcaatgtaaatgttacactgacattgtttaacaatgcatgtattaaaaaatcaacctgtcgcctactgagtttattagaag
agtcgaccgtttctactagtttgtagattttgttattttcaatttcattgtttaaaaacatgttaactactcgtttgagtttaagcgaaaaatccttgtccggatagacttgttcgcacagccaattgctaagagtggtt
ttgaccacggacaccttggtggtgaacgtcgtcgatttgaccagttcggtgaaaaagtttttcattaaattggacattttaacaaacacttatcaatctattgagctggtatttttgtttagaatcgcatcaagcgc
ttgctcgatctccaattttttttcggacgctcttagctttatgactcggtatgtcttctacggtagactcggtgttcttacttataatggccgggctgacgataataaacacgagaaacaatatgagcagatacaaa
aagatgctgttttcctttttgtcatacactaggctaaatatggccagtgcgcccaacaacaaatataaattcatttttattcccttactctattcgttgcgatagtacaacaacgattctcccgacgaaccggac
gaattgcgattatgctgcgcgtcgtcgtcgtcgttgttgttctcctcttcgctgctcgtttcgtctaaacctatattgtatttgttcaagtaatgtttggtgcttgcggaggattcgtggttcattaatttggccacttttt
gtaaaggcacgccgctattgtataggttactgctcaaataatgtcttatcatgttgctgcgcggccgttccatctcgacgcccgactcttcaaggagtcgcctgaaatctttgaagggcgtcgaggtgtttttag
atatttgcaaaatggtcgggtttcgtgaataaatctcgcgtgccaattccaacggtttcattttgatgttgttgagtgtgttattacgactgcgttttcgctttaaattaatcgtgtcgctgtgcagttttcctcttttaat
tagcacgttgagatcgtccacgctgagttggcgcgcttcgttgattcgcatacccgtccctaacatgatgcaaaacactatcgcgcccctaattagaccgcggtcgtgaacataatcgctgttgagcatttta
attttatcattaataaaatttaatatggtatctattacgtttttaagcattaaattcttttcctttttccctgatatttttgagctccttgtcgcgcggcagcataaccatgcggggaatttgtattcgggcaagttcatc
atgttggtgtaaaagtttatagtcaactgtagtgtttctttggtgaccgagcgaagttcgagcatgcgcctgcacagttcttggggatcaatgagaagtgtttggttttctatcgagtcaaactccttgtccaacga
gtacgacatgtcttccaggtgaacatcgtctaccgagcagtacacaattttaatgaatcgagacttgtaactttttaaagtggtgggcgcaaacggtttggggaacatgtacttgctccacagactgttgttttt
cacctcgtcgggcgtgcatcgttgccgatcggtggccaaatcgaacacggactcgaaccggggagcggattgaatttttattttccaagaattaaaattgttttcgttgcgaacattaaaaccgttcattgtgg
ttaatcaaatttattaaaaacaaaaggagaatcggtgtcaatactatccgaatattgttgttgttctcttaatattacgaaataatatattacatacagcagtaagaataaagctataaaagcgactacactaat
taaaattataattcccgccgacacgttgctcgtcgtgttgtcatagcccaccatgtcgtttattggcattttgtgaacgggctcgctaaattgttgcggttcgctggcagtatcgtcgttgagcgccaatttcaac
gggatgtattccaccttttcgtggttgcccaaccgatagtagggcacgtccaaattcatgtttacaacttatttgctaacaggaatttatgcaacaaaagtggtttggctttgatgagacgcaatttgaaatactt
gctgcatttacgcttaagattgtattccatgcgggcggcggtgttgtagtcgtacgcgctcgcgctgtgatacacgagccgtaaattggttgcgttgcgcaaacacttggcgccttgtttgttcgaatgctgtttt
atgcgtctgttaagattgctcgtgatgcccgtgtacaattttccattgtcttgccgcagaatgtacacgcaccacaccttgttggtgtacagagtcgtcgccatgattatgcagtgcgccctttcgtgttcggccg
agtggcgttaggcgcagccgcggcaataatcgcgttggcgtccttgttgtaatttatttgttgaaaaataaaacgtcttagagtttcgttttggaacgccaattcggtcaagctctcctggcaagcgcttttggtc
aaatgagcggccggcgaattgaccgcgttggcggccgacgttaagaaggtggcgttctggaacatgctgggctgcttgccggctcgcgtcgccagctcggccatgtaattgaatatgttggcagacgcag
atagcggcgccaaaaacgcaacgttctcttttaaactcatgactcgcgccctgtttttttcgttcagcacgtagtggtagtaatcgccgccgccggcaaacagatcgtcaatcacggcgttgatcagatcgtt
gatcatgttgatgtgcggaaagcgacgcgactcgactgcgctctgtatgtttggcggcagagtggcgtgcttgagcaacagagtcatgtaattgttggccagctgctgattgaaaggtaacggaatgggaat
gttgcacgtcaccgcttccgccaccatgtactggacggccagactgagttgtttggcggcctcggccaaagcgtctttgcccaacatatcagcgccaccgttgtaaaacttttgcgcgtacgccggcagcg
aatttagcacaaacgatggctgaaatatatttgaatcgctcgacagggactcggccgcgttgctctgtcccaactctttttgcaaccgaatcaggtggcgtatcatggtttcctccgattcaaaccgctttacc
acgtttacgctgattgggttcgtgtcgatgcacatgtcacgaatagtgtttataaaaagaatcatgagaggactaagttctgacatgtcattgcacctgtaatatctaataatcttttgaacaaaatccacacat
ttgttgtaccaaatagattcaccggcgtcgagcgtcggttctttgctcttgttgtacggtgcaatcgctaccgagtttgtgctgttgctgcggctcgtgtaatccatcctgttgtcgcgcgtggcgacggtcgtagg
caccgtcgccggcggcacgtacccgggcgcgttgtaagtttgcgcgctggtgaatatggccgttgccggattagagggatacctcagcggcggaggggtgttgtaataaaaattgccacgttcatctgtcat
actttttatttgtactcttatgattacaaaactcaatatacggattacttataatatagttgttgtgacaaaaaaagcgataataaaattaacaaaattatcaacaagttaatcatggaaaatttttcaacgttgaat
aacaacaacaaaatggcgcaggtcaacagcaccgtttgaaaactgacgcgccgacacaaaatgctttcgcaatttctaaaagccacattaaacgaattttcacctttgatataatcacgcagttctttttta
caacattcgtcgcacaaaattaacacctttataatgaggccgtcggtgtgtatcgtttgaaatgtccgcggttgactgcctggatgaaattcaaacgagtacccagtggacacgtgtatctgtgcaaaataat
gggctaatatcgaggcgcccgttttttaacctttacttttgatattttaataacattaatgttgttatttgcgtaatcagagtttttattgtggtgatcatcgtacaaataatgaagcaacagttcactatcgtatttaa
tcttgtttagcgttgtcaagtttttgtttcttaggcgttggagcgtctccgtcgtcgatattttcttcgaaatcgagtccaacaacgtcggcgtttccttcttgctcatcgatagcggcggcggaggcggcctctccg
tcgtcgtcattcgcggtttctacagtgcgtttgggcgacgacgtgtgtacagcagcgtccgtcttactattatcggaccgccaaatttttgtttgaaataacatttggcccttgttcaactttatttcggcgcagtta

aacattattgcattaagatcatattcgccgttttgcaccaaattgcacaaaacaccatagttgccgcacgacactgtagaatagcgttttgtacaacaatctgagttgcggcgagctagccaccttgataa
tatgggcgccaacgccccgtttttttaagtaatattcgtcttcaattataaaatctagtacgttttcatcttcactgttgatttgggcgttcacgatgatgtctggcgtaatgttgctcatgcttgccatttttcttataa
tagcgtttactttaatgtatttggcaatttattttgaatttgacgaaacgactttcaccaagcggctccaagtgatgactgaatatgtgaagcgcaccaacgcagacgaacccacacccgacgtaataggct
acgtgtcggatattatgcaaaacacttatattgtaacgtggttcaacaccgtcgacctttccacctatcacgaaagcgtgcatgatgaccggattgaaatttttgatttcttaaatcaaaaatttcaacctgttg
atcgaatcgtacacgatcgcgttagagcaaatgatgaaaatcccaacgagtttattttgagcggcgacaaggccgacgtgaccatgaaatgccccgcatattttaactttgattacgcacaactaaaatgt
gttcccgtgccgccgtgcgacaacaagtctgccggtctttatcccatggacgagcgtttgctggacacgttggtgttgaaccaacacttggacaaagattattctaccaacgcgcacttgtatcatcccacg
ttctatcttaggtgttttgcaaacggagcgcacgcagtcgaagaatgtccagataattacacgtttgacgcggaaaccggccagtgtaaagttaacgaattgtgtgaaaacaggccagacggctatatact
atcatactttccctccaatttgctcgtcaaccagtttatgcagtgcgtaaatgggcgccacgtggtgggcgaatgccccgcgaataaaatatttgatcgcaacttaatgtcgtgcgtggaagcgcatccgtgc
gcgtttaacggcgccggacacacgtacataacggccgatatcggcgacacgcaatatttcaaatgtttgaataataacgagtcacaactgataacgtgcatcaaccggatcagaaactctgacaaccag
tacgagtgttccggcgactccagatgcatagatttacccaacggtacgggccaacatgtattcaaacacgttgacgacgatatttcgtacaacagtggccaattggtgtgcgataattttgaagttatttccg
acatcgaatgtgatcaatcaaacgtgtttgaaaacgcgttgtttatggacaaatttagattaaacatgcaattcccaactgaggtgtttgacggcaccgcgtgcgtgccagccaccgcggacaatgtcaact
ttttacgttccacgtttgccattgaaaatattccaaaccattatggcatcgacatgcaaacctccatgttgggcacgaccgaaatggttaaacagttggtttccaaagatttgtcgttaaacaacgacgccatc
tttgctcaatggcttttgtatgcgagagacaaagacgccatcgggcttaacccgttcaccggcgagcctatcgactgttttggagacaacttgtacgatgtgtttgacgctagacgcgcaaacatttgtaacg
attcgggaacgagcgtttaaaaacgctcaattttggcgatggcgagtttttaaacgtattgagcagcacgctgaccggaaaagatgaggattatcgccaatttgtgctatatcctacgaaaacggccaaa
aaatcgtagaaaacgaacattttcagcgacgtatattgacaaatatactacagtcggacgtttgtgccgacctatatactacactttaccaaaaatatactacactaaactctaaatatactacaactccac
ttcaatataaccacactctcgtaaaacggcccaaaaaatatcgaaatatatggggcaaatacacgtttaaaaaacgctacgattccaaaaaacgctgcaactattccgcccgtgtttaatccctttgaaaac
cagccaaataacaggcaaaacgattctattctaccccctgtttaaccccttttcaaacgaccgacgccgtatggtacagcgaaccaggtggcgacgacgaccattgggtagtggcgccgccaaccgcacc
acctccaccgcccgagccagaaccagagccagaacccgagccagaacccgagccagagttaccgtcaccgctaatattagacaacaaagatttattttattcatgccactactcggttccgtttttcaa
gctaaccagttgtcatgcggaaaatgacgtcattattgatgctttaaacgagttacgcaacaacgttaaagtggacgctgattgcgaattggccaaagacctatcgcacgttttgaacgcgtacgcttatgtg
ggcaatgggattggttgtagatccgcgtacgacggagatgcgatagtggtaaaaaaagaagccgtgcctagtcacgtgtacgccaacctgaacacgcaatccaacgacggcgtcaaatacaaccgttg
gttgcacgtcaaaaacggccaatacatggcgtgtcccgaagaattgtacgataacaacgaatttaaatgtaacatagaatcggataaattatactatttggataatttacaagaagattccattgtataaac
attttatgtcgaaaacaaatgacatcattccggatcatgatttacgcgtagaattctacttgtaaagcaagttaaaataagccgtgtgcaaaaatgacatcagacaaatgacatcatctacctatcatgatca
tgttaataatcatgttttaaaatgacatcagcttatgactaataattgatcgtgcgttacaagtagaattctactcgtaaagcgagtttagttttgaaaaacaaatgagtcatcattaaacatgttaataatcgtgt
ataaaggatgacatcatccactaatcgtgcgttacaagtagaattctactcgtaaagcgagttcggttttgaaaaacaaatgacatcatttcttgattgtgttttacacgtagaattctactcgtaaagtatgttc
agtttaaaaaacaaatgacatcattttacagatgacatcatttcttgattatgttttacaagtagaattctactcgtaaagcaagtttagttttaaaaaacaaatgacatcatctcttgattatgttttacaagtaga
attctactcgtaaagcgagtttagttttgaaaaacaaatgacatcatctcttgattatgttttacaagtagaattctactcgtaaagcgagtttagttttcaaaaacaaatgacatcatcccttgatcatgcgttac
aagtagaattctactcgtaaagcgagttgaattttgattacaaatattttgtttatgatagcaagtataaataaccgcacaaagttaaatttttttcatttacttgtcaccatgtttcgaatatacccctaataacaca
actgtgcccggttgtttagtgggtgacattattcaagttcgttataaagatgtatcacatattcgctttttgtcagattatttatctttgatgcctaacgttgcgattgtaaacgaatatggacctaacaaccagttag
taataaaacgcaaaaacaaatcgctgaaaagcttgcaagatttgtgtctggacaaaatagccgtttcgctcaagaaacctttttcgtcagttaaaatcgttaaatgctgtttgtttgatgcgagacattatattt
cgctgggtttaccaattatttttaatccggctttgctacaaagaaaagtgccgcagcgcagcgtgggatatttcatgaattcaaaattggaaaggtttgccaattgtgatcggggtcatgtcgttgaagagaaa
caattgcagagtaatttgtatatagattattttgtatgatttgtggtttaaatgtttttaaaataaaagaataacaatttacacattgttttattacatggataatgttgtttgtttgacattaaaggttatcatggtgca
atgattaataataaaacaatattatgacattattttcctgttattttacaatataaaatcacaccaattgtgcaaagttttattatttgtttgtcgacggtcgaggggtcagcggcgtgtgcaacaataaaaaacat
gaagctgttaacaattttgattttattttattcatttttatgaatttgcaagcgctaccagattaccatcaagcaaataggtgtgtgttgctgggaactcgcattggatggaacgatgacaatagccaagatccc
aacgtatattggaaatggtgttaaataaaagtgaatatatttttataaaatttttatttaaaattccaagtaatccctgcaaacattaaacactgtaggtatttttaaatcttgccacatgcgaacaacgcacgg
cctgtcgtcgaacaccgctattacattatattttcctctgatatagttgttaaacaattttaattttaataaataatctttacaagtatcgtctgaaggcctcataaacaatttatatgatttaatatcaaaatacttt
caatccagtttcgagtgggctgttcacaaattacgcttctcccgctcataaacacgataattgcgtcgtggcaatttgccaaatacttaacgcaagtaataacgtctaagcgggcttcatcttgagcaactct
attatcaaaatcataaaacgatctatttgtgggcaaagctactgtaccgtctaaatcacataatacagcgcggggaaatttgtcgccgacaggaacgtaatattcgaaattatttacctttagaaactttttat
attgctttttaatagtttctggatttaatggaaatttatcagagcgtttataattgcgttcaagagccgtttccaaagaaacgtccatcaaacgcgttaaaaaatggtaattatgcgttgcggccatttttttgccac
atgtccaccgattgagtgttcaaattagtgtcgctgacaaccacgttggcaccacattttgcggcttttaaaaactgttcaatgcacattttggtaatttgttcttctttagtttgtctacatttccgcgattggttata
gaaagcgttcagttttgtataatcgccgtttaaaaacaacttaacgcgcacgtcgtctctgttgatttctgtatagccttttaaacttttggcatacgtgcttttgcccgaacccgaaatgcctatcaacaccaac
aattgttttgaagaaggcaatttaattgttggagcaagtttattatttaatgcctgcttagtcgatacaaattttataatatttttgatcattttaattttttcaggctcggttaattttaaaaattcgctctccacatcgat
cgtttgtgctttacgacatctgtacgctaaacatttccacggcaaagtttgcaccagttcgttgaaacgctgttgattcaaagtcaaacccgacaccataatatttattgtagactcgttggtgaacgtgtttcta
gcatcaacgtacggtttaatgacactttttaaatgcgggaaaagagctagaaagtcatcgtgttcgccatttataacaagctgcgccaatttagtaggattttcagcacggctctgatttttgtgcatgttcaaat
acacgtcgcttttaatcttgcatagtggcgcgttgtttttatcgtaaactacaaatccttcttccaaatttttcaactgggccgcgtgttcgacacattcttgcacagacgtaaactcgtaacatttggggtatttgc
aaaacggcaaattggaacagtaaaaataatcgcccgtttcgttgtttctgcttgccaaataccacaacgttggctgttcatcgtaaacggttacaattctgttgtgtttgcttgttaactcaaacatgtgagtcga
cgcgcagtctaaatattcgttacacaacgcttgaaattgattgtgggcctcgtcaagttgaagagcttgcaaaactaaacgtttaaacgtcacgtctgacacgcaaaggttttctgcaaaagcacttcctcgg
gtgctggcatgccattcgccgttgtacttgtagattttaattaaacttccgtcgatttttttcgtaaaacttaaaattctccttcgattggaacagtttgtgatgagcatcttcgccgccgatattttgtagcaattcttg
aaaattaaagaaacgatcgaaagaacgcgacacaacggcgtacgtgcggctgttaagaattaaaccgcgacattccacgaccacaggatgatctcgatcgcgttcaaacgattcgtaattaagaacca
tcaaatcgtgttcggtataatttttaattttgactttaaacttgtcacaaagatttttcactccgccgtttgcaagtagacgcgaaacgtgcaacatgattgctgtttaataatgcataccaatgctaaactgtctat
tatataaagtgcagtgataactttgttatcaacgcgttcgatgccgacatatataaacgcaatgtaacagtttttgctagtaccatcgcatacaacattatgaatacaaggggttgtgttaataataataaaatg
atatttatgaatgctttgggcttgcaacctcaaagtaaattgaaaattattgcacataaaatactagaaaaatgtaaacgtgacgcgtacacgcgtttcaagggcgtaaaggcgatcaagaatgaactaaa
aacatacaatcttacgttgcaacaatacaacgaggcgctcaatcagtgcgctttaaacgatagccgatggcgcgacacaaataattggcatcacgatattgaagaaggtgtgaaaataaacaagagac
atatatatagagttaattttaattctaaaacccaagaaattgaagaatattattacattaaagtagaatgttatgtaaacagttaattaatctacatttattgtaacatttgtggtaatagtggcgttggttatacattt
atatgattgtaatgttgtgtactcgttttgtaataaatttttgtgtttaatcaattcaatatttttatttgataaaaccttattttcgctactcaatttggcgtttttagacgcaagttttgcgtaatcgtcattgagcgatttt
agcgccttttcagttgtaattcgtttcagttgcaattctttaaaagatttatgcatgttgttgtagtcgcttttaattttgtctaacttttcttgcatagaaacgcttgtttgttgtaatttgtctaaatctaattgttgtttaat
gttgagctgcgtttgttcggcaatgtctacctgtagtttttttagtatcgcttgtgcttcagacagcatagtgtcgtcggcatttgcgttgttgtcttctgcgtcgtccaacagactttttttcaaacaacacactggcc
aaagaggccgcatcaaaattagcgtttattttattccattgtgcgacactcgacgcgctgcatttaatcacatccacaacgtttcggtttacgctgtaaacgttgaaatgcaaactttcaaccctacacaagg
gacatggtactttttttcgttttctaatcttgcgtatacacattgagcataattgatgtttgcacgtgtctagttctaatacgggtattatagtcaatctgtctattggttgcagaaaataatttttaatttctgcaaccg
aaaaacaaatgttgcattgcaatttaacaaactccatttttagacggctattcctccacctgcttcgcctgcaacaccaggcgcaggacctgccactgcgccgccgcccagagtagcgttaggatttgctct
tggtataaagtcgttgcgcaaaaagttgttttctgaattgattatttggtatcccaaaaacagcggaacgtacgtcgggtattcttcgtatccgctaagcgttctgtccagctcacgtgtgtcgccttcaaatttc
aaaacgtttctaatttgcaaacgattgggttgacttctcataatgtcactgcttcttatcgggttgtacaactcggggccgtcgggcacagacgcgaccagacccgtttcgtcaattatacacgtggcgcaatt
tctaaacctcaattcctccgtgtcgatttgcaagtactcgggcgctactgcgcgtcgaatcaaattttgcaaaaatccactgtaattgttaaataattgatcgccagcaccgcctcgaagcgctcgggcgttg
gtcacgtcaaagaaacgcaattcgtctcgcgacacccgcgaacaaaacgtgttcgggtttgtggtgtccagaatgctttttgtagttgcgtaaacgctgtgtataacgcgttgcgtgttgcttgtgaaaccttc
ggtatattttagattgtcgcatatagtgttaactgcgttttcgttgttatatatcaaatgaaagattagctgttcggcttgcatcatactgtttagattaaacacgtcttggtaattggttgcgcttggaattaaaattc
gcttgatacctctttctttatttccaactaaatgcctagcgatcgtcattttgaattgattgtcgtcttcgtcgaaaatgggcaaaaccatttttgacattttaaaacgttttatgaggtggttgttgcaaataaacca
tccatcgtcatgatacgcgtcgggcgaacacggcgatttgtatgttatgcacgcgtcgaacgacacgatggacgcgaaaatgcagcgattaactctcatttgtcgcggcgccatacccacgggcactagc
gccatattgttgccgttataaatatggactacggcgattttgtgattgagaaagaaatctcttattcaataaattttagccaagatttgttgtataaaattttaaattcttatattgttcctaattattcgctggcacaa

caatatttcgatttgtacgacgaaaacggctttcgcactcgtatacctattcagagcgcttgcaataacataatatcaagcgtgaaaaagactaattccaaacacaaaaaatttgtttattggcctaaagata
ccaacgcgttggtgccgttggtgtggagagaaagcaaagaaatcaaactgccttacaagactctttcgcacaacttgagtaaaataattaaagtgtacgtttaccaacacgataaaattgaaatcaaattt
gaacatgtatatttttcgaaaagtgacattgatctatttgattccacgatggcgaacaagatatccaaactgctgactttgttggaaaatggggacgcttcagagacgctgcaaaactcgcaagtgggcagc
gatgaaattttggcccgcatacgtctcgaatatgaatttgacgacgacgcgcccgacgacgcgcagctaaacgtgatgtgcaacataattgcggacatggaagcgttaaccgacgcgcaaaacatatca
ccgttcgtgccgttgaccacgttgattgacaagatggcccctcgaaaatttgaacgggaacaaaaaatagtgtacggcgacgacgcgttcgacaacgcgtccgtaaaaaaatgggcgctcaaattggac
ggtatgcggggcagaggtctgtttatgcgcaatttttgcattattcaaaccgacgatatgcaattctacaaaaccaaaatggccaatctgtttgcgctaaacaacattgtggcctttcaatgcgaggttatgga
caaacaaaagatttacattacagatttgctgcaagtgtttaaatacaaatacaacaatcgaacacagtacgaatgcggcgtgaacgcgtcatacgctatagatccggtgacggccatcgaatgtataaac
tacatgaacaacaacgtgcaaagcgtcacgttgaccgacacttgccccgcaattgaattgcggtttcagcaatttttgatccaccgctacagcagagcaattacatgaccgtgtccgtggacgggtatgtc
gtgctcgacaccgagttgagatacgtcaaatataaatggatgccaacaaccgagttagagtatgacgccgtgaataagtcgtttaacacactcaatgggccattgaacggtctcatgattttaaccgacttg
ccggagttactgcacgaaaacatttacgaatgtgtaatcacggacacgacaataaacgtgttgaaacatcgtcgcgaccgaatcgtgccaaattaaagcacgttaagcggatacaacgggcagtccga
gctgttaaagtcaatacaaccatcgttaacaaacgaatacgcattgttgtgacagctgaggatataaaaaggaatagagaagtaattgcaatgaaatatcccgttacaattccacggcacagcgtatgttg
ctcgagttctatcagttgcacacaacggcctaagaaaatttattaatgcttcatttgtatctatattagaaggataatacataggttcgcccaaaggactgggagaaggcggcggcgaaggtgtaggtgtagg
aggaataggagaaggcggcggcgaaggtgtaggtgttggaggaataggagaaggcggcggcgaaggtgtaggtgtaggaggaataggagaaggtggaggtgtaggtgtaggtgttggaggtataggtgt
tggaggaggtgtaggtgtaggtgttggaggtataggtgttggaggaggtgtaggcgaaggtggagaaggtgtaggagtaggtggaggtgtaggtaacggtacaattggtggagatgtaggtggtggtacaatt
ggtggatttggatacaattcctgaatgtcgtctaatatttttaaagttaataaaattattataaataaatttaatattattattattattattatcacaataatgtaccacatgttgcttaaatataaaaattaaacaaa
gaatgttgtattattgcaaatttaacaatttttgtattctccccatgtcatgcgttcgtaatgagcgggcggtttttatttctttgtatccacttgtaatcgttaatgtggttgtgaaaagtcatactgacgtaggcca
ttaaatttttcatgagcatattatttgacacaactgcaacatctgcgcctgccgtttcttgctggtacgaatcgacaaacgtaatgtctgtgccgtatttttctttgtcaagtgcaatttctataagctcaatgtggta
aatgatgaaacctttgacgttcatataatgatcgcggcacatggcgcactgtagtatgaaaaatacgttgtaaaatagcaccttcattgttttcaactgctgcatgacaaaatctaaactgcttttgtctcgcgt
atacaccatatcgtcgatgatgagactgagaaagtgcatggtgtcccatatggtagtaaacgtgtaagtaaaactcttgggctggcacgaacgcaaattgagttctgtggttttgtccataaattctatgcgaa
actgttgcaagtccatgtcgggggatgcgttaatggcccattcgatcaactgctgcacctcgtacttttgaatgtctttgtatttcatcaaacacgcaaaatggtataagtaagttgcttgcgaagacaacagtt
tggtgaggtgcgtcgatttagaggctcgcaaaaggtctatgagacgaaacgaatacaacagatagctgtctttgtaacgagaaaaaagcggcgtcagcggtatcatggcgactagcaaaacgatcgtgct
gtacttgtgtcaggcgccggccacagcgtcgttgtacgttagcgcagacacggacgccgacgagcctattatttatttcgaaaatattacagaatgtcttacggacgaccaatgcgacaagtttacttatttt
gctgaactcaaacaggagcaagccttatttatgaaaaaagtatacaaacacttggtgcttaaaaacgagggtgcttttaacaaacaccacgtattgttcgatgcaatgattatgtataagacatatgtgcatt
tggtcgacgagtctgcgttcggaagcaacgttatcaactattgcgaacagtttatcacggccatttttgaaatttttacgctcagcagtaaaatcgtcgtggccgtgcccgtcaattgggaaaacgataattta
agtgtacttttgaaacatttgcacaacctaaatctcattggaattgaaattgtaaattaaaacaaatcatgtggggaatcgtgttacttatcgttttgctcatactgttttatctttattggacgaatgcattaaattt
caattccttaaccgagtcgtcgcccagtttagggcagagcagcgactcggtggaattagacgagaacaaacaattaaacgtaaagctgaataacggccgggtggccaacttgcgcatcgcacacggcg
ataataaattgagccaagtgtatattgccgaaaaaccgctatctatagacgacatagtcaaagagggctccaacaaggtgggcactaacagcgtttttctgggcaccgtatacgactatggaatcaaatc
accaaacgcggccagcacatctagtaatgtaaccatgacgcgcggcgccgcaaactttgatatcaaggaattcaagtccatgtttatcgtattcaagggtgtgacgcccactaaaactgtagaggacaat
ggcatgttgcgattcgaagtcgacaacatgattgtgtgtttgatcgaccccaacacggcgccgctgtccgaacgagaggtgcgcgaattgcgcaaatctaattgcactttggtgtacacaagaaacgcgg
cagctcagcaagttttattggaaaataactttaccgtcattaatgctgaacaaaccgcctatctcaaaaactataaatcatacagagaaatgaattaataaaacaaaaagtctatttatataatatattattta
ttaacatacaaaatttggtacactagtgttcaaatcgtttctgttcaacgccattgtcatgttataaaacacatttgtagttttattgtaattatttttaaatttatttttaatttgctgtaataaaacttgttcattaaata
caaaagactttgaactacttgcgtttatattctttttataattgtactgaacaaacgaggggtgcaaaaagtttttgaaatgctgcacggcaatacctatcatctcctccattttgtcctctcctattgtaatagtgg
cactgcgcaccgttttaatgtttagaatgtaaatgagcgcatacagcggactattgttggtgctcaagcacattaggttgtgcttatgcatagggtcgttgctcagcagcgttttgtatactacaaagcccgtttt
ggggtcgcgtctgtacattagtacgtgcgacaaaaacaaacgcaccggcgtcacaagcgactcgtaatacatgctttctatcggaaactgtttggacttgatgtgttcgtacacggagccggcaaacttga
cgctgtctacaaacttatggttcgtgtaaacaatcaaaaatctgtcttgtacaccgtcgtcataatcgtccacgtacagcggcttgttgttaacaattaacattttgtagttggcttcatactttagcagcccttgg
tattttctgctcttggaatcgctcttgctcgaatcggcatgcttcttaaagtacgactcgctgcattgtttcaactcgttgatagtgtacaactgcgagttgagtttgctcacttccttgctcgctcgtttccttgttgga
ctctccgctgtggttgtcatcgtcaaacttgtgcatcaacaccaaatagtccaacagctcaaaaaacgacgacttgcccgaacccggttcgccgggcatgtaaatagccttctttccgtaatctacgggaa
tggccaaactagcggcgaaatgcatcaacataatcgcgttcgcgtgattaaaattggtgaagcgtttaaagtacaaatagccttcgacaatctttttcaaataattgtacgagtactccttcaagtccactttg
gacatgatgatgcgcatgtagaatcgagtcagccaagtgggcaaatcgtccgtgctgcgcgccaatatgattttgtcccaccacacattgtacttcttcaagatcattaacgcgtcggcgtggtgcgtgtaa
aatttggaaatgttatccgattcttcaaactgaacatcgggttcacgtgcaacatcatcgcgcaattcggttaaaaacaaacgtttatcattaaacttgtccatcaacatgtcgacatattcgattttgtgaattg
ttcgatacaagtactgaataattttgttgtgttctttggaaaaaaactctccgtgttggttaacaaattcgctgttcgtgcgaatcaacgtggtcgacacgtacgttttgttagtaaaaattagcatccaaatcaat
tcgctcaattctgcatcgttaccgaacatgtccgccatcaagcagacttttagcgcttttctattgatctttattttcttgtagcatttgcatttttggtcgagatcccgataccgttgaccgacacggtttgcattta
ggttgtgcaacatgtcggaaaccctgttcttgtttacgtacagagcgagcgtaatcagattttcatcgtccaaattccacaaatcgcgaaacaggttgtttaacgcgactcgcatatcggcttggcatgtgttg
caattgcccatgtagttaactatggccgtgttagtttttagcatttttacatctcggcacattttggcgatgtgataagttctataaatgctgagctcgtcggcgctagtagatagcatgtaattaaacgcgtcctc
gggcaaatacttttcgtcggtgggcttcttgaatgtctgcggcaacgtggtgcccaacaaaaatggacagctcgaatgaaagctgttggtgaacacgttgtacacaccgtgcgttgtcaagtacaagtatttc
caattgttaaattttatgttgctcaacttgtaacaattgcttttggtcaatttgaataggtcatcctctttctttacaatttgataatgtttgccgttgaaaaccaaattgactccggtcactacgttttccaattttcta
aagaatcctttacacacaatgtcaggcggcaagtttagcgccatcacattctcgtacgtgtacgcccacaattcatcgtgatccaaaatttcgtttttagccgactgagtcaaatatatcatgtagtgtatgcc
aaaataatagcccaacgatacgcacaatttggtatcgtcaaagtcaaaccaatgattgcaggccctattaaacactattttctcttgtttttgtaaggctcacatcgcttcaaagcttcattcaaagcttctttg
tcgcaggcaaataatgattcacacaaaagttccaaaaacagtttgatgtcggtttctctgtacgagaaattttcgttcttggtcaatatcttccacagtacatagattaaaaaatcaaaatttttaaatttgctttt
ttcaaagtattgttgtagaaggtttggatcgttggctcgttcgtgggtcgccaaaactttaaccatgttctcgtgaattgctataagccccaaattgatttgcgtttgaatgtagtctgcattttcgctgctcgccga
tataatgggtacgatgcgcggttttctggaacgcgtgtcgctcaagtccacgtcgtttttgtcaaaattgttgttctcgaacactctgaggcttttgaggttgacgttgacgatatgcttgtacttgggcaccgtaa
tgcattcctccaaattaatgtcgtccctaatgtaattgaaaaaatttttatccgaattgaccagctcgccattaactttgcacgtggccacagtgccgtcggccattttgagtataaacaagtcttcgtgagaat
cgtcaaacttggtttttccatttacaaacagcgtttgcggcggatcgtgattcgtgcgcaggctgagctcgacgttgagaaaacatttagggtcaaacacaaacaaatccacagggcctagtttttgttgtgt
atgattggtatcgtgggttcgatgacaattccaaattttatatttaaaaacagctgccatccgttaaaagagaaagcttgctttttgggccagttgggccaataatagtaatcgcccgcttgcacgcatttgttaa
tgtatccagggtcggtgctcttgaaaaaatcttcaaaattaatatactttttgtatgatgtcatagtgcttcttcaaaatgaaaggttttacaaaaatgcaaaaatcgttactttccaacacccagtcgtggccgtc
taatgtttgagctgcgtgtttctctgcaggttcttcggtgtcttcgcaagatgcgcccatgtcgtgtttcgcgcacggaccgttaaagttgtttctaattgtgtttaagaactgttgaaagttgttgacgtactcaaac
aatctacgtgttcctgttcgcgtgtttctaatgattaaatgatttgcatcttgcaagttgttaatctcgtacgttttgtcttgaggcacgtttttcaaaaaaaaattgtaaaatgttgtcaatcatgttggctatcgtgtttg
tacttttcgtgttaatttatttaataatttcgatcaaaaatcaccatccattcttacatagaatagaaacgctaatacaagatttcaacaacacattgttgtttggcgcgtatgtacagatttacgatttaagcacg
cccgcccgcaccgaacgattgtttattattgcgcccgaaaatgtggtgttgtataattttaacaaaacgctctattattacttggactcggcgaacgtgttttgtcccaacgagtttagcgtgaccacgttcacg
caatccactattaaaacgatcaacgagacgggaatatatgccaccgcatgcacgccggtcagcagcttgacgctaattgaacattttgcaacattaaaaaataacgtgcccgatcacacgctcgttctcg
atgtggtcgaccaacagattcagttttcaatactcgacattatcaattatttgatttacaatggctacgtggatttgttggccgaataacgcgtatatagacgcttgtacgttcatcgtagtaatcattttaatacat
ttgattgaactaaacatacatctgcaatgggtgaaagagtcactaaattttgcaatggaaaacggcgataaagaagacagcgacaatgaatagagtttatatttttatttaataaaatattgttcgtaatccat
aatgttttgtattatttcattgtgataatgttcccaatcttgcacgggggtggggcatcgtttgactttgacgtagaaatcgtacgcgtagttattagttggcagatcgtcgacaagtgtgatcgacttgaaaaagtt
tacatttttatcgctcaaatatttaattacaatttttggcgatttgggtatattgttgtcggatcgatgattgtgaatgtcaaaaacaaatttattttcaatgaaacgctttttaaattgtaatctacaatagcgttgtgt
gaattttgaactaaatcagagcgttcttcttgaacggtggaaccttcgctgataatgatatcaaaatagccttccaaatcgacgtctcgcatcgagtgtgctacatgatctctactgccatacgaccacaaga
ctaaaacgcaacccatctcgtgcaactcctgcaagctgtcatacacaaacggatctcgaatctcaacttgctcctcttcggttatgagagtgctgtccaaatcaaacacgaccacgtgcggaaatcccca
cgtcaaagattcgcttttgagagagaccactttgtagtgtggcaatagaaaccattctttaagaaacgaatacattggcggtttgttgctaagcacgcacatgtggcccaacactggcgttttgaatgcgcgtt

taatattgtgcctgatgtcgcgcatgtcgtcggcgggcgctttgaatatttgcatacagtaattgtaattgttttctatgatcttgcacagctgcgggtcgttgcaaaattgaaatattacatattcaaaaaatttat
acttttcaaagccaaggtatttgaggtcggcgtactcgcttaaaacgagaacatgtcgtttgatgatggcgtcgttaaggcgcaaacagatccatttgctttgaagcgaggaggccataatgtacaaaaatg
gaccagttacgccttatttaaactgtttaaagagtttcgtataaacaaaaactactctaaactaatagatttcttaacagaaaattttcccaacaacgtcaaaaacaaaacgttcaacttttcgtctaccggcc
atctgtttcactcgttgcacgcgtacgtgcccagcgtcagtgatttggtgaaagagcgcaaacaaattcgattgcagacagaatatttggcaaagctgttcaacaacacaataaacgatttcaaactgtaca
ctgagctgtacgagtttatcgaacggaccgaaggcgtcgattgctgttgtccgtgccagctattgcacaagagtctactcaacaccaaaaattacgtggaaaacttaaattgcaaactgtttgacataaagc
cgcccaaatttaaaaaggaaccttttgacaacattctttacaagtattccctaaattacaaaagtttgttgttgaaaaaaaaggaaaaacataccagcactgggtgtacacgcaaaaagaaaatcaaaca
caggcaaatattgaatgataaagttatttatttacaaaacagtaataaaaaataaactatttgagcttagcgggcttagtttaaaatcttgcagacatgattttgtaacagtcgaaagccaaacgagggcaggc
gacgaaatcgcttcgttcattcgctactgtcggctgtgtggaatgtctggttgttaatagtagcgtgttctgtaacttcggcgacctgtcgatgaacggctcctggatcttctgtatgtgcggggtctacccgggc
ggcgtctgtaacccgagcttctgcgcctgcgtgtcgaaccatatgtggtaccggttgaagaacggcgacggcgacgataaaccatgtttaaattgtgtaatttatgtagctgtaatttttaccttattaatattttt
tacgctttgcattcgacgactgaactcccaaatatatgtttaactcgtcttggtcgtttgaatttttgttgctgtgtttcctaatattttccatcaccttaaatatgttattgtaatcctcaatgttgaacttgcaattgga
cacggcatagttttccatagtcgtgtaaaacatggtattggctgcattgtaatacatcccgactgagcgggtacggatctatgtgtttgagcagcctgttcaaaaactctgcatcgtcgcaaaacggaatttcgg
taccgctgttgatgtattgttgcggctgcaacatttgtatcttttcgccgcgctcgatcaacaattcttcaagagtggtgcgtttgtcgcgctgtaaagccacgttttgtaacagcactattttcgcatatctcataa
tcggactgttgaaacagcgtgcaaacgacgaccgcataatatcgacggtcgtcaagtcgattgtggtcgaaggcatctccaacagagatcgcacggcgtccaacagcgtgtccgtttgaacctgcgtcat
ttgcggtctgcacgtgtagtcgtcaaacgtggtttcgagcagtttgaacaacgaatgatacttttccgatcgcagcaaaaatatcatggtcatgaccacgtcgctgattttgtattctgtagaactggtgctgttc
aacgaatagtgatggattagtttgcgagcagcatttctgtatcggcgcatgttgatcaactcttcggaaggctgcgcgggcgcggcggcgttggctcgcgcaaacaaatttattacgggacgcggcgtagg
ctgcgcggacgctggcgcggcgacgacgtccgcgtttcccgccgcgtactgagacgctatggcagcgttgttatttaaaattgtgttttgcgatttgcgagccacgtgcatcataaaatttatcaacacgtcg
gtgttcaactgcacgctttgatgttcgtcgcagagcaaaggaaatagctggggccatatcgccaattgcataggctcgtctatttttaaccgcaatttgtttatttccaaatacaacgcgatagcgctcatcgtg
accgacgacgcacacttactctgtaactatcacttggatcgtgttgtcgtaaacgcttcccaaaaagtctaacacgttgaccgtttcgattctattcaacttaattgtggacgcgttggcttgcatcggttccaa
cagactgcgcgctccgacagattgagtagacaaaatttttaaactttccgtcttattgggcgtaatgtcgttgattaacaacgacgcagccgtttgagaggccgcagtgttgatggtttgcaacatgtcgacgg
ccgccatttgcgtttgcgccgaaggtcttgctggcggcctgttgcggcggtttcttcgtgcttgcgacatgttgtcgtcagtgtccatatcggtatcatttattgaagcaatcatggttgagttcgataagcagag
atatttcgttgtccaattggtacttggtaatgatgtgccttataaatgtttcgggcacaatcatttctgtcattagcacgttacaaatatctattttgatcaatttcaatttatgaattaacagattaatgttttcgtccg
agtacttgctcatgatgaaacgacaaacgttgcggagttccaactccgctaccggatacgctttgttgggcaaactctctaaatagtgtctcaaataaaagccgatcaatacggtggacgctattttgttaac
ctttttcattttagtattgcggcccatttctatcatgaagtttttaaacggtagcaacagcctgtctccgttagcaacagtggagcagccgttgcattgcgcgctcaaaatactcaacacgcgctcgtgatcttct
tggcgcaatccgacggttgctttttgcattctttgacaaatggcacgcacatgtcgcgtttcgtgtacaaagaatacgctttgtcgcaaatcaagttatagaaaaattgcacaaatatctgcgtaatcaagttg
ttttcgttaataatgtcactttcgtttttgtaatcggttcgaagcaacacgtacaacatcagaggcatgccgaacatgggtcttaaaaaaatgtcccaaccatttttgcaagcccgcgtcgagggtgctcagcg
aggacgccaagtatttgcatttgcactcaaaacattgaattttgtttgcgggcttgcacgactgacacatgatcgcatccacgtcgggtgccggcgtcggattgtaatatttttgcaagtattgcataatggtcc
taaaatggggtacctgtttgataaactcgtcgcgcaaaaatatcgaaaaaatgtttttacattgtgtatgttgtctgtgttgttggcttgattctcaaaactactctttatggaaacaatacatttgttaaattctgtg
aaaaaagtaagacctttactgtccacgatcaagctttggttgaaatattttgaaaataaaaaacacaacgaatcgatttcatcttgtaacaattgcgcttcaaaacacacgttttcaaagcggtcgtaaatgtt
aaaccttaaactgtattgtaatctgtaagcgcacatggtgcattcgatataaccttataatatgaacgattccaattctctgttgattacgcgtttggcagcgcaaatactgtccagaaacatgcaaacggtgg
atgtgattgttgacgacaaaacgctcagtttggaagaaaaaaatagacacgttgaccagcatggtgttggctgtaaatagcccgccgcaatcgccgccgcgggtaacatccagcgacctggccgcatcga
tcattaaaaataacagcaaaatggtgggcaacgatttgaaatgcgatacaacgtgttgcgtatggccgtcgtttttgttaagcattatcccaagtattacaacgagacgaccgccggtttagttgccgaaat
agaaagtaatctgttgcaatatcaaaattatgtaaaccaaggcaattatcagaacattgagggttacgatagtttattaaataaggcggaagagtgttatgttaaaattgatagactatttaaagagagcatta
aaaaaatcatggacgacacggaagcgttcgaaagagaacaggaagcggagagattgagggccgaacaaactgccgcaaacgctcttctggagaggcgagcgcagacgtccgcagacgatgtcgtt
aatcgtgccgacgccaatattcccacggcatttagcgatccgcttccaggccccagcgcgccgcggtacatgtacgaaagttcagagtcggacacgtacatggaaaccgcccgacgtaccgccgaac
attacaccgatcaggacaaagactacaacgcggcgtacactgccgacgagtacaattccctggtcaagacggttcttttgcgtttaatcgaaaaggcgctggccactctaaaaaatcggttgcacataac
aactattgatcaattgaaaaagtttagagattatctgaatagcgatgctgatgctggagaatttcaaatatttttaaaccaggaagattgtgtgatactgaaaaatttgtcaaatttagcgtcaaagtttttcaac
gttcgttgcgtggccgacacgttagaggtaatgttggaagcgcttcgcaataatattgagttggtgcagcctgaaagcgatgccgtacggcgaatagtcataaaaatgacgcaagaaattaaagattcgag
cacgccgctgtacaacattgccatgtacaaaagcgattatgacgccataaaaaacaaaaacattaaaaccttgttcgacttgtacaacgacaggctgccaatcaatttcttggacacgtccgcaaccagt
ccagttcgcaaaacttccggcaagagatctgcggaagacgacttgttgccgactcgcagcagcaaacgtgccaatagacccgaaattaatgtaatatcgtcagaagacgagcaggaagatgatgacgt
tgaagatgtcgactacgaaaaagaaagtaaacgcagaaaattagaagacgaagattttctcaaattaaaagcattagaatttagcaaggacattgtcaacgaaaagcttcaaaaaattattgtggtcacc
gacggtatgaaacggctgtacgaatactgcaactgcaaaaattctttagagactttaccgagcgccgctaactatggcagcttgctcaaaaggctaaacctgtacaatctcgatcatatcgaaatgaatgt
aaatttttacgagttgctgtttccattgacactgtacaatgacaatgataacagtgacaaaacgctttctcatcaattggtaaattacatatttttggccagtaactattttcaaaactgcgctaaaaacttcaac
tatatgcgcgaaacttttaacgtgtttggcccgtttaaacaaatcgactttatggtcatgtttgttataaaatttaactttttatgcgacatgcgtaattttgccaaattaatcgacgagctggtgcccaacaaaca
gcccaacatgagaattcacagcgtgttggtcatgcgggataaaattgttaaactagcttttagtaatttacaatttcaaaccttttcaaagaaagacaagtcgcgcaacacaaaacatttgcaaagactaat
aatgttgatgaacgcaaactacaatgttatataataaaaaattataaaatatttttaatttttatttatattcagtacatttacacatattaacatattgtttatacaaattcttataatcattatgatttaaattgaattg
ttgtctaaacaaattaaacactttattaaacaataacttttcgttgtaatttttttactttgcacatgttataacaaaaaaattaaaattttcatcatgtctgatttgtctatggcgtcacagttgcttttaatgtaatcgca
agttaaccactcaaaaggacccttttctatttttaatttgtttaaatctttataatcagacttcagtttgtaaattagatttccacatcgaataataaatccttccagcgggctttggggaaacattaaagacttgaa
atttaacctttctacaaaatcgttgtacaaatatttgtgacacggaatagtattaaaccccacgttagtcaacaactcttgcgcctccacaaagggcacaaactccccgccgtataattgaatttcgtaagcg
tagtatttcaaactctctttctggtccacgtagttaattacgttaatgggtgtcgtttttgcgtcgtctttccaacccattaattcgccgtagacaataaaaccgtcattgaaccgcgcctgaagcgatcgcatgc
acgtttctaaatcttttcgaatgcggtaataattcataaaattgccgtccggtctgtaagtgtttcttgacccgtacgtaattttattttggttgcaaatgattctgaaattacaaccgtccaacttttcttgaacaat
aatttctttgtcggccaacgtaccttttttaccttgatctagatgcgacacagatggataaatttgatacacaattttattctcatcttcgggcattacgggtccgcgttcatttaacgcgtacatgacaatgttgtg
gcgaatgtcggtgcgctccggcggttctggcacgtggtgcagtctgtcctgcaattgttgcttccattgttgaaaaatattcggtccattcttgttgatactcgccgcgttgcatgagttttacgtacagttttaaaag
tttgacattctttacaaataacgttagagtttcgtcgattttgtatcctccattatttttgtttaaatccaatacatttaaatcgttcactaccagttgattgtttttatccatcgtaatttttatctcatcgcccacgttga
acaacatgtttaaaattttggtggatttcggcgcacgtttataatctaaataatattcaacgtacacgtaattgaacatgagctgcaacaatcctttggcattgttcaaaattttgtatctcatcaaagtataaata
attttcaccatcgacaccgtcatcaacttggttacaaactcgtacaattgcaagttttcaataccgtatttgtctttaaaatcttcacgtttactgaacatgcttaattcgggagattttccagtcaaaatgccaat
taatcccgtgtacaagtcaacgtatttgacatcgttgcccgattcatcttttgcatgtcgatttttcaaaagctctttattgtcgataaatttttcaaaggtctctcgatcacatttagtgtaaatatggtagtcagtgt
cgctgctttcgaccgcgtatcccttggcatggctgcccgtatcaatgcaaatgtacaccatgttagaatgtgctgcttactgtgcctgtatcaagccttatatacctcaaaatatttcacatttttgcatcatcgt
aaaatatacatgcatataattgtgtacaaaatatgactcattaatcgatcgtgcgttacaagtagaattctactggtaaagcaagttcggttgtgagccgtgtgcaaaacatgacatcataactaatcatgttt
ataatcatgtgcaaaatatgacatcatccgacgattgtgttttacaagtagaattctactcgtaaagcgagtttaaaaattttgtgacgtcaatgaaacaacgtgtaatatttttttacaatatttaagtgaaacatt
atgacttccaataattttgtggatgtggatacgtttgcaagacaattgattacagataaatgtagtgctctaatcaaagtgcggatctgttgccggcaaacattttagagattgtagagaaggccagagacaag
tattttgagggccaactcaaaaaaactatgaatacattaaaaaattatttttacgaaaaaatatatggacgattcgatagattataaagattttaacagacgcatcctattgatagtttttaaattcgctttaaac
aagagcacaatactttccatcgtacaaagagatcatcgagtggccattaaacgtttaaacaaaattaaccccgatttaaagagttctccgcgcaatgcttcagcattacaatgaatgtttggaaaatctaga
caatccagtcacggacgaacatcatttgttgacaaaagagttgctacaaaaatatttatcgaagcgtttgaatacagttacaccaacactaatgccatcagcatggacaaaacagatgaatttgattttatta
aaccggcattgaaacctttgccagatgcaagaccgccatcgcttttggccaacgtgatgaacgaacgtaaaagaaaattacaaaacaccaactcaacggcaaaatgtttgctaccagcaccaccgcc
acaattgcgtaaacttgaaaaaaagaatcatttattgcctttgttttctttgtaattatattgttgcatttctatttctaatatcatagttttctaataaagtagtttcatatttttgtttttgtacagtaattgtttcttggttta
acaagatcacaaccaataacataaagaataacacaatcataacaaaaattaaaaagccgcatactactagaacaaattctttaattagcgatcggtttctatttacaaattggccgagctgatcgccttca
gtcggcgagttgtgggcttggatgatgtcgacgatattgttgccggcgcgaccgcctgtcgctctcgatataatgtcggccgccgtcggtttcatgatgtgcttaactacaaataatagttgtacttgacgggcg

tcaccgtgatgccgctgctaaaacctccgtccgttaagacgcgttgcgttacaaaattaatgtttgtccgattagcgtagtcggaataatcaaacgtgttgggcggactaaaatcgggcatgttgatgggcac
aatgccgctggagctgatagcaatgctgtcgttcttgcaaaacagccgaatttttttgtagggctctgctttattcggcgcagacgacaccatctggtcaaagttgttcaatttttatgattacgttgggtaccaatt
gataggggaaaattattttctggaacattttgacaaagtccacaaccgtttggctatagtcgggaatgccgagcaaagactgcgcctgtttaatgtatttgagactggagcggtttactgtagcgcaattggat
ggcacgtcgcccttcataagccggcgcgttctctcccaattcaatttgttgtacaaattatcaatctcctcgtgcggcagattgattacatagcgcgcgggctgtttgcgatattgaaagatgcaaaaaaatgc
gtttcaacgacaatatcttcaccatggtggacgtttccagattgaaacataacaaaaagtcattgctttccaccaattctttaaaatgagacagcggaatttcacaagcgatcggtcgcaaattgctttttattg
gaggcggaacgctttgaccgttgcggtttttttagtaacgcgctgcacgcagattgcatgtccgtttcgggatacgtaaactcgatgggacatttggggttttcatggtgaacgatcatagtgttgcaataaaac
aagttgttggtcaggagcacgctaaaaacacgcgtttcgcccgcaccgatttcggtgatgggtaccaacgggttccagtagactatggtggcggacgctgtttttttggcgatcgactgtctatgttaacatc
atgctcgtgcctgtacactagcacagaattgaattttggaaattgtttttttgtcaatgtacaaccggtcgtcgtctgtgggcacgtacacgatcaagttttcgattaatttgttgcctacgtcgctttgcggttccac
caaattgtgagggaacgcaaaaaagcgatcgctaatacaaacttgaatctgaaacgggcactccatcgtgatgtatatgtcttacttcattagactttagattatttttaatttgtgaactcgtaccgtattcaat
agggtgtcgggcacgtaattgtaatggtaaaacagatcctgttgaacacgtgcgttgttcactacgattgaaatgcaaaaatacatcaagtacataaacactatgattagaaaggtagcagacagaaaata
tttcatctttaaatcttatgctagttgaataaaaatacatagtacttttatacgtttatttatatttgttttctttgttataaccgtaattgtaaaacttgtgatcgtgctcgccaggcataatttctttgcacatcagcttgc
gaatatatgtgacatcttcgtacaccgatttcttgatgttaccatcgtgaagcgttgtcggcttgagaggtttgcggtcgttgttgtaaaaatttttgcaccgaataattatccatagtgcagcacaggcaatgtca
ctgatgcatatgctttaatttttttattgcattcagttattatatgatttaataaacgtacacaatagcacgtttatcggttaaagataactttcaatatatataaaagtgtttgaattgcgagaccgtcaacataacgttt
atcaacgcgatgactaaacgacaatttgctttgctgtttgtgtggcaccacgacaaccaatttgtttgcaacacggacgaatacccgttttggcacaacattgaataccatgcacggcgctataaatgcatc
gttttgtactgtgtggaaaacgacggatcgctacaactgcccgtttgcaaaaacataaatctcataaattataaaaaaagcgtatcctcattattatggaaactgtgttgacagtatagtgaaacgtgctggca
aaaattgattatatgaaagtaactgcaatgttaaaccccccacctgttggacgtcgcgtacaattatttgctgttgatggacatggattgtgtggtgcaaagcgtgcaatggaaacaattgtcaaccgacacgt
attgttttgagccgttttacgactctcaaattaaatggttgtacgcgcccaaaagcggacaaagttttgatagttatcttgaaaactatgcaactctaattcgagtcaaacaagtgcagcaacatcgaaaaga
attaatactgcattgtgtggattttcttacaatgaaagcaaatgacaattttatggtgttcaaaaattatattaacatgattataaaagtgtatttgcaattttacaattacagatttcccatcaattttgaggacaac
acgatgaaaccttgtgtaaatttaacttttagacgtggcggcagttggaaaactcaactgcaacccgtatgcaattatgtttacaaaagtaaaaatatgccaaaatttattaaataaaacaaattaatttaaa
caagcgtttttattgacaatactcacatttgatattatttatatcaagaaatgatgtcatttgttttcaaaattgaactggctttacgagtagaattttacttgtaaaacacaatcaagaaatgatgtcatttttgta
cgtgattataaacatgtttaaacatggtacattgaacttaattttttgcaagttgataaacatgattaatgtacgactcatttgtttgtgcaagttgataaacgtgattaatatatgactcatatgtttgtgcaaaaatg
atgtcatcgtacaaactcgctttacgagtagaattctacttgtaacgcatgatcaaggatgatgtcatttgttttttttaaaattcaactcgctttacgagtagaattctacttgtaaaacacaatcgagggatgat
gtcatttgtagaatgatgtcatttgtttttcaaaaccgaactcgctttacgagtagaattctacttgtaacgcaagatcggtggatgatgtcatttttaaaaatgatgtcatcgtacaaactcgctttacgagtagaa
ttctacgtgtaaaacacgattacagcacttcgtagttgtatcgaaaattgttcaatggctctttgttaatgtcgtaattgattaatatgtcgtacaatttggcggcgttgtgtttgcacacgaccgtttttagttcttga
aacatttttttcgtgtatgtttagcatgttgtatttcagagtgcgatgtgtaatgctggtgacgagcatcaaaatgataaaatctaaagcggctaatttgtaatcccgttcatacgctctgtaatcgccaacaactct
gtggccagatctttttagattttgacaggcgttatggtacgaattgataatatttactatagtttctcttgttatcggtttgtcgattaaactgttaacaaacatcacgttgcccaagcgcgacggtttagacaccg
acttgtttttttgtctgttcaaatttgtacaaattaaaaacgctcatagactggtcgtcaggcagtgtgtcgttatacaaacaaaatggtaaaacgtttaattcgacaaacgacgagcacattaaagtttgttggct
gttaacgtcctggggatgtaaactgttattcataacgtaacacacttcaatgtcggaatgcttgttttcaaatttgtccttgtctacagtttcaatggtgattgagcgaggtttgagtttattttctaaattcatttggat
attttcaatatggtataccaccgacacgttgtgagccagcgatccttgattggttttaatcatattcaaaatattcatgatatggttgaaaaaagagtctgtcaaaacgtttgtgtcgttgttaaatatcgctttcca
gggtttactgttgcgtgactcaacgacggccgtgtaacataacaagcgcgccagttgcatgtgcgacaacttaatgttatcaatgtcggtgatgtttggcaccagattttcattgccgtcttccagtagcgtgct
cagttcggtcgagtagttattcaacgatcgattgtgcgattcaaacaagtttactatcgcaggttgtacatagtttttttatgtcgtcaaattgaattatatcgatcttgtccttgttctccagcataaacgacaaattt
tttaggtcgaatttaatatttggcgcgttttcgttggactttttgtaatttaacaacatcgccaacagtttgtgtaactcgccgttagcttgatctttgctaaacagtttattggtagcgtaattcacgttgtcgttcaaa
aacagcaactcgttgatgatcatttttttgtaaaagcgcgtacttgctcatgttgacagaatctcttacattgatcgtctaagagcggtcactacaggcgacatgacccaagggcaacaaatagacgcattaa
cccaaatacttgaaacttaccccaactactctgtataacaacatgagcattttaaaagttgtagaagcgtgcaatttggcacacactttttttaaaattgggttatttatttagggccaagacttgtttggatatcg
ctttagataatttggaactattgcgtcgaaagactaacataaaagaagtggcagtcatgttaaacaagaaaactacagagtgtttgcaattgaaacgaaaaatagataaaaaaaattgcacaacgtgttttaa
taaaaatttacactatcaaatgatgacatcataacgggttcaatattctgtgtgcaaaaataaatgacatcatatttcaaacttgttttacgcgtaaaattctactggtaaaacaagtttgagatggtatacagtt
aagtgttaacgcgtcactaaattcagacacgagcgcacgccctatatacatacaatttatcgcacaagatgcttaacgcgatctgtttataaactaaaacgcactgcaataaattttagcaagcatttgtatt
taatcaatcgaaccgtgcactgatataagaattaaaaaatgggtttgtttgcgtgttgcacaaaatacacaaggctgtcgaccgacacaaaaaatgaagtttccctatgttgcgttgtcgtacatcaacgtgacg
ctgtgcacctacaccgccatgttggtgggatacatggtaacattcaatgactccagcgaattgaaatatttacaatactggttgctgttgtcgtttttgatgtccgtggtgctaaacgctccgactctgtggacg
atgctcaaaaccacagaagcccatgaagtaatttacgaaatgaagctgttccacgccatgtactttagtaacgtgctgttgaattatgtggtgtttttggacaatcaaatgggtacaaattttgtttttgttaaca
atttaattcactgttgtgtactttttatgatatttgttgaattgcttatcctgttgggccacacaatgggcacgtacacggattatcaatatgtcaaatcgtgttatatggttatattgtttgtttcagttatgagtgttact
attgttatgggtttagagtgtttgaaaacgaaactaattgataacagtttgatgtttaacgcgtttgtgtgcgctttgtacattgtgattgcaataatgtggtctttaaaaaataatttgactagttattacgtttcaaat
ttacaaagtattcaagttgttccgttttcatacaacgatccgccgccaccgttctctaacattgtaatggatgacataaaaaaataaaaaataatttataaaaatgtttttttattctttcacaattctgtaaattctaa
acaaaaaaatataaatacaaacttattatgttgtcgtctaaataaacatcaatttgtaaatctggacacctattcatatcattgatattacagtctactatacaacaattaaaactaaccaaattatctttacaac
aattaaagcaattaaaacaatttaaataatcttcattgtcgtcgtataagtttatttgcactgtagacggtgttacacagcgatccattcgacgttcgtgttcgatcaactttctcgccaacttgtaccataaaaa
ttgtttggacaaaaagttttccaacaatggtaacggccaattcaacgtgacgatgcgcacgtcctcgggtatgcatttgttaaaaaacacacagctcgctttaccaaacgaaagcaaaggtactaaatatg
gcgccattggctgatttgttattccaagataattacaaataaactgatccgtcgtggggtgataactggcaggtgtcagctttaaataatcttcaacgttgttgtcgcgcaaaagtctgcattttacacgcgttgt
taatcccacgacttttgcatgtaaaatcggatccaaatactgcagaatcgtgtctataatttctaatggtaaacgtatgcgttttgctcgtgggcgctttgtaacgctcgacatcctaataacaactaacacaa
aactaaaatgatactcaatatattgcttttacagttcatctttaggtttaaactgtgcgtttatcgcgttgagcaagtcgccgttatcggcatcaatctcccaagcaaacaggccgcccaatttatttcggtcgac
atatttaacttttcctaacacagagtcgacgctgtcaaacgaaatcaaatcacctttacttttatcgaaaacgtacgacgcttgagcggcgctgtcaaacgtgtacacataattgttgagatctttttgaatttg
acgataatctacaacaccgtcctcccacgtgcccgataacttcgtataatgtatgctatacgaagttatggtaccgcggccgcgtagaggatctgttgatcagcagttcaacctgttgatagtacgtactaag
ctctcatgtttcacgtactaagctctcatgtttaacgtactaagctctcatgtttaacgaactaaaccctcatggctaacgtactaagctctcatggctaacgtactaagctctcatgtttcacgtactaagctct
catgtttgaacaataaaattaatataaatcagcaacttaaatagcctctaaggttttaagttttataagaaaaaaaaagaatatataaggcttttaaagcttttaaggtttaacggttgtggacaacaagccagg
gatgtaacgcactgagaagcccttagagcctctcaaagcaattttcagtgacacaggaacacttaacggctgacagaattagcttcacgctgccgcaagcactcagggcgcaagggctgctaaaggaa
gcggaacacgtagaaagccagtccgcagaaacggtgctgaccccggatgaatgtcaggctctgggaggcagaataaatgatcatatcgtcaattattacctccacggggagagcctgagcaaactggc
ctcaggcatttgagaagcacacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataagcggctatttaacgaccctgccctgaaccgacgaccgggtcgaatttgctttcgaa
tttctgccattcatccgcttattatcacttattcaggcgtagcaccaggcgtttaagggcaccaataactgccttaaaaaaattacgccccgccctgccactcatcgcagtactgttgtaattcattaagcatt
ctgccgacatggaagccatcacagacggcatgatgaacctgaatcgccagcggcatcagcaccttgtcgccttgcgtataatatttgcccatggtgaaaacgggggcgaagaagttgtccatattggcca
cgtttaaatcaaaactggtgaaactcacccagggattggctgagacgaaaaacatattctcaataaaccctttagggaaataggccaggttttcaccgtaacacgccacatcttgcgaatatatgtgtaga
aactgccggaaatcgtcgtggtattcactccagagcgatgaaaacgtttcagtttgctcatggaaaacggtgtaacaagggtgaacactatcccatatcaccagctcaccgtctttcattgccatacggaat
tccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaaaaggccgtaatatccagctgaacggtctggttataggtacattgagcaac
tgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacggtggtatatccagtgattttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtg
atcttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctcattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcgaagtgatcttccgt
cacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagatccggatatagttc
ctcctttcagcaaaaaacccctcaagacccgtttagaggccccaaggggttatgctagttattgctcagcggtggcagcagccaactcagcttcctttcgggctttgttagcagccggatcttctaggctca
agcagtgatcagatccagacatgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaaccattataagctgca
ataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggaggttttttaaagcaagtaaaacctctacaaatgtggtatggctgattatgatcagttatctagatccggtg

gatcccgggcccgcggtaccgtcgactgcagaattcgaagcttgagctcgagatctgagtccggacttgtacagctcgtccatgccgagagtgatcccggcggcggtcacgaactccagcaggaccat
gtgatcgcgcttctcgttggggtctttgctcagggcggactggtagctcaggtagtggttgtcgggcagcagcacggggccgtcgccgatgggggtgttctgctggtagtggtcggcgagctgcacgctgcc
gtcctcgatgttgtggcggatcttgaagttcaccttgatgccgttcttctgcttgtcggccatgatatagacgttgtggctgttgtagttgtactccagcttgtgccccaggatgttgccgtcctccttgaagtcgat
gcccttcagctcgatgcggttcaccagggtgtcgccctcgaacttcacctcggcgcgggtcttgtagttgccgtcgtccttgaagaagatggtgcgctcctggacgtagccttcgggcatggcggacttgaa
gaagtcgtgctgcttcatgtggtcggggtagcggcgaagcactgcaggccgtagccgaaggtggtcacgagggtgggccagggcacgggcagcttgccggtggtgcagatgaacttcagggtcagctt
gccgtaggtggcatcgccctcgccctcgccggacacgctgaacttgtggccgtttacgtcgccgtccagctcgaccaggatgggcaccaccccggtgaacagctcctcgcccttgctcaccatggtggc
gaccggtagccctttgaattccgcgcgcttcggaccgggatccgcgcccgatggtgggacggtatgaataatccggaatatttataggtttttttattacaaaactgttacgaaaacagtaaaatacttattta
tttgcgagatggttatcattttaattatctccatgatctattaatattccggagtaggtcgcgaatcgatactagcagctgatgcatagcatgcggtaccgggagatgggggaggctaactgaaacacggaag
gagacaataccggaaggaacccgcgctatgacggcaataaaaagacagaataaaacgcacgggtgttgggtcgtttgttcataaacgcggggttcggtcccagggctggcactctgtcgataccccac
cgagaccccattgggaccaatacgcccgcgtttcttccttttccccaccccaaccccccaagttcgggtgaaggcccagggctcgcagccaacgtcggggcggcaagccctgccatagccactacgggt
actccagataacttcgtataatgtatgctatacgaagttattcatatcgtctcagtagctcaaggtagagcgtagcgctctggatcgtatagatcttgctaaggttgtgagttcaagtctcgcctgagatattaaa
aaactttgtaattttaaaaattttatttttataatatacaattaaaaactatacaatttttttattattacattaataatgatacaatttttattattacatttaatattgtctattacggtttctaatcatacagtacaaaaat
aaaatcacaattaatataattacaaagttaactacatgaccaaacatgaacgaagtcaatttagcggccaattcgccttcagccatggaagtgatgtcgctcagactggtgccgacgccgccaaacttgg
tgttctccatggtggttatgaggttgctttttttgttgggcaataaacgaccagccgctggcatctttccaactgtcgtgataggtcgtgttgccgatggtcgggatccaaaactcgacgtcgtcgtcaattgctag
ttccttgtagttgctaaaatctatgcattgcgacgagtccgtgttggccacccaacgcccttctttgtagatgctgttgttgtagcaattactggtgtgtgccggcggattggtgcacggcatcagcaaaaacgt
gtcgtccgacaaaaatgttgaagaaacagagttgttcatgagattgccaatcaaacgctcgtccaccttggccacggagactatcaggtcgtgcagcatattgtttagcttgttgatgtgcgcatgcatcagc
tcaatgttcattttcagcaaatcgttttcgtacatcagctcctcttgaatatgcatcaggtcgcctttggtggcagtgtctccctctgtgtacttggctctaacgttgtggcgccaagtgggcggccgcttcttgact
cggtgctcgactttgcgtttaatgcatctgttaaacttgcagttccacgtgtttttagaaagatcatatatatcattgtcaatcaaacagtgttcgcgtgtcaccgactcggggttatttttgtcatctttaatgagca
gacacgcagcttttatttggcgcgtggtgaacgtagacttttgtttgagaatcatactcacgccgtctcgatgaagcacagtgtccacggtcacgttgatggggttgccctcagcgtccaaaatgtatacctgg
cactcgtccgtgtcgtcctggcactcgagcctgctgtacatttcgaagtggaaatgccgcatcgccacgatttgttgcacgtgtggtgcgcaaagtgattgttattctgccgcttcaccaactctttgcctttga
cccactggccgcggccctcgttgtcgcgaaaacagtcgtcgctgtcactgccccaacggtcgatcagctcttcgcccacctcgcactgctgcctgatgctccacataagcaaatcctctttgcccacattc
agcgttttcatggtttcttcgacgcgtgtgttgggatccagcgagccgccgttgtacgcatacgcctggtagtacccttgtagccgataatcacgttttcgttgtagtccgtctccacgatggtgatttccacgt
ccttttgcagcgtttccttgggcggggtaatgtccaagtttttaatcttgtacggacccgtcttcatttgcgcgttgcagtgctccgccgcaaaggcagaatgcgccgccgccgccaaaagcacatataaaa
caatagcgcttaccatcttgcttgtgtgttccttattgaagccttggtgtgactgatttactagtagcattgaggcatcttatatacccgaccgttatctggcctacgtgacacaaggcacgttgttagattaataa
tcttatctttttatcttaattgataagattatttttatctggctgttataaaaacgggatcatgaacacggacgctcagtcgacatcgaacacgcgcaacttcatgtactctcccgacagcagtctggaggtggtc
atcattaccaattcggacggcgatcacgatggctatctggaactaaccgccgccgccaaagtcatgtcaccttttcttagcaacggcagttcggccgtgtggaccaacgcggcgccctcgcacaaattga
ttaaaaacaataaaaattatattcatgtgtttggtttatttaaatatctgtcaaattacaatttaaataataaaaagcgtcctaaagagtattacacccttaaatcgattattagcgacttgcttatgggcgctcaa
ggcaaagtatttgatccgctttgcgaagtaaaaacgcaactgtgtgcgattcaggagagtctcaacgaggctatttcgattttgaacgttcatagcaacgatgcggccgccaacccgcctgcgccagacat
taacaagttgcaagaactgatacaagatttgcagtctgaatacaataaaaaaattacctttaccactgatacaattttggagaatttaaaaaatataaaggatttaatgtgcctgaataaataataataaggg
ttttgtacgatttcaacaatgaacttttgggccacgtttagcatttgtctggtgggttatttggtgtacgcgggacacttgaataacgagctacaagaaataaaatcaatattagtggtcatgtacgaatctatgg
aaaagcatttttccaatgtggtagacgaaattgattctcttaaaacggacacgtttatgatgttgagcaacttgcaaaataacacgattcgaacgtgggacgcagttgtaaaaaatggcaaaaaaatatcca
atctcgacgaaaaaattaacgtgttattaacaaaaaacggggtagttaacaacgtgctaaacgttcaataaacaacaccaactctattttgaccaatttagcgtcaagcgtcacaaacatcaacggtacg
ctcaacaatttgctagccgctatcgaaaacttagtaggcggcggcggcggtggcaattttaacgaagccgacagacaaaaactggacctcgtgtacactttggttaacgaaatcaaaaatatactcacgg
gaacgctgacaaaaaaataagcatgtccgacaaaacaccaacaaaaaagggtggcagccatgccatgacgttgcgagagcgcggcgtaacaaaacccccaaaaaagtctgaaaagttgcagcaat
acaagaaagccatcgctgccgagcaaacgctgcgcaccacagcagatgtttcttctttgcagaaccccggggagagtgccgtttttcaagagttggaaagattagagaatgcagttgtagtattagaaaat
gaacaaaaacgattgtatcccatattagatacgcctcttgataattttattgtcgcattcgtgaatccgacgtatcccatggcctattttgtcaataccgattacaaattaaaactagaatgtgccagaatcaga
agcgatttactttacaaaaacaaaaacgaagtcgctatcaacaggccaaagatatcgtcttttaaattgcaattgaacaacgtaattttagacactatagaaactattgaatacgatttacaaaataaagttc
tcacaattactgcacctgttcaagatcaagaactaagaaaatccattatttattttaatattttaaatagtgacagttgggaagtaccaaagtatatgaaaaaattgtttgatgaaatgcaattggaacctcccg
tcattttaccattaggtctttagatttggtaaggctagcacgtcgacatcatgtttgcgtcgttgacctcagagcaaaagctgttattaaaaaaatataaatttaacaattatgtgaaaacgatcgagttgagtca
agcgcagttggctcattggcgttcaaacaaagatattcagccaaaacctttggatcgtgcagaaattttacgtgtcgaaaaggccaccaggggacaaagcaaaaatgagctgtggacgctattgcgtttgg
atcgcaacacagcgtctgcatcgtccaactcgtccggcaacatgttacaacgaccagcgcttttgtttggaaacgcgcaagaaagtcacgtcaaagaaaccaacggcatcatgttagaccacatgcgc
gaaatcatagaaagtaaaattatgagcgcggtcgttgaaacggttttggattgcggcatgttctttagccccttgggtttgcacgccgcttcgcccgatgcgtatttttctctcgccgacggaacgtggatccc
agtggaaataaaatgtccgtacaattaccgagacacgaccgtggagcagatgcgtgtcgagttggggaacggcaatcgcaagtatcgcgtgaaacacaccgcgctgttggttaacaagaaaggcacgc
cccagttcgaaatggtcaaaacggatgcgcattacaagcaaatgcaacggcagatgtatgtgatgaacgcgcctatgggcttttacgtggtcaaattcaaacaaaatttggtggtggtttctgtgccgcgcg
acgaaacgttctgcaacaaagaactgtctacggaaaacaacgcgtacgtggcgtttgccgtggaaaactccaactgcgcgcgctaccaatgcgccgacaagcgacggctttcattcaaaacgcacag
ctgcaatcacaactatagtggtcaagaaatcgatgctatggtcgatcgcggaatatatttagattatggacatttaaaatgtgcgtactgtgattttagctcagacagtcgggaaacgtgcgattctgttttaaa
acgcgagcacaccaactgcaaaagttttaacttgaaacataaaaactttgacaatcctacatactttgattatgttaaaagattgcaaagtttgctaaagagtcaccactttagaaacgacgctaaaacact
tgcctattttggttactatttaactcatacaggaaccctgaagaccttttgctgcggatcgcaaaactcgtcgcccaccaaacacgatcatttaaacgactgtgtatattatttggaaataaaataacttgttga
ttgatatttaattttggcgacggatttttatatacacgagcggagcggtcacgttctgtaacatgagtgatcgtgtgtgtgttatctctggcagcgcgatagtggtcgcgaaaattacacgcgcgtcgtaacgtg
aacgtttatattataaatattcaacgttgcttgtattaagtgagcatttgagctttaccattgcaaaatgtgtgtaatttttccggtagaaatcgacgtgtcccagacgattattcgagattgtcaggtggacaaac
aaaccagagagttggtgtacattaacaagattatgaacacgcaattgacaaaacccgttctcatgatgtttaacatttcgggtcctatacgaagcgttacgcgcaagaacaacaatttgcgcgacagaata
aaatcaaaagtcgatgaacaatttgatcaactagaacgcgattacagcgatcaaatggatggattccacgatagcatcaagtattttaaagatgaacactattcggtaagttgccaaaatggcagcgtgtt
gaaaagcaagtttgctaaaattttaaagagtcatgattataccgataaaaagtctattgaagcttacgagaaatactgtttgcccaaattggtcgacgaacgcaacgactactacgtggcggtatgcgtgttg
aagccgggatttgagaacggcagcaaccaagtgctatctttcgagtacaacccgattggtaacaaagttattgtgccgtttgctcacgaaattaacgacacgggactttacgagtacgacgtcgtagctta
cgtggacagtgtgcagtttgatggcgaacaatttgaagagtttgtgcagagtttaatattgccgtcgtcgttcaaaaattcggaaaaggttttatattacaacgaagcgtcgaaaaacaaaagcatgatctac
aaggctttagagtttactacagaatcgagctggggcaaatccgaaaagtataattggaaaattttttgtaacggttttatttatgataaaaaatcaaaagtgttgtatgttaaattgcacaatgtaactagtgcac
tcaacaaaaatgtaatattaaacacaattaaataaatgttaaaatttattgcctaatattattttgtcattgcttgtcatttattaatttggatgatgtcatttgtttttaaaattgaactggctttacgagtagaattcta
cgcgtaaaacacaatcaagtatgagtcataatctgatgtcatgttttgtacacggctcataaccgaactggctttacgagtagaattctacttgtaatgcacgatcagtggatgatgtcatttgtttttcaaatcg
agatgatgtcatgttttgcacacggctcataaactcgctttacgagtagaattctacgtgtaacgcacgatcgattgatgagtcatttgttttgcaatatgatatcatacaatatgactcatttgtttttcaaaacc
gaacttgatttacgggtagaattctacttgtaaagcacaatcaaaaagatgatgtcatttgtttttcaaaactgaactcgctttacgagtagaattctacgtgtaaaacacaatcaagaaatgatgtcatttgtta
taaaaataaaagctgatgtcatgttttgcacatggctcataactaaactcgctttacgggtagaattctacgcgtaaaacgttatataggacttaaaataaacaacaatatataatgaaatatttattagattat
attatagcaatacatttacatttattataacaatactttttatttaatctgattatattataacgatacatttttatttagacattgttatttacaatattaattaactttttatacatttttaaatcataatatataatcattt
cgttgtgcatttcaaagcttttgatagcttcaaagtaatacatgaatttagagtattcaggaaaatgataaacgttggtaaacccgcatttggtacaatataacacgggatttttataatacagtttagtttttttac
acaatttgcaatagttgttagttgtaggcctcaaaggaaacgtgattgcgccgtccaatacctgggtaaactttttgactttaacagtggcaaacacggttcctttgatacccgaaaatcggttgtcttgcaga
gcggccatcatttcgcttggctcttgaagtataaaacagttgacgtcatccaccacgtcgggtctggtgcacatgcttcggtagcgctgcaacactatattggtgtatgtttccctgagaacgagaccgccgg
tggtgctaagatcgattgtttgaatgcgctcgttgggctctttgtgatttcgaattatgcgccgaattatttcaaacactttgcagttgtgatcgtcaattctcaattctttaacttccgtcgtgtgctctaaacttaca
gggaaaatgtattggtaaaaaaacctctctctggctaaatagctgaggtcgaccaaattgatagaaggatatatttcgtacgaggtttttggaacgttgtgatatagatagcatttttgacagcagatgtctatg

cggtcaggatcgtccaacggcttttcgatgtgaaccacaacatacaaaaaccattcgcgcgtgttgtctttgaatctataattgcaagtggtgcatcgcgaatcgctcatgtgctccatagtcttcttgtatttc
acaggcctgcttgcaaatttgcccgtcatgcgcatatctttgctgtttatgtagcccataatgtaattggtggaaaattttagcgtggctttcatgatgtcgcgttctaaatcgctcatgaaatgcatacgtagatc
gcgctcttgtttgaaatccagtttgtcgctgtacgcgggcaaaccttcaaacttgttcccaaactcgggcggcacaaaatatccatcttttctgttgacgactggtttttttacttacaatgctgctgtgctccaac
ggcttggccggagaggtgcacataggctgtttaggcggagagatgcgcgtaggtggtttgatgttagattttggcggcggacgaacaggcgacggcggcgagttggcggcaggcgctggcaaagatttgg
cacgacccttgcccccggtccttggcgcgtcaaaaatgttattctctcgaaaaaaacggttcattgtaactgttagttagcactcagaaatcaacacgatactgtgcacgttcagccatcgagaggctttata
tatggaaaccttatctatagagataagattgtatatgcgtaggagagcctggtcacgtaggcactttgcgcacggcactagggctgtggaggggacaggctatataaagcccgtttgcccaactcgtaaatc
agtatcaattgtgctccggcgcacacgctcgcttgcgcgccggatagtataagtaattgataacgggcaacgcaacatgataagaaccagcagtcacgtgctgaacgtccaggaaaatataatgacgtc
aaactgtgcgtcatcgccatattcgtgcgaggcaacgtccgcttgcgcagaagctcagcaggtaatgatcgataactttgttttctttcacatgtacaacgccgacatacaaattgacgcaaagctgcaatg
cggcgtgcgctcggccgcgtttgcaatgatcgacgataaacatttggaaatgtacaagcatagaatagagaataaatttttttattactatgatcaatgtgccgacattgccaaacccgaccgtctgcccgat
gacgacggcgcgtgctgtcaccattttatttttgatgcccaacgtattattcaatgtattaaagagattgaaagcgcgtacggcgtgcgtgatcgcggcaatgtaatagtgttttatccgtacttgaaacagttg
cgagacgcgttgaagctaattaaaaactcttttgcgtgttgttttaaaattataaattctatgcaaatgtacgtgaacgagttaatatcaaattgcctgttgtttattgaaaagctggaaactattaataaaactgtt
aaagttatgaatttgtttgtagacaatttggttttgtacgaatgcaatgtttgtaaagaaatatctacggatgaaagatttttaaagccaaaagaatgttgcgaatacgctatatgcaacgcgtgctgcgttaaca
tgtggaagacggccaccacgcacgcaaaatgtccagcgtgcaggacatcgtataaataagcacgcaacgcaaaatgagtggtggcggcaacttgttgactctggaaagagatcattttaaatatttatttt
tgaccagctattttgatttaaaagataatgaacatgttccttcagagcctatggcatttattcgcaattacttgaattgcacgtttgatttgctagacgatgccgtgctcatgaactatttcaattacttgcaaagc
atgcaattgaaacatttggtgggcagcacgtcgacaaacattttcaagtttgtaaagccacaatttagatttgtgtgcgatcgcacaactgtggacattttagaatttgacacgcgcatgtacataaaacccg
gcacgcccgtgtacgccacgaacctgttcacgtccaatccccgcaagatgatggctttcctgtacgctgaatttggcaaggtgtttaaaaataaaatattcgtaaacatcaacaactacggctgcgtgttgg
cgggcagtgccggtttcttgttcgacgatgcgtacgtggattggaatggtgtgcgaatgtgtgcggcgccgcgattagataacaacatgcatccgttccgactgtatctactgggcgaggacatggctaagc
actttgtcgataataatatactaccgccgcacccttctaacgcaaagactcgcaaaatcaacaattcaatgtttatgctgaaaaacttttacaaaggtctgccgctgttcaaatcaaagtacacggtggtga
acagcactaaaatcgtgacccgaaaacccaacgatatatttaatgagatagataaagaattaaatggcaactgtccgtttatcaagtttattcagcgcgactacatattcgacgcccagtttccgccagatt
tgcttgatttgctaaacgaatacatgaccaaaagctcgatcatgaaaataattaccaagtttgtgattgaagaaaaccccgctatgagcggtgaaatgtctcgcgagattattcttgatcgctactcagtaga
caattatcgcaagctgtacataaaaatggaaataaccaaccagtttcctgtcatgtacgatcatgaatcgtcgtacatttttgtgagcaaagactttttgcaattgaaaggcactatgaacgcgttctacgcgc
ccaagcagcgtatattaagtattttggcggtgaatcgtttgtttggcgccacggaaacgatcgactttcatcccaacctgctcgtgtaccggcagagttcgccgccggtccgtttgacgggcgacgtgtatgtt
gttgataagaacgaaaaagttttttggtcaaacacgtgttctcaaacacggtgcctgcatatcttttaataagaggtgattacgaaagttcgtctgacttgaaatcccttcgcgatttgaatccgtgggttcaga
acacgcttctcaaattattaatccccgactcggtacaataatatgatttacactgatcccactactggcgctacgactagcacagacgtcgtccgtccacaaactatttaaacaggctaactccaaacatgt
tcttgaccatcttggctgtagtagtaattattgctttaataattatatttgttcaatctagcagtaatggaaacagctcggggggtaatgtacctccaaacgccctgggggggttttgtaaatcctttaaacgctacc
atgcgagctaatccctttatgaacacgcctcaaaggcaaatgttgtagataagtgtataaaaaatgaaacgtatcaaatgcaacaaagttcgaacggtcaccgagattgtaaacagcgatgaaaaaatc
caaaagacctacgaattggctgaatttgatttaaaaaatctaagcagtttagaaagctatgaaactctaaaaattaaattggcgctcagcaaatacatggctatgctcagcaccctggaaatgactcaacc
gctgttggaaatatttagaaacaaagcagacactcggcagattgccgccgtggtgtttagcacattagcttttatacacaatagattccatccccttgttactaatttttactaacaaaatggagtttgtggtcac
tgaaaccaacgacacaagcattcccggagaacccattttgtttacggaaaacgaaggtgtgctgctgtgttccgtggacagaccgtctatcgttaaaatgctaagccgcgagtttgacaccgaggctttagt
aaactttgaaaacgacaactgcaacgtgcggatagccaagacgtttggcgcctctaagcgcaaaaacacgacgcgcagcgatgattacgagtcaaataaacaacccaattacgatatggatttgagcg
attttagcataactgaggttgaagccactcaatatttaactctgttgctgaccgtcgaacatgcctatttacattattatattttttaaaaattacggggtgtttgaatattgcaaatcgctaacggaccattcgcttt
ttaccaacaaattgcgatcgacaatgagcacaaaaacgtctaatttactgttaagcaaattcaaatttaccattgaagattttgacaaaataaactcaaattctgtaacatcagggtttaatatatatataattta
ataaataattaaataatatacaatgtttttattaattatattttttaatattaattaaaagtattaatatttaaaaaaaatgaatcaaattcatctaaagtgtcacagcgataaaatttgtcctaaagggtattttggcct
caacgccgatccctatgattgcacggcgtattatctgtgtccgcataaagtgcaaatgttttgcgaattaaatcacgaatttgacttggactccgccagctgcaagcctatcgtgtacgatcacacgggcag
cgggtgtacggctcgcatgtatagaaacttgttactatgaagagcgggtttccagttgcacaacactattatcgatttgcagttcgggacatataaatgtttaaatatatcgatgtctttgtgatgcgcgcgacattt
ttgtaggttattgataaaatgaacggatacgttgcccgacattatcattaaatccttggcgtagaatttgtcgggtccattgtccgtgtgcgctagcatgcccgtaacggacctcgtacttttggcttcaaaggttt
tgcgcacagacaaaatgtgccacacttgcagctctgcatgtgtgcgcgttaccacaaatcccaacggcgcagtgtacttgttgtatgcaaataaatctcgataaaggcgcggcgcgcgaatgcagctgat
cacgtacgctcctcgtgttccgttcaaggacggtgttatcgacctcagattaatgtttatcggccgactgttttcgtatccgctcaccaaacgcgtttttgcattaacattgtatgtcggcggatgttctatatcta
atttgaataaataaacgataaccgcgttggttttagagggcataataaaagaaatattgttatcgtgttcgccattagggcagtataaattgacgttcatgttggatattgtttcagttgcaagttgacactggcg
gcgacaagatcgtgaacaaccaagtgactatgacgcaaattaattttaacgcgtcgtacaccagcgcttcgacgccgtcccgagcgtcgttcgacaacagctattcagagttttgtgataaacaacccaa
cgactatttaagttattataaccatcccaccccggatggagccgacacggtgatatctgacagcgagactgcggcagcttcaaactttttggcaagcgtcaactcgttaactgataatgatttagtggaatgt
ttgctcaagaccactgataatctcgaagaagcagttagttctgcttattattcggaatcccttgagcagcctgttgtggagcaaccatcgcccagttctgcttatcatgcggaatcttttgagcattctgctggtg
tgaaccaaccatcggcaactggaactaaacggaagctggacgaatacttggacaattcacaaggtgtggtgggccagtttaacaaaattaaattgaggcctaaatacaagaaaagcacaattcaaagc
tgtgcaacccttgaacagacaattaatcacaacacgaacatttgcacggtcgcttcaactcaagaaattacgcattattttactaatgattttgcgccgtatttaatgcgtttcgacgacaacgactacaattc
caacaggttctccgaccatatgtccgaaactggttattacatgtttgtggttaaaaaaagtgaagtgaagccgtttgaaattatatttgccaagtacgtgagcaatgtggtttacgaatatacaaacaattatta
catggtagataatcgcgtgtttgtggtaacttttgataaaattaggtttatgatttcgtacaatttggttaaagaaaccggcatagaaattcctcattctcaagatgtgtgcaacgacgagacggctgcacaaaa
ttgtaaaaaaatgccatttcgtcgatgtgcaccacacgtttaaagctgctctgacttcatattttaatttagatatgtattacgcgcaaaccacatttgtgactttgttacaatcgttgggcgaaagaaaatgtgggt
ttcttttgagcaagttgtacgaaatgtatcaagataaaaaatttatttacttttgcctattatgcttagtcgtaaagagagtaatgaaattgagactgcatctaataatttcttgtatcgccgtatgtgagtcaaatatt
aaagtattcggaaagtgtgcagtttcccgacaatcccccaaacaaatatgtggtggacaatttaaatttaattgttaacaaaaaaaagtacgctcacgtacaaatacagcagcgtcgctaatcttttgtttaat
aattataaatatcatgacaatattgcgagtaataataacgcagaaaatttaaaaaaaggttaagaaggaggacggcagcatgcacattgtcgaacagtatttgactcagaatgtagataatgtaaagggtca
caattttatagtattgtctttcaaaaacgaggagcgattgactatagctaagaaaaacaaagagtttttattggatttctggcgaaattaaagatgtagacgttagtcaagtaattcaaaaatataatagatttaa
gcatcacatgtttgtaatcggtaaagtgaaccgaagagagagcactacattgcacaataatttgttaaaattgttagctttaatattacagggtctggttccgttgtccgacgctataacgtttgcggaacaaa
aactaaattgtaaatataaaaaattcgaatttaattaattatacatatattttgaatttaattaattatacatatattttatattatttttgtcttttattatcgaggggccgttgttggtgtggggtttgcatagaaataa
caatgggagttggcgacgttgctgcgccaacaccacctcctcctcctcctttcatcatgtatctgtagataaaataaaatattaaacctaaaaacaagaccgcgcctatcaacaaaatgataggcattaac
ttgccgctgacgctgtcactaacgttggacgatttgccgactaaaccttcatcgcccagtaaccaatctagacccaagtcgccaactaaatcaccaaacgagtaaggttcgatgcacatgagtgtttggcc
cgcaggaagatcgctaatatctacgtattgaggcgaatctgggtcggcggacggatcgctgccgcgacaaactgtttttctacttcatagttgaatccttggcacatgttggttagttcgggcggattgttagg
caacaaggggtcgaatgggcaaatggtaacatccgactgatttagattggggtcttgacgacaagtgcgctgcaataacaagcaggcctcggcgatttctccggcgtctttaccttgcacataataacttcc
gccggtgttattgatggcgttgattatatcttgtactagtgtggcggcgctaaacaagaaatagccgccggtggccaagagtatgcccgttcctcctacttttaagctttgcatgtaactatgtagacgggggttt
tgctgcagtgcgttttgaacaccttcgggcgtgcgcacgttggtttccgggaagttttgtttgactgcattggatcgcgtctgcttggtgtggtaattaaagtctggcacgttgtccacgcgccgcaattggctca
atgagtttatttgagggtctgaaatgccctgaaatactccgcgtatgttggggacatcattgttacgagtaattctgtttatgtctgaagtgctcacaaactggttgttagatagttgatagcccggctgaaatctg
ttgtttccaatgttgcgtacactgggcgcgttgagcacatttgtgaaaccggcgggagtgcttgttaaaagacgcgtattatcagtaataaaactggcctgattaggatacaatttattgactgcgcgaagattt
gaaaaaaaactcattttaaagcaaacttatttaataaatatatcacagtaaaggttttgcaaaactgccgtcgtcaatacaacacggcagcggcgtcatgttggtaaaatctaatcttctccttgctttagatt
ctgggcgagaaggcgcatttgttgtgtaagttatttcgacgtctgcattatttgttgtgtaaggtatctcgacgtatgaagcaactttaacattgttataattttttttaaatattgatgcgctccacggcgcgcgttg
atacggatgatatctctccattgtatgatcgctaaatttatataccgtttcaataaatatgttaaaacccaacatgttaattataatattcataatagtttgtttgttttcaataattatttttactgttttgaaatctaaa
agaggtgacgatgacgaatcagacgacgggttcagttgctataacaaaccaattggagtaaattttccgcatcctactagatgtgacgctttctacatgtgtgtcggttaaatcaaaaattagagttaatctg
ccctgaaggatttgaatttgatccagatgttaaaaattgtgttcctatatcagattatggatgtaccgctaaccaaaactaaaaataaaataaaatttatatagattaatgaaataaaatttatatagattaataa
aataaaatttatttaatatattatactatttatattatttacaacacttaacgtctagacataacagtttgtaacttagaaactaaatcagagttactgcgctcaaactctgaaaatttggcttgagactcggccac

ctgcttacgcaattgttcttgcagattattcacagtcgattgcaactcttctgatttcttggtagattcttgcaagtcatagtttgccttttgtaaatctaattcggcgacagcatgcttgtgtttaagcataatgtagt
cgctgtttaacatggtcattttatgttcaacttggctggtcttggctcgcagctcggacagttctttttgcaattgctccacatagttcaagtccgtggtgtgattgttgaccgtgttattttctaaaagctcgcgcca
atgctgtttgatggaatcctggttacgagtgacgttaatgggcataaattctacatacccgtgcttattgtacacgcgacaatctgatgaagtagcgctgcaaaaacatttgtacacagaattgtccataattat
cttgacataacacttgaaacacacagcatggttacaatgaatcgaagtcacaaacgaggaatttacgtttttagtgtctttaaaagtagtaaaacaaatattacacgaaacctctacttcttcttcgggttctg
attgctgctgctgctgctgctgcggctgcggagactgcggcgaggcaaacaaatctggcgactgtggtattacgtaattcggcgaataagatggactataagtgggagaccttggggcaatctcattcatca
gctgagcctcaagatctaaacctcgttgcagagccctctgcgcagctgtctccgacgcaatgttatcctggtactgctgggcagtgatgtcgggaaaccgttcacgatccacattttcactattaattagtatg
acgtcatcctcttgacttaatagcggatcgtcattgctaatgttaacctgaccgtgcacgtaatacgtgacaccctgacgatggtaggtgcgcgtcaacggctcgttgacgttcccgataatctgcacgttttc
ttcgctgacacgctgctcctgacgccgctcctgacggcgatggctgcgactgcttgaagacggctggctgcgactgcttgaagacggctgggcttcgggagatgttgtaaagttgatgcggcgacggctga
gagacagcctgtggcggcggctgctgctgggagtggcggcgttgatttggcgactcatggctgggctggtaggatactgttcactaggctgtgaggcttgaactgtgcttacgagtagaacggcagctgtatt
tatactgtttatcagtactgcacgactgataagacaatagtggtgggggaacttgccaggcaaaaatgaactttttttgtaatgcaaaaaagttgatagtgtagtagtatattgggagcgtatcgtacagtgtag
actattctaataaaatagtctacgatttgtagagattgtactgtatatggagtgtcaggcaaaagtgaactttttttgcattgcaaaaaaaattcattttaaatttatcatatcacaggctgcagtttctgttatctgtcc
cccactcaggcgtgcagctataaaagcaggcactcaccaactcgtaagcacagttcgttgtgaagtgaacacggagagcctgccaataagcaaaatgccaagggacaccaacaatcgccaccggtc
tacgccatatgaacgtcctacgcttgaagatctccgcagacagttgcaagacaatttggacagcataaaccgccgagacagaatgcaagaagaacaagaagaaaacctgcgctatcaagtgcgtaga
aggcagcgtcaaaaccagctccgctccatacaaatggaacagcagcgaatgatggcggaattaaacaacgagccggtgattaattttaaatttgagtgtagtgtgtgtttagaaacatattcccaacaatc
taacgatacttgtcctttttttgattccgactacgtgcgaccacggtttttgtttcaaatgcgtcatcaatctgcaaagcaacgcgatgaatattccgcattccactgtgtgctgtccattgtgcaatacccaggta
aaaatgtggcgttccttaaagcctaacgctgttgtgacgtgtaagttttacaagaaaactcaagaaagagttccgcccgtgcagcagtataaaaacattattaaagtgctacaagaacggagcgtgattagt
gtcgaagacaacgacaataattgtgacataaatatggagaatcaggcaaagatagctgctttggaagctgaattggaagaagaaaaaaaatcacagtgatcaagtagcttctgaaaaccgacagctgata
gaagaaaatactcgtctcaatgaacagattcaagagttgcagcatcaggtgaggacattggtgccgcaacgtggcattacggttaatcagcaaattggccgtgacgacagtgcgccagccgagctgaac
gagcgttttcgctcacttgtctattcgactatttcagagctgtttattgaaaatggcgttcatagtattcaaaattatgtttatgccggaacttctgctgctagttcatgtgatgtaaatgttactgttaattttgggtttg
aaaattaatgtgatatgaaatgtatatataaaaatgatggaataaataataaacatttttatactttttatgttttttttatttcatgtgattaagaaacttttaagatggatagtagtaattgtattaaaatagatgtaa
aatacgatatgccgttacattatcaatgtgacaataacgcagataaagacgttgtaaatgcgtatgacactatcgatgttgaccccaacaaaagatttataattaatcataatcacgaacaacaacaagtc
aatgaaacaaataaacaagttgtcgataaaacattcataaatgacacagcaacatacaattcttgcataataaaaatttaa

8.3 List of primers

Table 8.2 –List of used primers in this thesis displaying sequence and use. Sequences highlighted in *blue* ***represent primer overhangs.***

Name	Sequence 5' – 3'	Use
AmpFor	caggagctaaggaagctaaaatgagtattcaacatttccgtgtcgccc	Ampicillin amplification from pUC57
AmpRev	caataactgccttaaaaaaattaccaatgcttaatcagtgaggcacctatc	Ampicillin amplification from pUC57
pUCDMFor (5' Phosphorylated)	taactttaaataatgccaattatttaaagttaggtttaaacgtacccgtagtggctatgg	Insertion of RoxP site in pUADM
pUCDM_EcoRV_Rev (5' Phosphorylated)	gatatcggtaccgcggccgc	Insertion of RoxP site in pUADM
pUCDM_For	ttttttaaggcagttattggtgcccttaaacg	Amplify backbone of pUCDM
pUCDM_Rev	tttagcttccttagctcctgaaaatctcgataactc	Amplify backbone of pUCDM
AmCyan_For	aagcgcgcggaattcaaaggggatccatggcactgagcaacaaattcattgg	AmCyan amplification from pCPSS2185
AmCyan_Rev	agtgagctcgtcgacgtaggtctagattaatggtgatggtgatggtgttcgaac	AmCyan amplification from pCPSS2185
mTagBFP_For	aagcgcgcggaattcaaaggggatccatgagcgagctgattaaggagaacatg	mTagBFP
mTagBFP_Rev	agtgagctcgtcgacgtaggtctagactaggatccattaagcttgtgcccc	mTagBFP
mCherry_For	aagcgcgcggaattcaaaggggatccatggtgagcaagggcgag	mCherry

mCherry_Rev	agtgagctcgtcgacgtaggtctagattacttgtacagctcgtccatgcc	mCherry
ASeq1	gttaagtgttcctgtgtcactcaaaattg	Sequencing primer: pUADM_RoxP_AmCyan
ASeq2	cgttagccattgtcaccttgaagg	Sequencing primer: pUADM_RoxP_AmCyan
ASeq3	cataccgtcccaccatcgg	Sequencing primer: pUADM_RoxP_AmCyan, pUPDM_VloxP_EGFP, pUZDM_Vox_mCherry
ASeq4	gccaactcagcttcctttcg	Sequencing primer: pUADM_RoxP_AmCyan, pUZDM_Vox_mCherry
ASeq5	catgtctggatctgatcactgc	Sequencing primer: pUADM_RoxP_AmCyan
ASeq6	ctacgatacgggagggcttac	Sequencing primer: pUADM_RoxP_AmCyan
ASeq7	ctggcgaactacttactctagc	Sequencing primer: pUADM_RoxP_AmCyan
BSeq1	gagagcttagtacgtactatcaacagg	Sequencing primer: pUBDM_SLoxP mTagBFP, pUPDM_VloxP_EGFP
BSeq2	ccggattattcataccgtcccac	Sequencing primer: pUBDM_SLoxP mTagBFP
BSeq3	cgacaagcttgtcgagaagtac	Sequencing primer: pUBDM_SLoxP mTagBFP, pUPDM_VloxP_EGFP
BSeq4	cggccgcatcttcactgg	Sequencing primer: pUBDM_SLoxP mTagBFP
BSeq5	gaggctaactgaaacacggaagg	Sequencing primer: pUBDM_SLoxP mTagBFP, pUPDM_VloxP_EGFP
BSeq6	gtacagcgtctcggtgaagg	Sequencing primer: pUBDM_SLoxP mTagBFP
BSeq7	ggtgatgctgccaacttactg	Sequencing primer: pUBDM_SLoxP mTagBFP
BSeq8	gagaagcacacggtcacactg	Sequencing primer: pUBDM_SLoxP mTagBFP, pUPDM_VloxP_EGFP
PSeq1	cgtcgtccttgaagaagatggtg	Sequencing primer: pUPDM_VloxP_EGFP
PSeq2	ctcgacatcggcaaggtgtg	Sequencing primer: pUPDM_VloxP_EGFP

PSeq3	gaagagttcttgcagctcggtg	Sequencing primer: pUPDM_VloxP_EGFP
PSeq4	gactagtgagctcgtcgacg	Sequencing primer: pUPDM_VloxP_EGFP
ZSeq1	ccttagaggctatttaagttgctg	Sequencing primer: pUZDM_Vox_mCherry
ZSeq2	ccatgttatcctcctcgcc	Sequencing primer: pUZDM_Vox_mCherry
ZSeq3	gagatcaagcagaggctgaag	Sequencing primer: pUZDM_Vox_mCherry
ZSeq4	gacaacaccctggcctgg	Sequencing primer: pUZDM_Vox_mCherry
ZSeq5	gcgtacagctcgtccag	Sequencing primer: pUZDM_Vox_mCherry
pDSFrag4HomAFor	gcttccggctcgtatgttgtgtg	Amplify HA-Gent-HA fragment for removal of fragment 4 and HA-Frag4-HA
pDSFrag4HomARev	caggatccgaggcctacatgcag	Amplify HA-Gent-HA fragment for removal of fragment 4
RoxAmpFor	taactttaaataatgccaattatttaaagttaggtttaacggttgtggacaacaagcc	Amplify ampicillin with RoxP overhangs from pUC57
RoxAmpRev	taactttaaataattggcattatttaaagttagaggtcgaggtggcccgg	Amplify ampicillin with RoxP overhangs from pUC57
Frg4AmpInsertFor	atgccaattatttaaagttagctgtcgtaatcttggtcaaaaacgttatgttg	Amplify pDS-Frag4 with overhangs for ampicillin insertion by SLIC
Frg4AmpInsertRev	attggcattatttaaagttaggttattattccgattagtgttctataagtataatcatagttgctc	Amplify pDS-Frag4 with overhangs for ampicillin insertion by SLIC
NheIRoxAmpFor	taagcctgctagcagcatactaactttaaataatgccaattatttaaagttaggtttaacggttgtg	Amplify ampicillin fragment with overhangs for restriction digestion cloning
NotIRoxAmpRev	acttgagcggccgcgaacagtaactttaaataattggcattatttaaagttagaggtcgaggtg	Amplify ampicillin fragment with overhangs for restriction digestion cloning
NotIFrg4For	ctgttcgcggccgctcaagtgctgtcgtaatcttggtcaaaaacgttatgttg	Amplify pDS-Frag4 with overhangs for ampicillin insertion by restriction digestion cloning
NheIFrg4Rev	gtatgctgctagcaggcttaggttattattccgattagtgttctataagtataatcatagttgctc	Amplify pDS-Frag4 with overhangs for ampicillin insertion by restriction digestion cloning
pDSFrag4AmpRev	ggccatggcggaccgaatttaaatac	Amplify HA-Frag4-HA

Frag4CheckFor	gtgggccagcacactacg	Validation of fragment 4 modifications
Frag4CheckRev	cgcaaacgtgacgacgac	Validation of fragment 4 modifications
Frag4CheckFor2	cttttgaagcaaataaagtgggccagcacactacg	Validation of fragment 4 modifications
Frag4CheckRev2	gcagtgtcaaacgcaaacgtgacgacgac	Validation of fragment 4 modifications
Frag3HomAFor	gtgagcggataacaatttcacacaggaaac	Amplify HA-Gent-HA fragment for removal of fragment 3 and HA-Frag3-HA
Frag3GentCheckFor	ggcatgtaaagactatattcgcggcg	Validation of fragment 3 removal by GentR
Frag3CheckFor	cgtacgcgtttgtgtccaggtc	Validation of fragment 3 insertion
Frag3Rev	ggccgaagatggtacgcacg	Amplify HA-Gent-HA fragment for removal of fragment 3 and HA-Frag3-HA, validation of fragment 3 modifications
Frag3AmpCheckFor	gacagcagggctattgcacgc	Validation of fragment 3 modifications
ISceIBSDFor	taatacgactcactatagggtagggataacagggtaatattctctagacatcattaattcctaattttttg	Amplify BSD cassette with I-SceI overhangs
ISceIBSDRev	catagaaggcacagtcgaggattaccctgttatccctattagccctcccacacataac	Amplify BSD cassette with I-SceI overhangs
1HAFor	cgttgctgcgccaacaccacctcctcctcctcctttcatctaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac148 homology arms
1HARev	ttttgcaaaacctttactgtgatatatttattaaataagtttgctttaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac148 homology arms
2HAFor	gaagatttgaaaaaaaaactcattttaaagcaaacttatttaataaatatataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac149 homology arms
2HARev	ttttagatttcaaaacagtaaaaataattattgaaaacaaacaaactattcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac149 homology arms
3HAFor	gatatctctccattgtatgatcgctaaatttatataccgtttcaataaattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac150 homology arms
3HARev	aaattttatttcattaatctatataaattttattttatttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac150 homology arms
4HAFor	aggatcatatttagttgcgtttatgagataagattgaaagcacgtgtaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac1 homology arms
4HARev	caatttgaaaataggataaagtatttgttaaagaatgcaaataatattaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac1 homology arms
5HAFor	tcaagatttattaatttaattaatattatttgcattctttaacaaatacttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac2 homology arms

5HARev	tttgcgttaaaaaacaatagttgttgcaaaatcgacacacgtgtatcaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac2 homology arms
6HAFor	acacgtgtgtcgattttgcaacaactattgtttttaacgcaaactaaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac3 homology arms
6HARev	ggcgtcataactatttattaaataagataatttaaaaaatcgccgttaatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac3 homology arms
7HAFor	aatttgttacataatcaaattaccagtacagttattcggtttgaagcaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac15 homology arms
7HARev	cctctcaaaaaaaccaggtaaataatggtgaataacatttattgacgtaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac15 homology arms
8HAFor	tttatttttaaatttaaataaatgtattgttagaaaattgtgttgttttataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac18 homology arms
8HARev	ttgcgtgcgcatttgaacagtgtagtcgtgttgcgtcgccgctattcattcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac18 homology arms
9HAFor	taatataaggtaggacgccaagatttagttgattcaacaaacgttccatataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac19 homology arms
9HARev	tataaaaaataaaacaataattagaactgaacaagtttatttatttatttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac19 homology arms
10HAFor	actaaattaatactttatataatcaaaaataatactttatataatacttttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac20 homology arms
10HARev	tgtacagcgtgcccaagaaagttttgtgtacctcaccggacgcgtgcacacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac20 homology arms
11HAFor	tcttgggcacgctgtacaccggcttgggcacgctatatgtgttgccaaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac21 homology arms
11HARev	agtactcgatacatattgccattacttagttgttgaattaaatattcaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac21 homology arms
12HAFor	attttatttaacattttgaatatttaattcaacaactaagtaatggcaattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac22 homology arms
12HARev	actaactactgacgtaacatttaaatgacttattgtttgtgttatcacaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac22 homology arms
13HAFor	aacaataagtcatttaaatgttacgtcagtagttagtatataagccgtactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac23 homology arms
13HARev	aaataaaaaacattttattaaataatatatttatttcaattatacatgttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac23 homology arms
14HAFor	ataaaataattatttaaataaaaatgtttttattgtaaaatacacattgataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac32 homology arms
14HARev	caaatttttatataaaagcgccagtttgctgtggggacttaaagttcatccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac32 homology arms
15HAFor	agattagttaataatttgtgtgattagataactttttagggtattgcgcataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac33 homology arms

15HARev	tggacacaaacgcgtacgaattgatggactttgagtaagatgcattcacccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac33 homology arms
16HAFor	cgcggcgacaactgcttgtataatattgcccaatgtaaacatgcaaaatttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac39 homology arms
16HARev	gcttacggatttttttgattattgcctgccaaaatactaccgacgcaaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac39 homology arms
17HAFor	acgagcgcgcaataaccatagtttaatcgaagagaatagccgtcgccacataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac42 homology arms
17HARev	cgtcgcaaacatagcaagtagcatatcgggtgttcatgacgacttaccagcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac42 homology arms
18HAFor	caattacgaagacattaaaaaatttttcaacttgtagctggtaagtcgtctaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac43 homology arms
18HARev	agtcgtacagcggaggtagcgagcgggccaaaactcgaagcgcttgcgtccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac43 homology arms
19HAFor	aattggcaacggcgtttcgattccaatttacccccaccgcgctcaacaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac44 homology arms
19HARev	ccggcgacaagtaaaccaacacggcaatcaccatttcgtagggcaacatgcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac44 homology arms
20HAFor	taaatgtaaaagcggcattagatataaattaaaatacaaaactttgtaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac45 homology arms
20HARev	atttgataggtacaaaaaacatattaaaaatattacaactatgacaataacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac45 homology arms
21HAFor	gtgtgtttacaattaagaaaaatttgtattaataataagcaacattcgactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac46 homology arms
21HARev	ttttgtgttaacagtcagtcttttgaatatcacattctccctaaatataacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac46 homology arms
22HAFor	atattgttgttgttacaaaatagataataaaacaaaaaataaattaaatataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac47 homology arms
22HARev	aacgatatatatgcaacaaccttttaatattcacattaaaaagtgccaggcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac47 homology arms
23HAFor	gacgcaaattccatgcagcgcatccatttcgtactgcaaatgcggcgcaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac49 homology arms
23HARev	gtaaaattagttgtatcaaagagcagctgcaattagaatcactgctaaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac49 homology arms
24HAFor	cggcaagcgagattacacgaaatttgacaaattggtgaaaaaatattgactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac56 homology arms
24HARev	aatatctagcgtaaataaataaaaccacaaaatacaaacaacaaactttacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac56 homology arms
25HAFor	agcggtttatttcactcggaggtattatatcaggcagtcgaacgtgcgcgtaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac57 homology arms

25HARev	aagatgaggacgaagataattattacaactattaataaattaat tttttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac57 homology arms
26HAFor	tggatcgtgaactagaaataaaactatacaagtagaaaaaaa ttaatttataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac58 homology arms
26HARev	acgaaaacactagcacggattatgacactgaagacgacgag cgcgtcaaccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac58 homology arms
27HAFor	tttcgtagaattgtccgcgttatcgtcgctttcgtcgatgtcgtcc ctaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac59 homology arms
27HARev	gaaaccgaatacagtacaaccgaagacgaacttgcttaattct aattacacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac59 homology arms
28HAFor	cattttttcattgtataacatatcgggaatttgatacattgtaatta gaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac60 homology arms
28HARev	tgtgataaataaaacacaaacacaaagtttttatattgtcttttat tgatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac60 homology arms
29HAFor	ttcttacattacgagattcaacttgatactaaaattaattattaatt aaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac61 homology arms
29HARev	taaaaaatatattttacgcaccatatacgcatcgggttgatata gttaatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac61 homology arms
30HAFor	gcatgcaatcagtgttttattaattttagagcaacatgtacgata aattttaatacgactcactatagggtaggg	Amplify BSD cassette with gene 63 homology arms
30HARev	tattcgtcatttttttaatccaggattttaataaaacaaacaaaatt ttaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac63 homology arms
31HAFor	ttattacttcttgctaaatataatatgtaattaaaattttgtttgtttt ataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac64 homology arms
31HARev	attacgacagcggttattattccgattagtgttctataagtataat catacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac64 homology arms
32HAFor	gttggccgtctttttcgacacaaacattgcgaccgacacgaag agtaaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac69 homology arms
32HARev	tccatgttgaatgatgcctgatttgaagccgctggctttttataca cacgcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac69 homology arms
33HAFor	tataacgtgtgtataaaaagccagcggcttcaaatcaggcatc attcaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac70 homology arms
33HARev	aattgcatcaaattcattttatagtttattccataacaatatcctta aaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac70 homology arms
34HAFor	agttaataaatgtttaattttaaggatattgttatggaataaactat aaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac71 homology arms
34HARev	atatttatatatagcttatttttaaagaatgtcgtagtcgtttgcaa tgtcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac71 homology arms
35HAFor	aagcgagttgaattttgattacaaatattttgttttaatacgactc actatagggtaggg	Amplify BSD cassette with gene ac84 homology arms

35HARev	taatgtcaaacaaacaacattatccatgtaataaaacaatgtgtaaattgcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac84 homology arms
36HAFor	tttgtttgtcgacggtcgaggggtcagcggcgtgtgcaacaataaaaaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac85 homology arms
36HARev	tacttggaattttaaataaaaaaattttataaaaaaatatattcactttatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac85 homology arms
37HAFor	tggaaatggtgttaaataaaagtgaatatatttttttataaaattttattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac86 homology arms
37HARev	ctttatataatagacagtttagcattggtatgcattattaaacagcaatccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac86 homology arms
38HAFor	tatataaacgcaatgtaacagtttttgctagtaccatcgcatacaacatttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac87 homology arms
38HARev	gtataaccaacgccactattaccacaaatgttacaataaatgtagattaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac87 homology arms
39HAFor	aaattgaagaatattattacattaaagtagaatgttatgtaaacagttaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac88 homology arms
39HARev	cagattgactataatacccgtattagaactagacacgtgcaaacatcaatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac88 homology arms
40HAFor	ttgtaattattgaataataaaacaattataaatgctaaatttgttttttataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac7 homology arms
40HARev	tattgatagacatttccagtttgtgatattagtttgtgcgtctcattacacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac7 homology arms
41HAFor	gattacatgtttgctttgcatcaaaatggcgtaattaattaagaaacaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene 31 homology arms
41HARev	gagttagaacgatgacatcatttacatatggcaattattccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac31 homology arms
42HAFor	tccggacgagattgctagtggtaaacgtttaataattaaacatttgcaagtaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac17 homology arms
42HARev	ttttttattaatattataatttttatctacctttataaattttactacacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac17 homology arms
43HAFor	ccaaaacgacgcaaacgtgctcattttgaccgtaactttaacgagctaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac27 homology arms
43HARev	atttttttccacataatagccgttgtagtacacgttgaacaccattttacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac27 homology arms
44HAFor	gtgtcgtcaggacgtcaccgattttataaaaatatttgtggtgtaataaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac28 homology arms
44HARev	attatatattacaaaatatagttttattattattatttttaatacatgttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac28 homology arms
45HAFor	tgtattaaaaataataataataaaactatattttgtaatatataatgtattttattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac29 homology arms

45HARev	ctaatttggataaaagatgctgttttacctcatctgtatgttttctc aagcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac29 homology arms
46HAFor	gagaaaacatacagatgaggtaaaacagcatcttttatccaa attaggagtaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac30 homology arms
46HARev	gccgaccaaaacggtcaccttgcttcaaacagcgtggagctt cagtcgaccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac30 homology arms
47HAFor	caagggcgcattcacagcaaccgttgtcatttataagtaatagt gtaaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac35 homology arms
47HARev	ttaaataaaaaacttttatgtatatttaatggtttttattattgttatt acatagaaggcacagtcgagg	Amplify BSD cassette with gene ac35 homology arms
48HAFor	caataataaaaaccattaaatatacataaaagtttttttattaata cgactcactatagggtaggg	Amplify BSD cassette with gene ac36 homology arms
48HARev	tatcatcaagaatacaaacgggttgtttcaaaggtttacaagaa gcaaaccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac36 homology arms
49HAFor	aagcagtccgaatcacacccgcctaagtgcgtgcaatttttgg ggggcattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac38 homology arms
49HARev	tcaaacgatggtgctcgtgtgcgcaatttttttaagtgtcaaatat gtaaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac38 homology arms
50HAFor	tgcacgcttgcggcaaaaattgtgtggtgtgctccaaccgggtg acaaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac41 homology arms
50HARev	ggttattgcgcgctcgtaatctaagatatcacatttaacacgga cattatcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac41 homology arms
51HAFor	aacgactacgacattctttaaaaataagctatatataaatattg cattgttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac72 homology arms
51HARev	tggtgtattaatttaattacaataaaacattatgtacaataatatg tcttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac72 homology arms
52HAFor	tattgtggtgatcatcgtacaaataatgaagcaacagttcacta tcgtattaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac82 homology arms
52HARev	aaattgccaaatacattaaagtaaacgctattataagaaaaat ggcaagccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac82 homology arms
53HAFor	gacacgacaataaacgtgttgaaacatcgtcgcgaccgaatc gtgccaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac91 homology arms
53HARev	aaatttgcaataatacaacattctttgtttaatttttatatttaagc aaccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac91 homology arms
54HAFor	cgcattaacccaaatacttgaaacttaccccaactactctgtat aacaactaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac120 homology arms
54HARev	atgatgtcatttatttttgcacacagaatattgaacccgttatgat gtcacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac120 homology arms
55HAFor	aagcatttgtatttaatcaatcgaaccgtgcactgatataagaa ttaaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac124 homology arms

55HARev	tttgtttagaatttacagaattgtgaaagaataaaaaacattttta taaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac124 homology arms
56HAFor	cattgtaatggatgacataaaaaataaaaaataatttataaaa atgttttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac125 homology arms
56HARev	aaaagcaatatattgagtatcattttagttttgtgttagttgttatta ggcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac125 homology arms
57HAFor	atttaatgtgcctgaataaataataataagggttttgtacgatttc aacataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac130 homology arms
57HARev	gatgtttgtgacgcttgacgctaaattggtcaaaatagagttggt gttgtcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac130 homology arms
58HAFor	taaaataaaatttatttaatatattatactatttatattatttacaac actaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac151 homology arms
58HARev	cacagttcaagcctcacagcctagtgaacagtatcctaccag cccagcccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac151 homology arms
59HAFor	taagcacagttcgttgtgaagtgaacacggagagcctgccaat aagcaaataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac153 homology arms
59HARev	aaatgtttattatttattccatcatttttatatatacatttcatatcac acatagaaggcacagtcgagg	Amplify BSD cassette with gene ac153 homology arms
60HAFor	tttatactttttatgtttttttatttcatgtgattaagaaacttttaagt aatacgactcactatagggtaggg	Amplify BSD cassette with gene ac154 homology arms
60HARev	gagggataatgtcatttgttattctcaaatatgatgtcatcataga aggcacagtcgagg	Amplify BSD cassette with gene ac154 homology arms
61HAFor	tattaacggcgatttttaaattatcttatttaataaatagttatga cgctaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac4 homology arms
61HARev	taattcttgacgtgcgatacataatatattattggataaagttgca ttaacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac4 homology arms
62HAFor	ccatcttgtaagttagtttcatttaatgcaactttatccaataatat atttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac5 homology arms
62HARev	atgtagctttcttgtctaaacatgacgctctaatgagcggactcc acacgcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac5 homology arms
63HAFor	acgcaaactacaatgttatataataaaaaattataaaatattttt aattttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac105 homology arms
63HARev	tttgaggtatataaggcttgatacaggcacagtaagcagcaca ttctaaccatagaaggcacagtcgagg	Amplify BSD cassette with gene ac105 homology arms
64HAFor	cgtcaatgaaacaacgtgtaatatttttttacaatatttaagtgaa acatttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac106/107 homology arms
64HARev	ctactttattagaaaactatgatattagaaatagaaatgcaaca atataacatagaaggcacagtcgagg	Amplify BSD cassette with gene ac106/107 homology arms
65HAFor	aaacgggcactccatcgtgatgtatatgtcttacttcattagact ttagataatacgactcactatagggtaggg	Amplify BSD cassette with gene ac110 homology arms

65HARev	aaacgtataaaagtactatgtattttattcaactagcataagatttaaagcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac110 homology arms
66HAFor	atctttaaatcttatgctagttgaataaaatacatagtacttttatacgttaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac111 homology arms
66HARev	gcaataaaaaattaaagcatatgcatcagtgacattgcctgtgctgcactcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac111 homology arms
67HAFor	tataaaagtgtttgaattgcgagaccgtcaacataacgtttatcaacgcgtaatacgactcactatagggtaggg	Amplify BSD cassette with gene ac112/113 homology arms
67HARev	tatcaaatgtgagtattgtcaataaaaacgcttgtttaaattaatttgttcatagaaggcacagtcgagg	Amplify BSD cassette with gene ac112/113 homology arms
BSDCheckRev	cggccaccacggatcagaattacggag	BSD cassette integration PCR check
1CheckFor	cgcgcaaaccacatttgtgactttgttacaatcg	Gene ac148 BSD cassette integration PCR check
2/3CheckFor	gccctgaaatactccgcgtatgttggggac	Gene ac149 and ac150 BSD cassette integration PCR check
4/5CheckFor	cagcttatgactcaagttatgagccgtgtgc	Gene ac1 and ac2 BSD cassette integration PCR check
6/61CheckFor	gcggctgcaaatacagcgggtcg	Gene ac3 and ac4 BSD cassette integration PCR check
7CheckFor	cggttgccggtatacatgactcttgacacg	Gene ac15 BSD cassette integration PCR check
8CheckFor	gccggacgatgaaggtacaacaatgcc	Gene ac18 BSD cassette integration PCR check
9CheckFor	ccaacggatccgacatgtaaacttgacgc	Gene ac19 BSD cassette integration PCR check
10/11CheckFor	gaccgcgtgcaatttgatcaactcgttg	Gene ac20 and ac21 BSD cassette integration PCR check and ac20VloxP modification in baculoviral genome
12CheckFor	catctagcgccggtagatgaccatttttgaacac	Gene ac22 BSD cassette integration PCR check
13CheckFor	gtcacgcgtgttacgcacattgtgc	Gene ac23 BSD cassette integration PCR check
14CheckFor	ccgatcatccgttgagcaaaacaaccgg	Gene ac32 BSD cassette integration PCR check and ac32Tn7 modification in baculoviral genome
15CheckFor	cacgtgatggcgttttgaatgacgatgcg	Gene ac33 BSD cassette integration PCR check
16CheckFor	gcgtcgcaacaatcgcgatgacctc	Gene ac39 BSD cassette integration PCR check

17CheckFor	cgaccagtttgagcatccaaatccggc	Gene ac42 BSD cassette integration PCR check
18/19CheckFor	gtggaagacaggattttggccgagacgac	Gene ac43 and ac44 BSD cassette integration PCR check
20CheckFor	caacatgcacgacggtcgcgttaagac	Gene ac45 BSD cassette integration PCR check
21CheckFor	gtgtgctgtttgccagcaactctgc	Gene ac46 BSD cassette integration PCR check
22CheckFor	gtcgaaggtcattcggattccctgagagc	Gene ac47 BSD cassette integration PCR check
23CheckFor	gcaaataataatctagtcgttgggcgacctcgc	Gene ac49 BSD cassette integration PCR check
24/25CheckFor	cgcgaaaacgtaaacaaccagctacacgc	Gene ac56 and ac57 BSD cassette integration PCR check
26/27CheckFor	ctttcacattacttgtgatctgtcgcttgaaacgc	Gene ac58 and ac59 BSD cassette integration PCR check
28/29CheckFor	catcttcgctggtgtcataatgcggtgg	Gene ac60 and ac61 BSD cassette integration PCR check
30/31CheckFor	gcggtgctaaaggcagtcaaaatctgattaaaagc	Gene ac63 and ac64 BSD cassette integration PCR check
32CheckFor	ggtcgccatattgctgttgtcgatatgtggg	Gene ac69 BSD cassette integration PCR check
33CheckFor	ggacgcgaaaacgaacaagaacgtctcaac	Gene ac70 BSD cassette integration PCR check
34CheckFor	gagcgtttgaagatgcagtcgctgtacg	Gene ac71 BSD cassette integration PCR check
35CheckFor	gcgtgtcccgaagaattgtacgataacaacg	Gene ac84 BSD cassette integration PCR check
36CheckFor	ctttgctacaaagaaaagtgccgcagcg	Gene ac85 BSD cassette integration PCR check
37CheckFor	ccaattgtgatcggggtcatgtcgttgaagag	Gene ac86 BSD cassette integration PCR check
38/39CheckFor	ctggcatgccattcgccgttgtacttg	Gene ac87 and ac88 BSD cassette integration PCR check and ac87RoxP and ac87Tn7 modification in baculoviral genome
40CheckFor	caatgtaccgcgcggcggtatg	Gene ac7 BSD cassette integration PCR check
41CheckFor	gctcagacaccaacgaccagatcaaatgtg	Gene ac31 BSD cassette integration PCR check

42CheckFor	gccgagttggttagaaaacttgaaagtgcacag	Gene ac17 BSD cassette integration PCR check
43CheckFor	gtgcgctgtctgctgtggttcaagtctaattg	Gene ac27 BSD cassette integration PCR check
44CheckFor	cacgggtcgcggagacgaaactg	Gene ac28 BSD cassette integration PCR check
45/46CheckFor	cactgcaacggcagatactactggcc	Gene ac29 and ac30 BSD cassette integration PCR check
47/48CheckFor	cgtgcgtgtaacaaagtcgacagcgaaac	Gene ac35 and ac36 BSD cassette integration PCR check
49CheckFor	gctggtccatgttggatttgttgtaagagctcg	Gene ac38 BSD cassette integration PCR check
50CheckFor	gacccaacgtgtagtgatcttcgtaggcg	Gene ac41 BSD cassette integration PCR check
51CheckFor	cccaacgccggcacaaacaaaattgc	Gene ac72 BSD cassette integration PCR check
52CheckFor	ggtcaacagcaccgtttgaaaactgacgc	Gene ac82 BSD cassette integration PCR check
53CheckFor	caacattgtggcctttcaatgcgaggttatgg	Gene ac91 BSD cassette integration PCR check
54/55CheckFor	cgttgtgagccagcgatccttgattggttttaatc	Gene ac120 and ac124 BSD cassette integration PCR check
56CheckFor	ccacgccatgtactttagtaacgtgctgttg	Gene ac125 BSD cassette integration PCR check
57CheckFor	caccttttcttagcaacggcagttcggc	Gene ac130 BSD cassette integration PCR check
58CheckFor	gcgcgcgttgatacggatgatatctctcc	Gene ac151 BSD cassette integration PCR check
59CheckFor	gaggcttgaactgtgcttacgagtagaacgg	Gene ac153 BSD cassette integration PCR check
60CheckFor	cacagtgatcaagtagcttctgaaaaccgacagc	Gene ac154 BSD cassette integration PCR check
62CheckFor	cacgcgacgcacaagtatctgtacacc	Gene ac5 BSD cassette integration PCR check
63CheckFor	gaagacgagcaggaagatgatgacgttgaagatg	Gene ac105 BSD cassette integration PCR check
64CheckFor	catcgttgcccgattcatcttttgcatgtcg	Gene ac106/107 BSD cassette integration PCR check

65/66CheckFor	gatgggacatttggggttttcatggtgaacg	Gene ac110 and ac111 BSD cassette integration PCR check and ac110Tn7 modification in baculoviral genome
67CheckFor	gaacacgtgcgttgttcactacgattgaaatgc	Gene ac112/113 BSD cassette integration PCR check and ac112/113Tn7 modification in baculoviral genome
ac20VloxPFor	gttttcccagtcgagctcgatatcac	Amplify ac20VloxP fragment from pDS_ac20VloxP
ac20VloxPRev	ctgtacagcgtgcccaagaaag	Amplify ac20VloxP fragment from pDS_ac20VloxP
ac20VloxPFragSeq1	gatcgttgtcagaagtaagttggc	Sequencing primer: ac20VloxP, ac87RoxP fragment and ac32, ac87, ac110 and ac112/113 Tn7 fragment
ac20VloxPFragSeq2	ggttgagtactcaccagtcacag	Sequencing primer: ac20VloxP, ac87RoxP fragment and ac32, ac87, ac110 and ac112/113 Tn7 fragment
Check10Rev	gtgcttacgatactggctgtgcaatgctg	Validate ac20VloxP modification in baculoviral genome
ac20VloxPBacSeq1	cttttatcgcgccttcggac	Sequencing primer: ac20VloxP and ac87RoxP fragment in baculoviral genome
ac20VloxPBacSeq2	cgacggggttacgattcaac	Sequencing primer: ac20VloxP and ac87RoxP fragment in baculoviral genome
PVloxPmcherry1	gaattagcttcacgctgccg	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry2	ctgggccaacttttggcg	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry3	gcttccttagctcctgaaaatctc	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry4	gcattcattttatgtttcaggttcagg	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry5	ccgtcctcgaagttcatcac	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry6	ccgatggtgggacggtatg	Sequencing primer: pUPDM_VloxP_mCherry
PVloxPmcherry7	caagccctgccatagccac	Sequencing primer: pUPDM_VloxP_mCherry

ac87RoxPFor	ggttttcccagtcgagctcgatatc	Amplify ac87RoxP fragment from pDS_ac87RoxP
ac87RoxPRev	gaaattgttatccgctggtaccgatatc	Amplify ac87RoxP fragment from pDS_ac87RoxP
Tn7FragFor	cgtctcgagactagttccgtttaaac	Amplify ac32, ac87, ac110 and ac112/113 Tn7 fragments from pDS_ac(32, 87, 111, 112/113)RoxPTn7
Tn7FragRev	ctgccaggcacatgggtttaaac	Amplify ac32, ac87, ac110 and ac112/113 Tn7 fragments from pDS_ac(32, 87, 111, 112/113)RoxPTn7
Tn7Seq1	cctatagtgagtcgtattacgcgc	Sequencing primer: ac32, ac87, ac110 and ac112/113 Tn7 fragment
Tn7Seq2	ggggatgtgctgcaaggc	Sequencing primer: ac32, ac87, ac110 and ac112/113 Tn7 fragment
Tn7Seq3	gccgctctagaactagtggatcc	Sequencing primer: ac32, ac87, ac110 and ac112/113 Tn7 fragment
ac32Tn7CheckRev	gacgcggctggccgaattaatctaacc	Validation of ac32Tn7 modification
ac87Tn7CheckRev	gacattgccgaacaaacgcagctcaac	Validation of ac87RoxP and ac87Tn7 modification
ac110Tn7CheckRev	gccgacaacgcttcacgatggtaacatc	Validation of ac110Tn7 modification
67CheckRev2	ggtcgtgtgcaaacacaacgccg	Validation of ac111/112Tn7 modification
YFPCheckFor	cacgcgttgttaatcccacgacttttgc	Validation of EYFP insertion into LoxP site in the baculoviral genome
YFPCheckRev	gcaagatctatacgatccagagcgctacgc	Validation of EYFP insertion into LoxP site in the baculoviral genome

8.4 AcMNPV genes

The following compiled tables (Table 8.3 – 8.10) are based on the information found in the Baculovirus Molecular Biology Book, 3rd edition, Chapter 12: The AcMNPV genome: Gene content, conservation, and function.[358]

Table 8.3 – Planned gene deletions/retentions within Fragment 1.

Fragment 1		
Ac114	Retained	Homologues found in all group I NPV genomes. ODV and BV associated protein.[359]
Ac115/pif-3	Deleted	Per os infectivity factor. ODV-associated protein. Not required for *in vitro* infection.[310]
Ac116	Deleted	BmNPV (Bm95) deletion mutant showed no defects.[285]
Ac117	Deleted	BmNPV (Bm96) deletion mutant produced infectious viruses in cultured cells but with minor motility detriment in larvae.[285]
Ac118	Deleted	Homolog only found in three group 1 NPVs (Ac-, Ro- and PlxyNPV).[360]
Ac119/pif-1	Deleted	Per os infectivity factor important for oral infectivity and ODV midgut binding. Deletion mutant infectious *in vitro* but not *in vivo*.[39]
Ac120	Retained	Homologs found in all group 1 and most group 2 genomes. It is likely to be dispensable, as an insertion/deletion mutation of this gene in BmNPV (Bm98) had no apparent consequence on infectivity.[285]
hr4c	Retained	Homologous region –early transcription enhancer.
Ac121	Deleted	Ac121 homologs only found in BmNPV and PlxyNPV. A BmNPV knock-out appeared viable.[285,361]
Ac122	Deleted	Homologs present in most Group I NPVs. A BmNPV insertional mutant had no effect on infectivity.[285]
Ac123/pk-2	Deleted	Protein kinase, phosphotransferase. Deletion mutant infectious in cultured cells.[362]
Ac124	Retained	Transcriptional regulator for chitinase. Deletion mutant had no effect on BV or OBs. A BmNPV deletion virus was viable, thus appeared to be a non-essential gene.[285,338]
Ac125/lef-7	Retained	Late expression factor involved in host DNA damage control in order to promote viral replication. When deleted, infection was unaffected in Tn368 cells, but in Sf21 and Se1c cells DNA replication was 10% of wt. Similar results were seen in a BmNPV lef-7 deletion mutant. In AcMNPV, the deletion resulted in the accumulation of phosphorylated H2AX and activation of the DNA damage response pathway that led to a major reduction in late gene expression and reduced infectious virus production by 100-fold.[301,363]
Ac126/ChiA	Disrupted	Host liquification and horizontal transmission. Regulates v-cath. Already disrupted in MultiBac - replaced N terminus with AmpRLoxP.[364–366]
Ac127/v-cath	Deleted	Cysteine protease involved in host liquification. Already disrupted in MultiBac – AmpRLoxP at N terminus.[365,366]
Ac128/GP64	Retained	Entry of BV into and out of cells. Deletion is lethal, only replicated in a single cell and cannot bud out and infect surrounding cells.[367]
Ac129/p24	Retained	Capsid associated protein. It seems to be a non-essential gene in BmNPV.[285,368]
Fragment 1 (CONT.)		

Ac130/GP16	Retained	In OpMNPV, the ac130 homolog (Op128) is glycosylated and localized near the nuclear membrane in the cytoplasm. Although it appeared to be associated with envelopes of nucleocapsids in the cytoplasm, it was not associated with either ODV or BV. It seems to be a non-essential gene in BmNPV.[285,369]
Ac131/PEP	Deleted	Major OB calyx protein. Seals surface of polyhedra and provides additional stability. [49,370]
Ac132	Retained	ODV and BV associated. Role is progeny virus production.[371]
Ac133/alk-exo	Retained	Found in all baculoviral genomes. Alkaline with 5'-3' exonuclease and endonuclease activity. Deletion was shown to be lethal.[372]
Ac134/p94	Deleted	Knock-out showed no effect on ability of AcMNPV to infect larvae or cultured cells, thus deemed non-essential.[373]
Ac135/p35	Retained	Apoptosis inhibitor. Deletion mutants showed severely compromised BV production in Sf21 cells. [374–376]
hr5	Retained	Homologous region – early transcription enhancer.
Ac136/p26	Deleted	Nuclease cleaves host 2'-3'-CGAMP. AcMNPV deletion mutant showed no differences from wt in both cultured cells and larvae.[377]
Ac137/p10	Deleted	Required for proper formation of polyhedra envelope in ODV and nuclear lysis. Deletion prevents cell lysis, thus elongating protein production. [378]
Ac138/p74-pif	Deleted	Per os infectivity factor. Oral infectivity of larvae and midgut attachment. Deletion mutants infectious *in vitro* but not in larvae. [379]
Ac139/ME53	Retained	BV and ODV associated. AcMNPV deletion mutant fails to replicate its DNA and does not produce nucleocapsids.[380]

Table 8.4 – Planned gene deletions/retentions within Fragment 2A.

Fragment 2A		
Ac147/IE1	Retained	Trans-activation of early promoters. Deletion mutant not viable.[381,382]
Ac148/Odv-e56/pif-5	Deleted	Per os infectivity factor localised to envelopes of ODV. Insertion mutant was infectious.[310,311,383,384]
Ac149	Deleted	BmNPV deletion mutant produced viable virus.[285]
Ac150	Deleted	Related to Ac145. AcMNPV deletion mutant showed lesser infectivity by ODV in larvae, but no difference in BV production.[385]
Ac151/IE-2	Retained	Deletion mutant resulted in delayed infection. IE-2 contains a predicted RING finger domain and was required for optimal origin specific plasmid DNA replication in Sf-21 cells but had little effect in Tn-368 cells. IE-2 deletion mutants worked differently in Sf-21 cells in which the infection was delayed vs. Tn-5B1-4 cells, in which the infection was unaffected. In insect larvae, the mutant viruses were significantly less infectious than wt, which seemed to be due to a lack of virions in the occlusion bodies.[301,302]
Ac152	Retained	Associated with the nuclear localisation of G-actin. Deletion resulted in decreased BV titres.[386]
Ac153/Pe38	Retained	Found to activate DNA replication in transient assay. Deletion mutant displayed a delay in DNA replication, 99% reduction in BV production and reduced oral infectivity. [298–300,387]
Ac154	Retained	Knock-out did not affect BV production but progeny virus was reduced. Stimulates anti-apoptotic effect on host cells.[388]
hr1	Retained	Homologous region – early transcription enhancer.
Fragment 2A (CONT.)		

Ac1/ptp	Deleted	RNA 5'-triphosphatase, hydrolyses ATP to ADP and GTP to GDP. Deletion mutant showed small effect on ODV production but otherwise viable.[389]
Ac2/Bro	Deleted	Baculovirus Repeated ORF. Knock-out mutants are viable but differences in polyhedron production in infected cells was detected.[285,390]
Ac3/ctl	Deleted	Deletion mutant performed as wt virus.[391]
Ac4	Retained	May be involved in nuclear localisation of G-actin from transiently transfected results.[386]

Table 8.5 – Planned gene deletions/retentions within Fragment 2B.

Fragment 2B		
Ac14/lef-1	Retained	Late expression factor. Core gene found in all baculoviruses. DNA primase activity. Deletion mutant could not be isolated.[392]
Ac15/egt	Deleted	Ecdysteroid UDP-glucosyltransferase. Prevents molting and pupation in infected larvae. Knock-out mutants are viable. Deletion of Ac15 spontaneously occurs in cell culture.[210,248,393,394]
Ac16/bv/odv-e26	Retained	Found in all Group I NPVs. Deletion mutant displays a delay in DNA synthesis and reduced BV titers.[395]
Ac17	Retained	Found in most Group I and II NPVs. Ac17 deletion resulted in reduced BV titers.[51]
Ac18	Deleted	Homologs found in all Group I and II NPVs. Deletion did not affect infectivity.[396]
Ac19	Deleted	Found in all Group I and most Group II NPVs. Deletion mutant was viable, but OB and ODV production was reduced.[285]
Ac20/arif1 ct	Deleted	Actin-Rearrangement-inducing factors. C-terminal deletion or LacZ gene insertion resulted in viable viruses.[312]
Ac21/arif1 nt	Deleted	
Ac22/pif-2	Deleted	Per os infectivity factor. Required for oral infectivity and not in cell culture.[39]
Ac23/F protein	Deleted	Homologs found in most baculoviruses. Knock-out mutants are infectious *in vitro* and *in vivo*, but larvae killing is extended.[397,398]
Ac24/PKIP	Retained	Protein kinase interacting protein. Homologs found in all Group I and II NPVs. Deletion mutants could not be isolated in AcMNPV and show a decreased infection spread in BmNPV.[399]

Table 8.6 – Planned gene deletions/retentions within Fragment 3.

Fragment 3		
hr2	Retained	Homologous region – early transcription enhancer.
Ac32/fgf	Deleted	Homology to fibroblast growth factor (fgf). AcMNPV fgf is secreted and stimulates insect cell motility. Although *fgf* deletion in AcMNPV showed no changes from wt in cultured cells, the time of larvae death was delayed in two insect species.[306-308]
Ac33/hisP	Deleted	Polynucleotide kinase. Ac33 also shows structural similarity to histidinol-phosphatase (an enzyme in the histidine biosynthesis pathway). No homologs in BmNPV or HearSNPV.[400]
Ac34	Retained	Orthologs found in all Group I and II NPVs. AcMNPV deletion mutant resulted in delayed late gene expression and reduced BV titers, but not DNA synthesis.[401,402]
Fragment 3 (CONT.)		

Ac35/v-ubi	Retained	Viral ubiquitin. Interacts with ac66 and ac141 suggesting involvement in signal for BV formation and exit from nuclei but no influence on viral replication. Deletion mutant experienced reduced BV titers.[285,403]
Ac36/39K/ pp31	Retained	Homologs found in all NPVs. Capable of binding DNA and stimulated late gene expression in transient transcription assay. Deletion mutants showed a reduction in late gene transcription, BV production and improper formation of the virogenic stroma.[282-284]
Ac37/lef-11	Retained	Late expression factor required for late gene expression. Knock-out virus failed to replicate.[404]
Ac38	Retained	Homology to ADP-ribose pyrophosphate. Deletion mutant produced BVs at 1% to wt. A BmNPV deletion for this gene produced a viable virus but the protein expression was reduced.[285,295]
Ac39/p43	Deleted	Deletion had no effects on growth curves or virus production.[405]
Ac40/p47	Retained	Subunit of the baculovirus polymerase. Essential gene and deletion mutants could not be isolated, and mutations caused defects in late gene expression.[13]
Ac41/lef-12	Retained	Late expression factor. Interruption or mutation of ATG start codon of *lef-12* resulted in viable viruses in both *S. frugiperda* and *T. ni* cells, although reduced yields of BV were observed (20-40% of wt) and the infection cycle appeared to be slowed. A BmNPV Knock-out mutant was viable, but decreased protein expression was observed.[285,291]
Ac42/gta	Deleted	Homologous to 'global transactivator', the DEAD-like helicase superfamily that are enzymes involved in ATP-dependent unwinding of DNA or RNA. Deletion of the Ac43 homolog from BmNPV (Bm33) did not cause any defects in BV or ODV production in BmN cells. However, it took longer to kill larvae in *B. mori*.[406]
Ac43	Deleted	Involved in late and very late gene expression. Deletion of BmNPV homolog (Bm34) resulted in reduced occlusion body production and prolonged killing of larvae. Ac43 mutant did not appear to affect BV production.[407,408]
Ac44	Deleted	Shows homology to a 64.2 kDa inhibitor of apoptosis/RING-finger protein in the genome of *Spodoptera frugiperda* ascovirus 1a. Ac44 may be dispensable as a BmNPV homolog (Bm35) knock out appeared normal.[285,409]
Ac45	Deleted	Sequences located within Ac45 appeared to be required for Ac41 expression in a transient late transcription assay. A BmNPV homolog (Bm36) deletion showed no difference between wt.[285,410]
Ac46/odv-e66	Deleted	Component of ODV envelopes. BmNPV deletion homolog (Bm37), was viable, but took more time to kill insect larvae and <30% less expression as wt was seen at 96 hpt.[410,411]
Ac47	Deleted	27% identity to homologs of a protein called translin-associated factor X (Trax). TRAX interacts with translin, which might be involved in responses to DNA damage, transport of RNA, and control of translation, but its function in AcMNPV is not known. BmNPV homolog (Bm38) knock out appeared to be normal.[285,303]
Ac48	Retained	Found in most genomes of Group I NPVs. Function unknown.[409]
Ac49/pcna	Deleted	Ac49 shows homology to proliferating cell nuclear antigen (PCNA). In AcMNPV it is not an essential gene and did not seem to elevate DNA replication in transient replication assays.[304,305]
Ac50/lef-8	Retained	Late expression factor. Component of the late baculovirus RNA polymerase. Deletion mutant was not viable.[13]

Table 8.7 – Planned gene deletions/retentions within Fragment 4.

Fragment 4		
Ac55	Retained	Homologs found in all Group I and most Group II NPVs. On further research, it is likely non-essential as a BmNPV knock-out virus was viable.[285]
Ac56	Deleted	Likely non-essential as a BmNPV mutant deleted for this gene appeared to be normal.[285]
Ac57	Deleted	It appears to be nonessential as a BmNPV mutant deleted for this gene displays no difference between wt. It also appeared to be non-essential for HearSNPV.[285,412]
Ac58	Deleted	These two orfs are likely a single gene, as homologs are fused in other baculoviruses and they were also found to be joined when the region was resequenced in the C-6 strain. It localized to nuclei of infected cells and was associated with AcMNPV ODV and BV. Likely nonessential because a BmNPV mutant deleted for this gene appeared to be normal.[22,413,414]
Ac59	Deleted	
Ac60	Deleted	Ac60 is similar to Ac58/59. It appears to be dispensable because a BmNPV mutant deleted for this gene behaved normally.[285]
Ac61/FP25	Deleted	BV associated and its deletion showed to be not lethal but resulted in a 'few polyhedra phenotype'. FP mutants are defective in virion occlusion and nucleocapsid envelopment in nuclei and release two-to five-fold more infectious BV than wt in infected Sf9 cells. However, protein expression could be impaired due to involvement in polh and p10 expression.[26,27]
Ac62/lef-9	Retained	Late expression factor. Baculovirus RNA-polymerase subunit. Essential gene as a deletion mutant could not be isolated.[415,416]
Ac63	Deleted	Associated with BV envelopes. Likely nonessential as BmNPV mutant for this gene functioned as wt.[417]
Ac64/gp37	Deleted	In AcMNPV it is expressed in the late phase of replication. The gp37 homolog in SpltNPV and CpGV has been reported to contain chitin binding domains and is capable of binding to chitin and enhancing per os infections. GP37 was reported to be polyhedron associated in AcMNPV and to be N-glycosylated. It was shown that the *gp37* gene is nonessential for replication in cell culture or *T. ni* larvae. Similar results were observed for the BmNPV homolog (Bm52).[285,418-422]
Ac65/dnapol	Retained	DNA polymerase. Essential as deletion prevents viral replication.[423]

Table 8.8 – Planned gene deletions/retentions within Fragment 5.

Fragment 5		
Ac67/lef-3	Retained	Late expression factor essential for DNA replication. Deletion mutants are not viable.[392,424]
Ac68/pif-6	Terminator disrupted	Per os infectivity factor. Deletion mutants experienced abnormal polyhedra and few virions but no changes in BV titers were noted.[42]
Ac69/pmt/mtase	Deleted	Methyltransferase that was found to stimulate late gene transcription in a transient assay. AcMNPV and BmNPV mutant of this gene replicated normally in cell culture.[285,425,426]
Ac70/hcf-1	Deleted	AcMNPV hcf-1 mutants replicate normally in both Sf-21 cells and *S. frugiperda* larvae. However, in *T. ni* cells replication was impaired and the mutant showed a significantly reduced infectivity where larvae died more slowly than when infected with wt.[427]
Ac71/iap-2	Deleted	Inhibitor of apoptosis-2 (*iap-2*). *Iap-2* mutants exhibited no effect on viral replication in cell culture; however, the presence of another apoptotic suppressor, *p35,* may have counteracted the effect. In contrast, a BmNPV iap-2 (bm58) null mutant indicated that it was required for replication and <30% less expression was observed than wildtype at 96 hpt.[285,287,428]
Ac72	Retained	Homologs found in most Group I NPVs. Cells infected with a BmNPV null mutant failed to lyse and larvae did not liquify.[285]

Table 8.9 – Planned gene deletions/retentions within Fragment 6.

Fragment 6		
Ac83/pif-8	Retained	Per os infectivity factor required for larvae infection. Additionally, contains cis acting nucleotide sequences required for nucleocapsid assembly and thus deletion mutant did not produce BVs.[429]
hr3	Retained	Homologous region – early transcription enhancer.
Ac84	Deleted	The orf is only found in a few other NPVs and it is not found in the BmNPV or the HearSNPV genomes. No other information found.[409]
Ac85	Deleted	Encodes a 53 aa and is only found in two other NPVs that are AcMNPV variants. Not found in the BmNPV or HearSNPV genomes. No other information found.[409]
Ac86/pnk/pkl	Deleted	Ac86 encodes a protein with RNA ligase, polynucleotide 5'-kinase, and polynucleotide 3'-phosphatase activities and may be part of an RNA repair pathway. Expressed early in infection and appears non-essential when the coding region was replaced with a LacZ gene. (Promoter of the LacZ gene could have rescued the expression of neighbouring genes). This gene is not found in the BmNPV or HearSNPV genomes.[430,431]
Ac87/p15	Deleted	The homolog in BmNPV (Bm70) may encode a capsid protein called p15. It is possibly non-essential, as a deletion mutant in BmNPV (Bm70) appeared normal.[285,309]
Ac88/cg30	Deleted	Ac88 predictions show zinc finger and leucine finger domains. It was found to be associated with AcMNPV, but not in HearNPV ODV. Deletion of ac88 from AcMNPV resulted in only subtle differences from wt, but a BmNPV (Bm71) deletion mutant resulted in a 10 to 100- fold reduction in titer and showed a longer lethal time.[22,288,432,433]
Ac89/VP39	Retained	Ac89 encodes the major capsid protein, present in all baculoviruses. Required for correct DNA synthesis and nucleocapsid assembly. Deletion mutant in BmNPV resulted in no BV production.[434,435]

Table 8.10 – Remaining genes of the AcMNPV baculovirus that were not included within the initial planned fragment design.

Remaining Genes	
Ac5	Homologs found in most Group I NPVs. Deletion in AcMNPV and BmNPV (Bm134) appeared to produce normal virus, however the number of occlude virions was reduced.[40,285]
Ac6/lef-2	Accessory factor for DNA primase (lef-1). AcMNPV and BmNPV deletion mutants were not infectious.[112,301]
Ac7	This gene is only found in three baculovirus genomes in addition to AcMNPV. Deletion from AcMNPV did not affect replication in cell culture or in *T. ni* larvae, however it was only a partial (177 nucleotides) deletion.[313]
Ac8/polyhedrin	Stabilises virions in the environment, allowing them to persist indefinitely. Not essential in cell culture. Contains strong promoter thus locus utilised for recombinant protein production.[436]
Ac9/or1629 pp78/83	Found in all Group I and II NPVs. Involved in nuclear actin assembly thus movement of virions into the cytoplasm of infected cells. Recombinant bacmids did not function without restoration of the gene.[437]
Ac10/PK-1	Protein kinase-1. Deletion mutants produced no viral progeny, although DNA replication was unaffected. Catalytic domain was required for infectivity.[438]
Ac11	Ac11 interacts with the components of ESCRT-III complex and may be involved in nucleocapsid release from nuclear membrane, thus required for ODV and BV production.[439]
Ac12	Contains and F-box domain and interacts with S phase kinase associated protein1. Considered non-essential.[337]
Ac13	Shows predicted coiled coils domains and structural similarity to some membrane proteins. In AcMNPV and BmNPV (Bm5), the deletion mutant viruses seemed normal. However, in BmNPV, lower BV titers and fewer OB were noted.[440]
Ac25/DBP	Single stranded DNA binding protein capable of unwinding and annealing DNA. Ac25 mutant failed to produce virions due to defective nucleocapsids.[441,442]
Ac26	Present in all Group I and most Group II NPVs. A BmNPV homolog (Bm17) is expressed in the late stage and localised to the nucleus and cytoplasm. Bm17 deletion lead to viable viruses but spread of infection was delayed.[285,443]
Ac27/iap-1	Inhibitor of apoptosis. Found in Group I NPVs only, unlike iap-2 that is found in Group I and II NPVs. Iap-1 deletion mutants showed similar replication in cells and larvae compared to the wt. Transient expression in virus induced apoptosis.[444,445]
Ac28/lef-6	Late expression factor required for transcription of late genes. A deletion mutant was infectious but severely compromised.[285,290,446]
Ac29	Show 80% probability to amphiphysin BAR domain from Drosophila, which are sensors of membrane curvature. A BmNPV deletion mutant (Bm20) was viable.[285]
Ac30	Predicted homology to genes encoding tryptophan repeat gene family. A deletion mutant of the ortholog (Bm21) produced normal virus, but the survival time was longer compared to wt.[447]
Ac31/SOD	Ac31 has homology to super oxide dismutase. Conflicting results are seen where in one study, deletion of the *sod* gene from BmNPV (Bm23) indicated that it was essential for replication in BmN cells, however, another report indicated that a Bm23-deleted bacmid was viable.[285,289]

Remaining Genes (CONT.)	
Ac51/DnaJ domain protein	Late gene found in the cytoplasm and nuclei of infected cells. Required for the egress of nucleocapsids from the nuclei, as 1000-fold reduction in BV production was seen upon Ac51 deletion. DNA replication, nucleocapsid assembly and ODV formation was not affected.[285,448]
Ac52	Homologs found in half of the Group I and Group II NPVs. Ortholog (Bm41) deletion resulted in a 1000- fold reduction of BV production due to disrupted nucleocapsid envelopment and polyhedron formation. Survival time in larvae was also prolonged.[449]
Ac53	Homologs found in all baculoviruses and shown to be essential Deletion mutant was capable of DNA replication but produced defective virions.[450]
Ac53a/lef-10	Late expression factor. Named Ac53a as it was not originally identified in the sequencing data of AcMNPV due to small (78aa) orf and overlapping region to Ac54. Required for late gene expression. A BmNPV ortholog (Bm42a) knock-out bacmid was not viable.[285,451]
Ac54/Capsid protein	Homologs found in all baculovirus. Capsid protein required for nucleocapsid assembly. When deleted abnormal capsid structures were found.[30,452]
Ac66	Homologs present in all baculoviruses. Oriented in the opposite direction as DNA polymerase (Ac65) and 5' end overlaps with Ac65. ODV and BV associated, with ubiquitination present in BV. Involved in egress of virions and enucleation of polyhedra. Deletion mutant showed severely compromised BV titers and infection did not spread from cell.[453-455]
Ac73	Shows 99% probability to BAG (Bcl-2-associated anthanogene) protein, which regulate molecular chaperones. A BmNPV (Bm59) deletion mutant virus appeared normal.[456]
Ac74	BV associated. In BmNPV, deletion of ortholog Bm60 leads to reduced and delayed DNA synthesis, 10-fold lower BV production and longer time to kill larvae.[457,458]
Ac75	In BmNPV it is associated with ODV and BV and present in the ring zone of infected cells. May be involved in the nuclear egress of nucleocapsids and formation of intranuclear microvesicles. A BmNPV knock-out mutant produced no BV and virions were retained in nuclei.[443,459,460]
Ac76	Similarly to Ac75, Ac76 is present in the ring zone of infected cells and when deleted no BV production was detected and intranuclear microvesicles failed to form. Functions as a Type II integral membrane protein and may be involved in nuclear entry and egress of BV.[461,462]
Ac77/Vlf-1	Very late factor-1 with homologs found in all baculoviruses. Member of the lambda integrase family of protein that are site-specific DNA recombinases involved in catalysing DNA rearrangements, integration and excision of viral genomes and decatenation of newly replicated chromosomes. Binds to regulatory region of very late genes and influences hyperexpression. A mutant virus was able to synthesise DNA but produced tube-like capsids that lacked DNA.[111,463,464]
Ac78	90% similarity in structure to an integrin transmembrane domain. May be involved in nuclear entry and egress of BV and egress of nucleocapsids from nuclei. A knock-out virus showed unaffected DNA synthesis, but nucleocapsids were trapped to nuclei, BV not produced and polyhedra lacked occluded virions.[465,466]
Ac79	99% probability in similarity to an endonuclease domain of UvrC involved in DNA repair. Ac79 deletion reduced BV production, resulted in smaller plaques and showed disformed tube-like capsids. One BmNPV deletion study showed it was essential, whereas another showed it was able to produce BV and spread infection but reduced efficiency.[467,468]

Remaining Genes (CONT.)	
Ac80/GP41	A tegument protein modified with an O-linked-N-acetylglucosamine found between virion envelope and capsid protein. Required for nuclear egress of nucleocapsids. A BmNPV mutant was able to produce BV and spread infection but with reduced efficiency.[33,34]
Ac81	Homologs present in all baculoviruses. A BmNPV deletion mutant (Bm67) was severely compromised and did not produce BV.[285,469]
Ac82/TLP	Telokin-like protein. No sequence homology to telokin or structural similarity but reacted to telokin polyclonal antiserum, which may be an artefact of the antibody. Shows nuclear localization and Ac82 mutant virus resulted in reduced BV production. In BmNPV, BV production was also delayed as well as DNA replication.[296,297,470]
Ac90/lef-4	Component of the late baculovirus RNA polymerase. Found to be an mRNA capping enzyme and essential for late gene transcription.[13,471,472]
Ac91	A BmNPV homolog (Bm74) deletion resulted in no changes in virus replication but longer survival time of larvae was observed.[473,474]
Ac92/p33	A flavin adenine dinucleotide- linked sulfhydryl oxidase implicated in protection of cells from oxidative stress during apoptosis. Present in all baculoviruses and associated with ODV and BV. Essential as a deletion mutant could not be isolated.[475-477]
Ac93	Present in all baculoviruses. May be involved in nuclear entry and egress of nucleocapsids and formation of intranuclear microvesicles. An AcMNPV and BmNPV deletion mutant did not produce infectious BV.[35,285]
Ac94/ODV-e25	Encoded by all baculoviruses and found associated with ODV and BV. The N-terminal sequence seems to be a nuclear targeting sequence. A deletion mutant resulted in a 100-fold reduction in BV and polyhedra that lacked virions.[37,478,479]
Ac95/DNA helicase	Homologs present in all baculoviruses. Required for transient DNA replication and dependant on lef-3 for transport into nucleus. An essential gene as a deletion mutant could not be isolated.[283,480,481]
Ac96/pif-4	Per os infectivity factor with homologs present in all baculoviruses. Ac96 deletion mutant could replicate in cell lines but not insect larvae.[41,482,483]
Ac97	Small (56aa) protein present in only AcMNPV. The lack of homologs in other baculoviruses may suggest it is not a functional orf.[358]
Ac98/38K	Orthologs present in all baculoviruses and interacts with itself, encoded proteins from Ac54, Ac89 and Ac104. Associated with BV and ODV nucleocapsids. A deletion mutant showed normal DNA replication but nucleocapsid assembly was interrupted thus unable to produce infectious virions.[435,484]
Ac99/lef-5	Homologs present in all baculoviruses. Functions as an initiation factor for baculovirus polymerase. A knock-out virus was able to express early genes and replicate DNA, but late gene transcription was interrupted, and no infectious virions detected.[485-487]
Ac100/p6.9	Small (55aa) arginine/serine/threonine-rich DNA binding protein. Homologs present in all baculoviruses. Similar to protamines that are involved in the production of highly condensed DNA. ODV associated when unphosphorylated and hyperphosphorylated after synthesis, which is required for high levels of expression of late genes. A deletion mutant fails to produce nucleocapsids and infectious virions.[16,488,489]
Ac101/BV/ODV-C42	Orthologs present in all baculovirus. Associated with the capsid of both ODV and BV. Binds the actin nucleation factor PP78/83 and transports it into the nuclei. Ac101 knock-out bacmid resulted in virus with normal DNA synthesis but failed to form nucleocapsids.[24,29,490]
Ac102/p12	ODV-associated protein involved in the nuclear localisation of G-actin. Essential gene as a deletion mutant could not be isolated in AcMNPV and could not spread infection from one cell in BmNPV.[285,491,492]

Remaining Genes (CONT.)	
Ac103/p45	Homologs present in all baculoviruses genomes. Deletion was lethal as no infectious BV were detected and envelopment of ODV and incorporation into OB was interrupted.[493]
Ac104/vp80	Capsid associated and interacts with 38K. Localised to actin scaffolds in nuclei that connect the virogenic stroma to the nuclear envelope. Deletion is lethal as nucleocapsids were unable to move from the virogenic stroma.[23,494]
Ac105/RNA editing ligase	Early transcribed gene and may be involved in nuclear localisation of G-actin. Deletion mutants in AcMNPV and BmNPV suggests it is a non-essential gene.[285,491,495]
Ac106/107	These two orfs are likely a single gene as homologs are fused in many baculoviruses. Deletion of a BmNPV homolog (Bm90) resulted in a mutant that was unable to spread between cells indicating that it is an essential gene for virus replication.[285,383]
Ac108/pif-9	Per os infectivity factor. Interacts with pif-8 and forms part of the PIF complex. A deletion mutant in AcMNPV was not orally infectious, although infectious in BmNPV, the survival time of larvae was extended.[38,44,496]
Ac109	Homologs present in all baculoviruses. Found to be ODV and BV associated in AcMNPV. A deletion mutant experienced unaffected DNA replication but defects in nucleocapsid and polyhedra formation were detected.[497-499]
Ac110/pif-7	Per of infectivity factor found in all baculoviruses. A deletion mutant had no effect on BV and ODV formation, however inoculation of the OB failed to infect larvae.[46,311,500]
Ac111	Ac111 deletion had no effect on BV production or per os infectivity in *S. exigua* larvae, but infectivity was reduced by 5-fold in *T.ni* larvae, indicating it might be a per os infectivity factor depending on the host range. Deletion of BmNPV homolog had no effect on virus replication.[359]
Ac112/113	The two orfs are likely joined as found by sequencing data. A related gene is not found in BmNPV.[383,501]
Ac140	Encodes a 60aa protein and is only found in AcMNPV. No other information is found.[358]
Ac141/exon0	BV and ODV nucleocapsid associated and interacts with BV/ODV-C42 (Ac101) and fp25 (Ac61). May be involved in egress of nucleocapsids due to interaction with microtubules. A deletion mutant results in no spread of infection from the initial cell.[502,503]
Ac142/p49	Homologs present in all baculoviruses and found associated with ODV and BV virions. A knock-out mutant displays no difference in DNA synthesis to wt, however nucleocapsid formation was affected. A BmNPV homolog (Bm118) deletion mutant failed to produce BV and lacked virions in the polyhedra.[24,504,505]
Ac143/ODV-e18	Orthologs present in all baculoviruses and found associated with ODV and BV. Upon Ac143 deletion, no BV production is detected.[25,506]
Ac144/ ODV-ec27	Homologs found in all baculoviruses and found associated with BV and ODV for at least 3 different viruses. Deletion of Ac144 results in no evident nucleocapsids, however DNA synthesis is unaffected.[24,25]
Ac145	Related to Ac150. Both are predicted to encode a protein with a chitin binding domain. Ac145 deletion mutant experienced 6-fold drop in infectivity in larvae but when BV was injected intrahemocoelically, the mutant was as infectious as wt suggesting the genes play a role in oral infection.[385,507]
Ac146	Associated with BV and ODV, but not the envelopes of ODV. Ac146 deletion mutant did not produce infectious BV.[508,509]

8.5 Expected PCR sizes for validation of blasticidin integration

Table 8.11 PCR validation fragment sizes of blasticidin integration.

The gene orf, forward primer used and expected PCR product of the colony PCR checks for blasticidin integration. See Figure 4.2 for the PCR products. Ac124 - No correct PCR product. Problems with design of primer as another binding site present 800 bases further. Still proceeded to test but noted to check again if the virus variant was functional.

Gene orf	Forward primer	Product size (bp)
Ac148	1CheckFor	1316
Ac149	2/3CheckFor	678
Ac150	2/3CheckFor	967
Ac1	4/5CheckFor	726
Ac2	4/5CheckFor	1264
Ac3	6/61CheckFor	701
Ac15	7CheckFor	851
Ac18	8CheckFor	774
Ac19	9CheckFor	745
Ac20	10/11CheckFor	786
Ac21	10/11CheckFor	1078
Ac22	12CheckFor	831
Ac23	13CheckFor	773
Ac32	14CheckFor	1231
Ac33	15CheckFor	768
Ac39	16CheckFor	818
Ac42	17CheckFor	809
Ac43	18/19CheckFor	756
Ac44	18/19CheckFor	970
Ac45	20CheckFor	790
Ac46	21CheckFor	861
Ac47	22CheckFor	803
Ac49	23CheckFor	921
Ac56	24/25CheckFor	823
Ac57	24/25CheckFor	1262
Ac58	26/27CheckFor	748
Ac59	26/27CheckFor	1055
Ac60	28/29CheckFor	889
Ac61	28/29CheckFor	1299
Ac63	30/31CheckFor	766
Ac64	30/31CheckFor	1239
Ac69	32CheckFor	795
Ac70	33CheckFor	785
Ac71	34CheckFor	764
Ac84	35CheckFor	1228
Ac85	36CheckFor	770
Ac86	37CheckFor	898
Ac87	38/39CheckFor	891
Ac88	38/39CheckFor	1272
Ac4	6/61CheckFor	912

Ac5	62CheckFor	764
Ac7	40CheckFor	703
Ac17	42CheckFor	747
Ac27	43CheckFor	846
Ac28	44CheckFor	775
Ac29	45/46CheckFor	808
Ac30	45/46CheckFor	1077
Ac31	41CheckFor	779
Ac35	47/48CheckFor	776
Ac36	47/48CheckFor	1056
Ac38	49CheckFor	833
Ac41	50CheckFor	829
Ac72	51CheckFor	779
Ac82	52CheckFor	771
Ac91	53CheckFor	908
Ac105	63CheckFor	1142
Ac106/107	64CheckFor	907
Ac110	65/66CheckFor	814
Ac111	65/66CheckFor	1033
Ac112/113	67CheckFor	904
Ac120	54/55CheckFor	1090
Ac124	54/55CheckFor	1622
Ac125	56CheckFor	868
Ac130	57CheckFor	871
Ac151	58CheckFor	858
Ac153	59CheckFor	797
Ac154	60CheckFor	806