2016

Sturgeon (Acipenseridae) phylogeny, biogeography, & ontogeny.

Katie May Laumann
College of William and Mary - Virginia Institute of Marine Science

Recommended Citation

Laumann, Katie May, "Sturgeon (Acipenseridae) phylogeny, biogeography, & ontogeny." (2016). *Dissertations, Theses, and Masters Projects.* Paper 1539616731.
https://dx.doi.org/doi:10.25773/v5-23d7-b702

Sturgeon (Acipenseridae)
Phylogeny, Biogeography, & Ontogeny

A Dissertation
Presented to

The Faculty of the School of Marine Science
The College of William & Mary in Virginia

In Partial Fulfillment
Of the requirements for the Degree of
Doctor of Philosophy

By

Katie May Laumann

2016

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Katie May Laumann

Approved by the committee, April 2016

Eric Hilton, Ph.D.
Committee Chairman / Advisor

Casey Dillman, Ph.D.
Smithsonian Institution
Washington, D.C.

J. Emmett Duffy

J. Emmett Duffy, Ph.D.
Smithsonian Institution
Washington, D.C.

John Graves, Ph.D.

Lynne Parenti, Ph.D.
Smithsonian Institution
Washington, D.C.

DEDICATION

For my family, my wonderful parents Bonnie May Kersta and Rupert Laumann and my sister Jessie May Laumann, and most of all for my husband Jon Lefcheck, and Berlioz May Fett Laumann-Lefcheck, who have provided unequaled love and support.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES 4

LIST OF FIGURES 5

ABSTRACT 7

Sturgeon (Acipenseridae): 9

Phylogeny, Biogeography, & Ontogeny 9

INTRODUCTION 10

- References 13

Figures 14

CHAPTER 1. Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data 16

- Abstract 17
- 1. Introduction 19
- 2. Methods 29
  - 2.1 Taxon Sampling and Summary 29
  - 2.2 DNA Purification and Sequencing of *A. fulvescens* and *H. dauricus* 30
  - 2.3 DNA Purification and Sequencing for Other Species 31
  - 2.4 Alignments and Model Tests 32
  - 2.5 Phylogenetic Analyses 33
  - 2.6 Tree Comparisons 34
  - 2.7 Comparisons to Previous Studies 35
- 3. Results 36
  - 3.1 Sequences and Alignment 36
  - 3.2 Full mitogenome results 37
  - 3.2.1 Phylogenetic Hypothesis Support Values 37
  - 3.2.2 Full Mitogenome Hypotheses 37
  - 3.3 Gene Tree Analysis and Comparison to the Full Mitogenome 39
  - 3.3.1 tRNA Gene Family Phylogeny 39

3.3.2 rRNA Gene Family Phylogeny 40
3.3.3 Protein-coding Gene Family Phylogeny 41
3.4 Tree Comparisons: Influence of Individual Gene Regions on Gene Tree topology 41
3.5 Gene Influence on Bayesian Full Mitogenome Tree 41
3.5.1 Gene Families 41
3.5.2 Individual Genes 42
3.6 Comparison to Birstein et al. (2002) 44
3.6.1 Parsimony 44
3.6.2 Maximum Likelihood 46
3.6.3 Combined Evidence 48
4. Discussion 49
4.1 Phylogenetic structure within Acipenseridae 49
4.2 Influence of Individual Genes on Phylogeny 53
4.3 Taxon coverage 54
4.4 Mitogenome Coverage 56
4.5 Use of Mitochondrial Sequences: Benefits and Drawbacks 56
5. Concluding Remarks 57
Acknowledgements 59
References 60
Tables 75
Figures 81
CHAPTER 2. The Biogeography of Extant Sturgeons 90
Abstract 91
1.0 Introduction 92
1.1 Sturgeons and Biogeography 92
1.2 History of Areas Inhabited by Sturgeons 97
1.2.1 Geologic History: Formation of the Atlantic and Pacific Oceans and Ponto-Caspian Region 97
1.2.2 Geologic History: The Arctic Ocean and Bering Sea 99
1.2.3 Recent Geologic History: Baltic and North Seas 99

1.2.4 Recent Geologic History: East China, Yellow, and Japan Seas 101
1.2.5 Recent Geologic History: Interior North America, The Mississippi and Missouri Rivers and their tributaries 101
2.0 Methods 102
2.1 Phylogenies for Comparison 103
2.2 Species Ranges and Phylogenies 104
3.0 Results 105
3.1 Areas of Endemism 105
3.2 Intersection Tree 108
4.0 Discussion and Conclusions 108
4.1 Anadromy 108
4.2 Area Relationships 109
4.2.1 Atlantic Clade 109
4.2.1a Atlantic Clade: North American and Baltic/North Sea/Ponto-Caspian Regions 109
4.2.1b Atlantic Clade: North American Regions 110
4.2.1c Atlantic Clade: Baltic/North Sea and Ponto-Caspian Regions 110
4.3 Pacific Clade 111
4.3.1 Pacific Clade: Asia and Pacific North America 111
4.3.2 Pacific Clade: Pacific and Arctic 112
4.4 Bering Sea 114
4.5 Conclusions 115
Acknowledgements 116
References 117
Tables 130
Figures 133
CHAPTER 3. Utility of Ontogenetic Characters in Phylogenetic Studies 154
Abstract 155
1. Introduction 156
2. Methods 160
2.1 Taxon Sampling and Data Collection 160

2.1.1 Morphological Characters 160
2.1.2 Behavioral Characters 161
2.2 Metrics for Phylogenetic Signal in Individual Traits 164
2.3 Tree Comparison Metrics 165
2.4 Phylogenetic Comparative Tests 166
3. Results 167
3.1 Traits 167
3.1.1 Morphological Ontogenetic Traits: Age 167
3.1.2 Morphological Ontogenetic Traits: Length 168
3.1.3 Behavioral Ontogenetic Traits: Age 168
3.2 Pruned (reference) topologies 169
3.3 Phylogenetic Signal in Individual Traits 169
3.4 Tree Comparisons 170
3.4.1 Morphological Character Tree Comparisons 170
3.4.2 Behavioral Character Tree Comparisons 170
3.5 Correlations between Characters 171
4. Discussion and Conclusion 171
4.1 Phylogenetic Signal 171
4.3 Concluding Remarks 175
Acknowledgements 176
References 177
Tables 185
Figures 189
CONCLUSION 199
References 202
Appendix 1 203

ACKNOWLEDGEMENTS

Many thanks to my committee, Drs. Eric Hilton, John Graves, Emmett Duffy, Casey Dillman, and Lynne Parenti, for their assistance, advice, and support throughout this work. Thanks also to Jan McDowell for support, mentorship, and friendship, and to Dr. Masaki Miya for mentorship and the amazing opportunity to work with his lab at the Natural History Museum and Institute in Chiba, Japan. Thanks to my family my good friends Derek and Rebekah Loftis and Ashley Morton for encouragement and support. Finally, thank you Jon and Berlioz.

LIST OF TABLES

CHAPTER 1. Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data 16

Table 1. GenBank Accession Numbers 75

Table 2. Gene Families, Lengths, and Best Models 76

Table 3. Rzhetsky-Nei Distances: Bayesian Mitogenome Topology and Gene-Family Topologies 77

Table 4. Full Mitogenome Hypothesis Clade Constituents, and Presence in Gene-Family Trees. 77

Table 5. Rzhetsky-Nei Distances between Bayesian Mitogenome Topology and Gene Trees 78

Table 6. Birstein et al. 2002 Parsimony Clade Constituents. 79

Table 7. Birstein et al. 2002 Maximum Likelihood Clade Constituents. 80

Table 8. Birstein et al. 2002 Combined Evidence Clade Constituents. 80

CHAPTER 2. The Biogeography of Extant Sturgeons 90

Table 1: Sturgeon species ranges. 130

Table 2: Lamprey species ranges. 131

Table 3: Pinniped species ranges. 132

CHAPTER 3. Utility of Ontogenetic Characters in Phylogenetic Studies 154

Table 1: Character matrix of three behavioral traits. 185

Table 2: Earliest occurrences of ontogenetic changes in four species. 186

Table 3: Calculations of phylogenetic signal in morphological characters based on age at which milestones first occur. *s indicate significance 187

Table 4. Calculations of phylogenetic signal in morphological characters based on total length (mm) at which milestones first occur. 188

LIST OF FIGURES

INTRODUCTION 10

Figures 14

Figure 1. *Acipenser gueldenstaedtii* (Russian sturgeon) with protruding jaw 14

Figure 2. *Acipenser gueldenstaedtii* (Russian sturgeon) with distinctive scutes and barbels .. 15

CHAPTER 1. Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data 16

Figure 1: Parsimony tree based on 12s, 16s, cyt*b*, D-loop, and NADH 5 81

Figure 2: Maximum Likelihood tree based on 12s, 16s, cyt*b*, D-loop, and NADH 5 82

Figure 3: Combined Evidence tree based on 12s, 16s, cyt*b*, D-loop, NADH 5, and morphological data 83

Figure 4: Bayesian full mitogenome phylogeny. 84

Figure 5: Maximum Likelihood Full mitogenome phylogeny. 85

Figure 6: Parsimony full mitogenome phylogeny 86

Figure 7: Bayesian phylogenetic hypothesis based on concatenated tRNA sequences. 87

Figure 8: Bayesian phylogenetic hypothesis based on concatenated rRNA sequences. 88

Figure 9: Bayesian phylogenetic hypothesis based on concatenated protein-coding sequences. 89

CHAPTER 2. The Biogeography of Extant Sturgeons 90

Table 1: Sturgeon species ranges 130

Table 2: Lamprey species ranges 131

Table 3: Pinniped species ranges 132

Figure 1: Bayesian full mitogenome phylogeny. 133

Figure 2: Croizat's global biogeography map 134

Figure 3: Areagram of Croizat's (1958) biogeographic regions 135

Figure 4: Map of sturgeon Areas of Endemism; 136

Figure 5: Morphological Lamprey Phylogeny 137

Figure 6: Molecular Pinniped Phylogeny 138

Figure 7: Range maps of sturgeon species included in biogeographic analyses 139
Figure 8: Sturgeon Areagram 150
Figure 9: Lamprey Areagram 151
Figure 10: Pinniped Areagram 152
Figure 11: Sturgeon-Lamprey-Pinniped Intersection Tree 153
CHAPTER 3. Utility of Ontogenetic Characters in Phylogenetic Studies 154
Figure 1: Bayesian full mitogenome phylogeny. 189
Figure 2: Example of cleared and stained specimens 190
Figure 3: Full mitogenome Bayesian phylogeny, pruned 191
Figure 4: Full mitogenome Bayesian phylogeny 192
Figure 5: Topology shown by single-character (yolk sac and teeth) trees 193
Figure 6: Topology shown by single-character (jaws elements) tree 194
Figure 7: Topology shown by single-character (migration) tree 195
Figure 8: Topology shown by single-character (nocturnality) trees 196
Figure 9: Topology shown by single-character (phototaxis) trees 197
Figure 10: Phylogenetic Independent Contrast values 198
CONCLUSIONS 199

ABSTRACT

Thought of as "ancient" fishes, 25 broadly recognized extant sturgeon species are classified in four genera (*Acipenser, Huso, Pseudoscaphirhynchus,* and *Scaphirhynchus*). Molecular and morphological analyses have led to broad but conflicting changes to sturgeon phylogeny. For example, the position of *Scaphirhynchus* among other sturgeons had been contentious, and various sets of sturgeon species have been proposed to make up the subfamily Husinae. Here, a molecular phylogeny of sturgeons, based on the full mitogenome, is presented. In this phylogeny, *Scaphirhynchus* is recovered with strong support as basal to the other sturgeons. *Huso huso* is recovered as basal within a clade containing *P. kaufmanni* and several species of *Acipenser*, and is proposed as a new, monotypic subfamily Husinae. This phylogeny is used to examine phylogenetic signal in individual genes and in gene families. The protein coding genes as a unit, and individually, along with 16s rRNA, show phylogenetic signal most similar to that of the full mitogenome. The phylogeny, along with evolutionary relationships of pinnipeds and lampreys, provides the basis for the exploration of sturgeon biogeography. Relationships among geographic areas inhabited by sturgeons are found, finding two sets of related areas- a Pacific area group and an Atlantic group. Relationships of areas within and between these groups reflect area relationships proposed by previous biogeographic and geologic studies. Phylogenetic signal is tested amongst ontogenetic characters, and is recovered in the timing at which larval sturgeon teeth are completely resorbed, indicating that the timing of ontogenetic milestones can carry signal. The phylogeny is used to remove confounding signal from, and

investigate correlations among, behavioral and morphological ontogenetic characters.

Correlation is found between one pair of characters.

Sturgeon (Acipenseridae):

Phylogeny, Biogeography, & Ontogeny

INTRODUCTION

The 25 broadly recognized extant sturgeon species are classified in four genera (*Acipenser, Huso, Pseudoscaphirhynchus,* and *Scaphirhynchus*) and are widely acknowledged to form the monophyletic family Acipenseridae. Thought of as "ancient" fishes, sturgeons occupy a basal position relative to other ray-finned fishes (Actinopterygii). Together with their sister group, the paddlefishes (Polyodontidae), they form the extant sister-group to Neopterygii (the group including gars, bowfins, and teleostean fishes; Grande and Bemis 1996). Sturgeons range throughout the freshwaters and coastal regions of Europe, Asia, and North America. Most sturgeons are diadromous, moving between fresh and saltwater for feeding (primarily downstream movement) and reproductive (upstream movement) purposes (Bemis and Kynard 1997, Peng 2007). Sturgeons have protrusible jaws that enable suction feeding (Figure 1), chemosensory barbels (Figure 2), and bony plates called scutes along their bodies in rows (Figure 2). These charismatic fishes are culturally and economically important (e.g., as the source of true caviar). The oldest known sturgeon fossils are from the Late Cretaceous (85 million years; Hilton and Grande 2006), but the sister-group relationship between Acipenseridae and Polyodontidae, combined with the existence of paddlefish fossils from the Early Cretaceous, indicates that sturgeons are much older (Peng *et al.* 2007). Populations of many sturgeon species are currently in sharp decline due in part to overfishing and habitat alterations (Holzkamm and Waisberg 2004), and all species of sturgeons are provided conservation protection (IUCN 2014). Even with the imperiled status of these fishes, the study of their evolutionary history through examination of character states and cladistic methods began only relatively recently (Findeis 1997), and a well-resolved, broadly accepted phylogeny for the family remains elusive.

The first explicitly cladistic study of the family Acipenseridae, based on morphological data (Findeis 1997), recovered the genus *Huso* as the sister group of the other three genera and recovered *Scaphirhynchus* and *Pseudoscaphirhynchus* as sister taxa (forming the tribe Scaphirhynchini, recognized as the subfamily Scaphirhynchinae by some authors). Molecular data were introduced into the study of sturgeon phylogenetics at about the same time (e.g., Birstein and DeSalle 1998) and recent analyses using both molecular and morphological data (independent of one another and together) have resulted in markedly different phylogenetic hypotheses than those of Findeis (1997) (Hilton *et al.* 2011). Several broad changes to Findeis' phylogeny of Acipenseridae have been proposed across morphological and molecular studies, for example the dissolution of Scaphirhynchini. The details of these differences, however, vary from study to study; for example, the position of *Scaphirhynchus* is has been debated. Some differences in hypothesized relationships of sturgeons may be due to incomplete taxon sampling and the use of different molecular markers and morphological characters across studies.

Variation in the details of recent sturgeon phylogenetic hypotheses illustrates the need for further examination. The goal of this dissertation research was to use modern phylogenetic methods and various types of data to work toward a well-resolved sturgeon phylogeny. I constructed a molecular phylogeny, sequencing the full mitogenome and including near-complete taxon sampling, explored the effects of using different segments of mitochondrial DNA on phylogeny, and different optimality criteria for phylogenetic hypotheses. I used the resulting full mitogenome phylogenetic hypothesis to investigate area relationships of different areas inhabited by sturgeons. Finally, I investigated jaw development in five North American species in the genus *Acipenser,* and identified and tested potential morphological and behavioral ontogenetic characters for use in phylogenetic studies.

References

Bemis W.E. and Kynard B. 1997. Sturgeon rivers: an introduction to acipenseriform biogeography and life history. Environmental Biology of Fishes, 48:167-183.

Birstein V.J. and DeSalle R. 1998. Molecular phylogeny of Acipenserinae. Molecular Phylogenetics and Evolution, 9(1): 141-155.

Findeis, E.K. 1997. Osteology and phylogenetic interrelationships of sturgeons (Acipenseridae). Environmental Biology of Fishes, 48(1-4): 73-126

Grande, L. and W.E. Bemis. 1996. Interrelationships of Acipenseriformes with comments on "Chondrostei". pages 85-115. In: M.L.J. Stiassny, L.R. Parenti & G.D. Johnson (eds.) Interrelationships of Fishes. Academic Press, New York.

Hilton E.J., Grande L., Bemis W.E. 2011. Skeletal anatomy of the shortnose sturgeon *Acipenser brevirostrum* Lesueur, 1818, and the systematics of sturgeons (Acipenseriformes, Acipenseridae). Fieldiana Life and Earth Sciences, 3: 1-168.

Holzkamm, T.E. and L.G. Waisberg. 2004. Native American utilization of sturgeon. In G.T.O. LeBreton, F.W.H. Beamish and R.S. McKinley (eds.). Sturgeon and Paddlefish of North America. Kluwer Academic Publishers, Dordreht. Fish and Fisheries Series 27: 22–39.

(IUCN) The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>.

Nelson J. S. 2006. Fishes of the World, 4th ed. John Wiley and Sons, New York.

Peng Z., Ludwig A., Wang D., Diogo R. , Wei Q., He S. 2007. Age and biogeography of major clades in sturgeons and paddlefshes (Pisces: Acipenseriformes). Molecular Phylogenetics and Evolution, 42: 854-862.

Figures

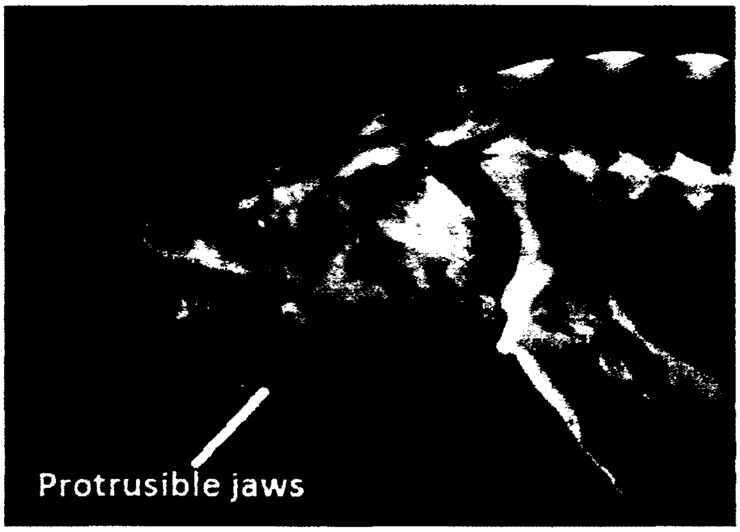


Figure 1. Acipenser gueldenstaedtii (Russian sturgeon) with protruding jaw, enabling suction feeding. Photo taken at Volga River Exhibit at the Tennesse Aquarium in Chatanooga.

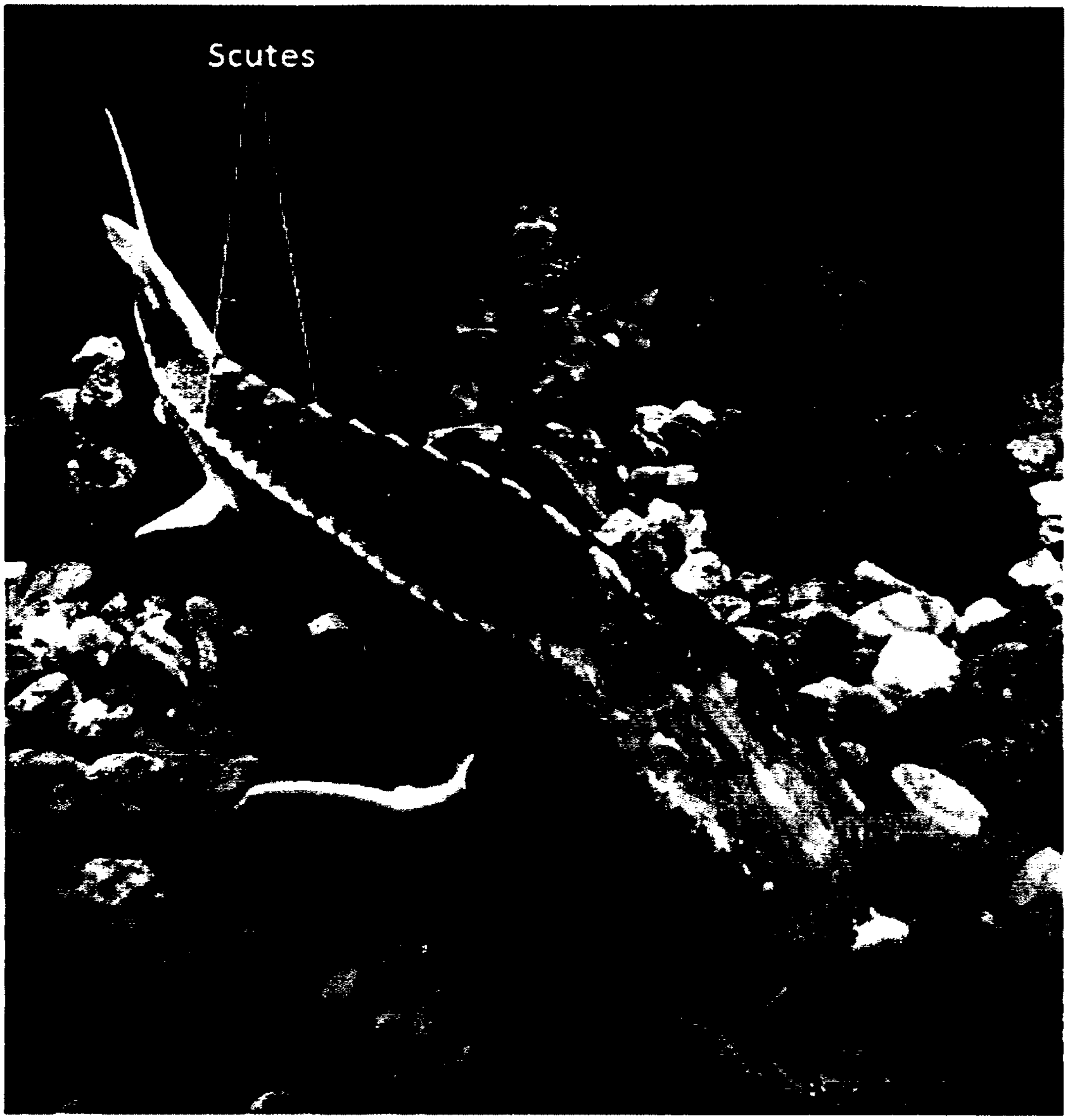


Figure 2. Acipenser gueldenstaedtii (Russian sturgeon) with distinctive scutes and barbels indicated. Photo taken at Volga River Exhibit at the Tennesse Aquarium in Chatanooga.

CHAPTER 1.

Phylogeny of Sturgeons (Actinopterygii: Acipenseriformes: Acipenseridae) Based on Full Mitogenomic Data

Abstract

The 25 commonly recognized species of sturgeons, family Acipenseridae (Actinopterygii, Acipenseriformes), are traditionally classified in four genera (*Acipenser, Huso, Scaphirhynchus,* and *Pseudoscaphirhynchus*). Recent phylogenetic studies of the family, using both morphological and molecular data, generally recover *Acipenser* as paraphyletic, *Huso* nested within *Acipenser,* and *Pseudoscaphirhynchus* as sister to *A. stellatus*. The details of phylogenetic relationships among species within the family, however, vary from study to study. Previous molecular-based studies of the family have been limited to the use of fragments of the mitochondrial gene to estimate the phylogenetic relationships among sturgeons. In this study, I present a phylogenetic hypothesis based on analysis of the complete mitogenomic sequences of 23 sturgeon taxa (22 species and both subspecies of *A. oxyrinchus*) and discuss the implications of the proposed relationships. This analysis supports the inclusion of *Huso* spp. in *"Acipenser"* and proposes that "Scaphirhynchini" is nonmonophyletic. *Huso huso* is recovered as sister to a clade that includes *A. ruthenus, A. fulvescens, A. baerii, A. naccarii, A. gueldenstaedtii, A. brevirostrum,* and *A. stellatus* and *P. kaufmanni* as sister species. *Huso dauricus* is recovered in a clade with *A. medirostris* and *A. mikadoi,* and this clade is sister to *A. dabryanus, A. sinensis, A. transmontanus,* and *A. schrenckii. Scaphirhynchus* is recovered as sister to all other sturgeons. *Acipenser sturio* and both subspecies of *A. oxyrinchus* form another monophyletic group. I explore the influence of each mitochondrial gene region on full mitogenome topologies. Although they are linked within the same molecule, individual rRNA, tRNA, and protein-coding mitochondrial genes do not all exhibit the same phylogenetic signal as one another or as the full mitogenome as a whole. Several genes, including all of the protein-coding and rRNA genes, as well as D-loop, recover significant phylogenetic structure among sturgeons. Of those genes that

recover structure, the protein coding genes and 16s recover relationships similar to those found by the full mitogenome, whereas 12s and D-loop recover structure different from that shown by the full mitogenome. Other genes, including the tRNAs, recover very little phylogenetic structure, showing either complete polytomies or several groups, within which there islittle or no resolution.

1. Introduction

The 25 commonly recognized extant species of sturgeon, traditionally classified in four genera (*Acipenser, Huso, Pseudoscaphirhynchus,* and *Scaphirhynchus*), form the monophyletic family Acipenseridae Bonaparte 1831 (Berg 1904, Bemis *et al.* 1997, Birstein *et al.* 1997, Findeis 1997, Hilton *et al.* 2011). Two monotypic genera, †*Priscosturion* (Grande and Hilton 2006) and †*Protoscaphirhynchus* (Wilimovsky 1956), are known only as fossils. The paddlefishes (Polyodontidae), two entirely fossil families (†Chondrosteidae, †Peipiaosteidae), and Acipenseridae together form the Acipenseriformes, which is broadly regarded as the extant sister-group to Neopterygii (gars, bowfins, and teleosts; Grande and Bemis 1996, Bemis *et al.* 1997, Nelson 2006, Hilton and Forey 2009, Hilton *et al.* 2011). The earliest fossil evidence of sturgeons dates to the Late Cretaceous, 85 million years ago ago (Hilton and Grande 2006), but these fishes are likely older, as polyodontid fossils date back to the Early Cretaceous (Bemis *et al.* 1997, Grande *et al.* 2002). Due in part to overfishing and habitat alterations (Holzkamm and Waisberg 2004), nineteen species have been listed as endangered or critically endangered by the IUCN (2014) and at least one (*P. fedtschenkoi*) is likely extinct (IUCN 2014). International trade of all sturgeon species is closely monitored (ICES 12.7 CoP16).

A broad understanding of sturgeon evolutionary history may have implications for management and conservation efforts. These efforts focus on regions, species, and the evolutionary process itself (Forest *et al.* 2007). Conservation managers recognize that species represent evolutionary history, and treat the loss of "old" taxa, taxa undergoing rapid diversification, evolutionarily distinct taxa, and taxa at high risk of extinction as priorities for conservation (Crozier 1997, Isaac *et al.* 2007, Forest *et al.* 2007, Rodrigues and Gaston 2002, Moritz 1994, Rodrigues and Gaston 2002). Determining whether a species or group of species fits into these categories requires phylogenetic information, and although incomplete phylogenies may be useful, hypotheses based on comprehensive datasets are the most informative for management and conservation planning (Moritz 1994, Rodrigues *et al.* 2011). As conservation and management strategies aim to conserve phylogenetic diversity (evolutionary differences across all biodiversity, Lefcheck *et al.* 2014), phylogenetic hypotheses are playing a larger role in conservation efforts (Faith 1992, Avise 2010). Understanding phylogeny also allows conservation managers and planners to estimate historic population size, identify trends in gene flow, and determine how bottlenecks have historically impacted currently imperiled species (Moritz 1994). These data can help conservation managers determine the best methods for conservation of various taxa (Moritz 1994), helping maintain gene flow in extant populations and preventing or mitigating the effects of bottlenecks. All of the criteria listed above apply to sturgeons- they are considered "ancient" fishes, and are declining across their ranges. A well-supported sturgeon phylogeny will help resource managers to determine the conservation priority of these fishes as well as helping identify avenues for the protection of these species (for example by helping to determine when hybridization with escaped or released hatchery-raised individuals might pose a risk to wild populations, and mitigate this threat). In addition to conservation and management value, a better understanding of sturgeon phylogeny may be broadly informative for the evolution of all fishes, as sturgeons generally represent the most species-rich clade of basal ray-finned fishes (Actinopterygii).

There have been numerous attempts to resolve the evolutionary history of Acipenseridae from many perspectives; however, a consensus on sturgeon phylogeny has not been reached. Based on morphological data, Findeis (1997) found *Huso* to be sister to all other sturgeons and recovered *Pseudoscaphirhynchus* and *Scaphirhynchus* as a monophyletic group, Scaphirhynchini (also sometimes classified as a subfamily, Scaphirhynchinae; e.g., see Mayden and Kuhajda 1996). Findeis (1997) included the genus *Acipenser* in his analysis as a terminal taxon but noted that it was likely nonmonophyletic as there were no synapomorphies supporting this genus. Subsequent morphological studies have led to various phylogenetic hypotheses. Mayden and Kuhajda (1996) recovered *Acipenser* and *Huso* as sister genera and found *Pseudoscaphirhynchus* and *Scaphirhynchus* to form a clade, though they failed to recover monophyly for *Pseudoscaphirhynchus*. Choudhury and Dick (1998) also recovered Scaphyrhynchini as monophyletic, and found distinct clades within *Acipenser*; *Huso* was not included in their study. Artyukhin (2006) recovered Scaphirhynchini as sister to the other sturgeons, and recovered distinct subclades within *Acipenser*. *Acipenser oxyrinchus* and *A. sturio* were hypothesized to be sister to the other sturgeon clades, which included *H. huso* and *H. dauricus* as sister taxa, grouped with *A. nudiventris, A. schrenckii,* and *A. ruthenus*. Other studies have recovered alternate clades within *Acipenser*. Hilton *et al.* (2011) also recovered *Scaphyrhinchus* as sister to the other sturgeons, finding (*H. huso, (A. ruthenus, A. baerii)*) to form a monophyletic group with *(A. stellatus, (P. kaufmanni, P. hermanni))*.

Molecular studies of sturgeon phylogenetics have led to additional phylogenetic hypotheses; these studies are based on mitochondrial DNA sequences. In fishes, the mitochondrial genome (also referred to as mitogenome or mitochondrion) is composed of 37 genes that are exclusively maternally inherited as a single linkage-group, and is considered to be a single locus. These genes are distributed amongst three gene families, the tRNAs, rRNAs, and protein coding genes, and evolve at different rates and with different nucleotide substitution patterns. Most animal mitochondrial genomes, those of sturgeons included, contain 22 tRNAs: tRNA-Phe, -Val, -Leu, -Ile, -Gln, -Met, -Trp, -Ala, -Asn, -Cys, -Tyr, -Ser, -Asp, -Lys, -Gly, -Arg, -His, -Ser (2), -Leu (2), -Glu, -Thr, -Pro. Two genes, 12s and 16s, make up the rRNA (ribosomal ribonucleic acid) gene "family". The protein coding gene family comprises 13 genes: ATP6 and ATP8, COI, COII, COIII, cyt*b, and* NADH's 1-6 and NADH 4L. Mitochondrial DNA, is not protected from oxidative damage by chromatin, and so is exposed to free radicals and susceptible to injury (Gredilla *et al.* 2010). Although mechanisms to mitigate this damage have been identified in the mitochondrion, recent studies suggest that these mechanisms cannot repair damage after it is caused (Gredilla *et al.* 2010). In addition to decreased damage repair as compared to nuclear DNA, proofreading mechanisms during mitochondrial DNA replication are limited. As this portion of the genome does not undergo recombination, mismatches cannot be repaired based on a reference or template sequence. The proofreading mechanisms that are present in mitochondrial DNA have been shown to be inefficient in some mammal cells (Song *et al.* 2005). In contrast to this, nuclear DNA is packaged and protected by chromatin and therefore not exposed to the same levels of damage. Recombination allows for corrections to mismatches in nuclear DNA replication, and additional proofreading mechanisms are efficient, leading to high fidelity in nuclear DNA replication (Caravas 2012). This leads to a high rate of substitution and evolution in the mitochondrial genome as compared to the nuclear genome (Avise 1998, Caravas 2012).

Mitochondrial DNA is invaluable to phylogenetic studies, but various concerns have been raised about relying exclusively on mitochondrial genes to elucidate phylogeny (e.g. Caraveas 2012, Moore 1995). Pitfalls concerning mitochondrial DNA could lead to the recovery of unresolved or incorrect topologies. Even if a well-supported mitochondrial phylogeny is recovered, in some situations this phylogeny may not represent the species phylogeny (Caravas 2012). The high rate of substitution in mitochondrial DNA could result in saturation or overwriting of phylogenetic signal, weakening the display of this signal (DeFilippis and Moore 2000). Composition bias of adenine and thymine present in mitochondrial DNA may lead to the same issue (DeFilippis and Moore 2000). The reliance on few or short mitochondrial sequences, frequent in past molecular phylogenetic studies of sturgeons, could lead to low resolution in the resulting phylogeny (Vilstrup *et al.* 2011). Including multiple genes from the mitogenome, however, only provides a single independent estimate of phylogenetic relationships because these genes are linked and do not evolve strictly independently of one another (Moore 1995).

Additional challenges to the use of mitochondrial DNA in phylogenetic studies include hybridization, or reproduction between different species. The use of mitochondrial DNA in phylogenetic studies of taxa that exhibit hybridization may result in the recovery of a gene tree that is different from the species tree (Baack and Rieseberg 2007, Caravas 2012, Moore 1995). Hybridization can lead to introgression, the incorporation of genetic material from one species into the genome of another (Baack and Rieseberg 2007). Hybridization and introgression are especially problematic when using mitochondrial DNA alone to infer phylogeny because if the mitochondrial DNA of one species is incorporated into a hybrid or an individual of another species, that mitochondrial DNA will not reflect the evolutionary history of the species. If undetected, this will mislead phylogenetic analyses (Moore 1995). If hybridization is detected, it can be accounted for in phylogenetic studies (Moore 1995), but the detection of hybridization and introgression requires comparison between topologies based on unlinked genes (Caravas 2012). Because all mitochondrial genes are linked, hybridization and introgression are undetectable based on mitochondrial sequence data alone. Mitochondrial phylogenies must be compared with nuclear phylogenies to search for congruence. If nuclear and mitochondrial phylogenies present the same topology, introgression is unlikely. If they present conflicting hypotheses, introgression should be explored as one possible cause of incongruence (Duchene 2011, Moore 1995, Caravas 2012, Baack and Rieseberg 2007).

Despite these potential pitfalls, there are several advantages to the use of mitochondrial sequences in phylogenetic analyses. Many copies of the mitochondrion are present in each cell, making amplification of mitochondrial sequences relatively simpler than amplification of nuclear genes (Caravas 2012). Mitochondrial gene order is generally conserved in vertebrates, and the mitochondrion lacks introns and can therefore be sequenced in larger units can than nuclear genes (Caravas 2012). Recent technological innovations allow for the sequencing of the full mitogenome at relatively low cost and in relatively short amounts of time (Mardis 2008). Mitochondrial genes and the mitochondrial genome as a whole have been used to recover the phylogenies of many taxonomic groups, including Diptera (Caravas 2012), reptiles and amphibians (Roos *et al.* 2007, Zhang *et al.* 2005, Zhang and Wake 2009, Zhang *et al.* 2013), delphinids (Duchene *et al.* 2011), teleostean and other fishes (Inoue *et al.* 2003, Ishiguro *et al* 2003, Miya *et al.* 2001, Inoue *et al.* 2010, Lavoue *et al.* 2007), and various groups of birds (DeFilippis and Moore 2000).

In addition to phylogenetic reconstruction, the utility of mitochondrial genes in species identification, for example, is indisputable (Wolf 1999, Brown 1996, Miya and Nishida 2000). Cytochrome *c* oxidase I (COI), for example, is used for DNA barcoding (Ward 2009). Among fishes, COI has been shown to distinguish among 98% of marine and 93% of freshwater species (Ward 2009). Additionally, a hypervariable region of rRNA was recently found to successfully identify, to species, at least 168 taxonomically diverse teleostean fishes, including closely related congeners (Miya and Nishida 2015). Although individual mitochondrial genes used for species identification may not carry enough information to resolve phylogenies, they may be useful in phylogenetic investigations when considered alongside other genes. Genes that differentiate between species or even subspecies add resolution to "shallow" branches in a phylogeny, enhancing the phylogeny by providing support for fine-scale relationships (Hajibabaei *et al.* 2007). Some barcoding genes, COI in particular, have been shown to be extremely useful in phylogenetic studies focusing on relationships at the species level (e.g. in caddisflies, Hogg *et al.*

2009, and in the families Apionidae, Ptaszynska *et al.* 2012 and Sarcophagidae, Meiklejohn *et al.* 2013).

Although mitochondrial genes are linked, they evolve at different rates (Xia 1998, Sloan *et al.* 2009). Recent studies generally conclude that the common practice of relying on one or only a few mitochondrial genes to inform phylogeny is not ideal, as some genes are misleading and inappropriate for phylogenetic reconstruction because they recover topologies with low resolution (Duchene *et al.* 2011, Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008). Selection of genes to be used in a phylogenetic study is important not only because phylogenetic signal may vary widely from gene to gene, but also because different genes resolve relationships at different evolutionary scales.

Various studies spotlight the different phylogenetic signals carried by different mitochondrial genes by comparing topologies derived using these genes with other topologies recovered by full mitogenomic, morphological, or other data. Zardoya and Meyer (1996) compared tetrapod and mammal topologies recovered by all protein-coding genes combined and individual protein-coding genes. They found that NADH4, NADH5, NADH2, cyt*b,* and COI recovered topologies consistent with those previously proposed based on morphological and genetic data, whereas topologies based on ATP6, ATP8, NADH3, and NADH4L diverged significantly from previous topologies. Miya and Nishida (2000) evaluated the similarity of single protein-coding gene topologies and a tRNA topology to the full mitogenome topology of eight teleosts and found that only NADH5 and the concatenated tRNA sequences recovered the same topology as the full mitogenome. Rastorguev *et al.* (2008) similarly compared individual protein coding gene topologies to the full mitogenome topology of 10 sturgeons and found that COI, NADH4, and ATP6 recovered topologies similar to that of the full mitogenome. Analyses of NADH6, in contrast, recovered a very different topology from that of the full mitogenome.

Using the oceanic dolphins (family Delphinidae) and the killer whales (genus *Orcinus*), Duchene *et al.* (2011) compared topologies inferred from the full mitogenome to those derived from 12s, 16s, and individual protein coding genes. They found different results for the two different taxonomic groups, with COI, *cytb*, NADH3 and ATP6 recovering topologies (with varying levels of support) similar to that of the full mitogenome for *Orcinus*, and NADH1, COI and NADH4 doing so with the delphinids. Cleary, individual mitochondrial genes carry different phylogenetic signal from one another, with some carrying similar signal to that of the full mitogenome.

Efforts to identify factors contributing to the different phylogenetic hypotheses recovered by different genes have found that neither gene length nor evolutionary rate accurately predict how similar a gene's topology will be to that derived from the full mitogenome (Zardoya and Meyer 1996). The depth of the phylogeny under consideration may play a role, as might the completeness of the taxon sampling (Zardoya and Meyer 1996). Therefore, when selecting genes to be used in a phylogenetic study, the taxonomic group, number of taxa for which data are available, and target resolution of the phylogeny should be considered.

Longer DNA sequences than those that have previously been used may help resolve phylogenetic relationships, particularly higher-level relationships (Miya and Nishida 2015). The use of more data, particularly those from molecular gene regions that evolve at different rates, can provide resolution at multiple phylogenetic levels and reduce biases that may be gene-specific (Hajibabaei *et al.* 2007). Previous studies were limited because there was no straightforward method to sequence the entire mitogenome until fish-versatile primers were designed to sequence the full mitogenome via a combination of long-and short PCR (Miya and Nishida 1999, 2015). The development of this methodology has allowed exploration of individual mitogenomes, mitochondrial gene arrangement, and reconstruction of phylogeny.

The utility of complete mitogenomic data in phylogenetic analysis of fishes has been demonstrated in studies of both higher-level relationships (e.g. teleostean phylogenetic relationships, Miya and Nishida 2000b) and in population-level relationships (e.g. Jacobsen *et al.* 2012). More recently, next-generation sequencing (NGS) methods have been applied to full-mitogenome sequencing techniques (Ahn *et al.* 2009, Groenenberg *et al.* 2012, Feldemeyer *et al.* 2010, Jex *et al.* 2010, Iorizzo *et al.* 2012, Varshney *et al.* 2009, Miya and Nishida 2015, Lindqvist *et al.* 2010, Webb and Rosenthal 2011, Botero-Castro *et al.* 2013, Shen *et al.* 2015).

Mitochondrial data have been used frequently in analyses of sturgeon phylogenies. Of ten recent molecular phylogenetic studies of sturgeons (Birstein and DeSalle 1998, Birstein *et al.* 2002, Dillman *et al.* 2007, Fontana *et al.* 2001, Krieger *et al.* 2000, Krieger *et al.* 2008, Ludwig *et al.* 2001, Rastorguev *et al.* 2008, Simons *et al.* 2001, Zhang *et al.* 2000), most have relied on short segments of the mitochondrial genome (<5000 base pairs, or <30% of the mitogenome), with just one (Rastorguev *et al.* 2008) considering the full mitogenome (but with limited taxon sampling). None of the studies included more than 22 of the 25 extant species, and only three (Birstein *et al.* 2002, Dillman *et al.* 2007, Krieger *et al.* 2008) considered specimens from all four recognized genera. Within these studies, the most commonly used mitochondrial genes were cyt*b*, NADH4, NADH4L, NADH5, the rRNAs, and D-loop. Single-gene trees resulting from analyses of these sequences suggest that each of these genes carries different phylogenetic signals, despite the fact that they all occur in the mitochondrial genome. Further, the strength of phylogenetic signal among these genes varies and when two genes are considered together, the signal carried by one may be stronger than that from another, "swamping" it out. This effect may be reversed or negated when more than two genes are considered together.

Along with recent morphological studies, these molecular analyses have expanded our understanding of sturgeon phylogeny (Hilton *et al.* 2011). Among the most significant findings of these studies include *H. huso* nested within *"Acipenser"* (e.g., Birstein and DeSalle 1998, Birstein *et al.* 2002, Dillman *et al.* 2007, Krieger *et al.* 2008, Hilton *et al.* 2011), which Vasil'eva *et al* (2009) formally recognized as *A. huso*, and break up Scaphirhynchini by recovering *A. stellatus* as sister to *Pseudoscaphirhynchus* spp. (e.g., Birstein *et al.* 2002, Dillman *et al.* 2007, Krieger *et al.* 2008, Hilton *et al.* 2011). However, the hypotheses of relationships among species vary from study to study. For example, various relationships of *Huso* within *"Acipenser"* have been proposed, including the recovery of *H. huso* as sister to *A. ruthenus* (in a redefined Husinae; Hilton *et al.* 2011) or as sister to other species of *"Acipenser"* (e.g., Fontana *et al.* 2001). Further, different relationships have been proposed regarding the position of genera formerly in the tribe Scaphirhynchini as well as the monophyly of species groups included in the genus *"Acipenser."* Molecular-based studies of sturgeon relationships have all used different mitochondrial gene regions or different combinations of gene regions, and therefore each study includes different characters leading to differences in the phylogenies proposed. Nuclear data have not been used in phylogenetic analyses of sturgeons, in part due to difficulties sequencing nuclear genes in these polyploid fishes (Krieger *et al.* 2006).

The goal of this study was to apply full mitogenome sequencing methods to complement previous studies of sturgeon phylogeny. Combining both long and short PCR methods and NGS methods, I build upon past studies by using more mitochondrial DNA sequence data than has been available previously. This allowed for the comparison of phylogenetic signal carried by individual mitochondrial genes to the signal of the full mitogenome, and led to the resolution of polytomies recovered in some recent sturgeon phylogenetic studies (Birstein *et al.* 2002, Dillman *et al.* 2007, Ludwig *et al.* 2001, and Zhang *et al.* 2000). I conducted a phylogenetic analysis of Acipenseridae using the full mitogenome of 23 sturgeon species and subspecies representing all four currently recognized sturgeon genera. I investigated phylogenetic signal present in individual genes, and determined which mitogenome segments exert the strongest signal recovering this phylogenetic hypothesis. To determine the relative phylogenetic signal of these three topologies, data were explored by assessing the topologies based on 1) all protein coding gene regions analyzed together, and sequences of 2) all ribosomal RNA and 3) all transfer RNA sequences. Additionally, I compared my topology to three topologies from a previous study (Birstein *et al.* 2002).

2. Methods

2.1 Taxon Sampling and Summary

Tissue samples were obtained for vouchered specimens of 22 sturgeon taxa (*A. o. desotoi, A. o. oxyrinchus, A. sturio, H. huso, H. dauricus, A. brevirostrum, A. gueldenstaedtii, A. naccarii, A. fulvescens, A. mikadoi, A. medirostris, A. schrenckii, A. transmontanus, A. sinensis, A. baerii, A. ruthenus, A. stellatus, S. albus, S. platorynchus, S. suttkusi, P. kaufmanni,* and *P. hermanni*) and one outgroup species (*P. spathula*). Total genomic DNA was extracted from muscle or, in some cases, gill tissue using the Qiagen DNeasy Blood & Tissue Kit following manufacturer's protocol (Qiagen 2006).

Methods described below were applied in an attempt to obtain full mitogenome sequences for each of these species. Full mitogenome sequencing was successful for all of these taxa except *A. baerii, A. ruthenus, A. stellatus,* and *P. hermanni.* Full mitogenome sequences for four ingroup species (*A. baerii,* Chen *et al.* 2011*; A. dabryanus,* Peng *et al.* 2007, *A. ruthenus,* Li *et al.* 2013; and *A. stellatus,* Arnason *et al.* 2007) and four outgroup species including both extant paddlefishes (*Polyodon spathula,* Inoue *et al.* 2003 and *Psephurus gladius,* Peng *et al.* 2007) and two Polypteriformes (*Polypterus delhezi,* Chen and Lin 2010 and *P. endlicherii,* Chen and Lin 2010) were available on GenBank® (Accession numbers are shown in Table 1). To evaluate the validity of sequences obtained from GenBank®, I sequenced a 16s segment of rRNA from tissue extracted from voucher specimens of *A. baerii* and *A. stellatus*. Sequencing failed for for *A. ruthenus* and *P. hermanni*. I aligned the sequences with the full mitogenome sequences on GenBank® and calculated the number of base pair differences. There were no base pair differences between the GenBank® *A. baerii* 16s sequence and the 562 base pair segment I sequenced. The GenBank® 16s sequence from *A. stellatus* was also a 100% match to the 610 base pair sequence I obtained.

2.2 DNA Purification and Sequencing of A. fulvescens *and* H. dauricus

I attempted to use a long PCR technique (Cheng *et al.* 1994), described below, to amplify the complete mitogenome of the total genomic DNA extractions. This method was successful only for *A. fulvescens* and *H. dauricus*.

Two general fish primers were designed and used to amplify the mitochondrial genome in two long-PCR reactions (Miya and Nishida 1996). Long PCR was done in a BioRad S1000 thermal cycler using a 20-µl reaction (2.4 µl extracted DNA, 2 µl of each primer, 3.2 µl dNTPs, 2 µl MgCL2, 6.24 µl water, 2 µl TaKaRa LA Taq ™ buffer, and 0.16 µl TaKaRa LA Taq Polymerase ™) and running 30 cycles of denaturation at 98° C for 10 seconds followed by annealing and extension combined at 60° C for 12 minutes. The two long PCR products from each of these species were mixed together and diluted with TE buffer (1:5) for use as template in subsequent PCR reactions.

Twenty-four sets of general fish primers (Inoue *et al.* 2003) were used for short PCR of long PCR products. PCR was carried out in a BioRad S1000 thermal cycler (Applied Biosystems) using a 10.4-µl reaction (1 µl dilute long PCR product, 1 µl of each primer, 0.8 µl dNTPs, 5.53 µl water, 1 µl TaKaRa Z Taq ™ buffer, and 0.07 µl TaKaRa Z Taq Polymerase ™) and using the following thermal-cycler protocol: initial denaturation at 94°C for 2 minutes, followed by 30 cycles of 98°C for 1 second, 50°C for 5 seconds, and 72°C for 15 seconds. PCR products were purified using a QIAquick PCR Purification Kit and following manufacturers' instructions. Direct-cycle sequencing using dye-labeled terminators (Applied Biosystems Inc.) was conducted using 36 internal sequencing primers. Labeled fragments were analyzed on an Applied Biosystems 3130xl Genetic Analyzer._DNA sequences were edited by eye and contigs were assembled using Sequencher 5.1. Sequences were annotated using mitoannotator in MitoFish (Iwasaki *et al.* 2013).

2.3 DNA Purification and Sequencing for Other Species

AnNGS method was used to obtain mitogenome sequences from the extracted total genomic DNA of *A. o. desotoi, A. o. oxyrinchus, A. sturio, P. kaufmanni, P. hermanni, H. huso, A. brevirostrum, A. gueldenstaedtii, A. naccarii, S. albus, S. platorynchus, S. suttkusi. A. mikadoi, A. medirostris, A. schrenckii, A. transmontanus,* and *A. sinensis*.

Before beginning library preparation for NGS, extracted DNA was quantified using the Qubit 2.0 fluorometer and the Qubit dsDNA BR Assay kit. Samples were prepared for sequencing using the Nextera XT DNA Library Preparation Kit v2 Set A following the manufacturers protocol for tagmentation of input DNA, PCR amplification and cleanup, and library validation, normalization, and pooling (illumine 2012). Tagmentation is an enzymatic reaction in which DNA is simultaneously fragmented and tagged with adapters; the subsequent PCR cycles amplify the tagged segments (illumina 2015). Libraries were sequenced using the Illumina MiSeq System. All MiSeq data were subsequently trimmed and merged in CLC Genomics workbench using the *de novo* assembly algorithm. This algorithm uses *de Bruijn* graphs, first identifying contigs (sets of overlapping sequecnce segments) in the data read in, creating a sequence of contigs, and mapping the reads through alignment to the contigs (QIAGEN 2016).

2.4 Alignments and Model Tests

MitoAnnotator (Iwasaki *et al.* 2013) was used to rotate each mitogenome so that all mitogenomes started at tRNA-Phe. Rotated sequences were aligned in MAFFT version 7. The resulting alignment was edited by eye in MacVector6 (Rastogi 1999). The edited alignment was used to generate a consensus sequence using MacVector6. The resulting consensus sequence was then re-annotated in MitoAnnotator to identify individual mitochondrial genes (Iwasaki *et al.* 2013). Based on this annotation, individual gene alignments were separated.

The program DAMBE was used to test for saturation in the full mitogenome alignment and individual gene alignments (Xia and Lemey 2009). Alignments were imported into DAMBE. First, proportion of invariant sites was estimated for each alignment using the option "Seq.Analysis|Substitution rates over sites|Estimate proportion of invariant sites" using the "new tree" and neighbor joining criteria. Actual tests for saturation were conducted using the option "Seq.Analysis|Measure substitution saturation|Test by Xia *et al.*" and entering the appropriate estimated proportion of invariant sites. Saturation was not detected in the full mitogenome or individual gene alignments.

The best models of nucleotide evolution were determined for each gene using in jmodeltest 2.1.3 (Posada 2008) and BIC criteria. The full mitochondrial genome was partitioned by gene, and the best model of evolution for each gene was used in all Bayesian analyses. Models used are listed in Table 2.

2.5 Phylogenetic Analyses

Parsimony analyses were run on the full mitogenome alignment in PAUP* 4.0 (Swofford 2003) with 1000 bootstrap replicates. Gaps were treated as missing data. Stepwise addition was used to obtain starting trees for branch swapping. Only the best trees found were kept. A 50% majority-rule consensus tree was calculated and is shown in the results section.

Maximum likelihood analyses were run on the full mitogenome alignment using the RAxML-HPC2 (Stamatakis 2006) on XSEDE tool on the Cipres Science Gateway (Miller *et al.* 2011). Gaps were treated as missing data. Nonparametric bootstrapping was enabled and the seed value was set at 12345. A general time-reversible model was identified by jmodeltest 2.1.3 (Posada 2008) as the most appropriate model parameter for the full mitogenome alignment, and was chosen using the setting "GTRCAT" to allow for rate heterogeneity. 1000 bootstrap replicates were run to generate support values. The best-scoring tree was retained, and is presented in the results section.

Bayesian phylogenetic analyses were run for the following four data configurations: 1) the edited alignments of the full mitogenome; and gene family trees made up of edited and concatenated alignments of 2) all protein coding genes; 3) ribosomal RNA sequences, and 4) transfer RNA sequences. Individual gene trees were also constructed using Bayesian methods for each individual mitogchondrial gene. Analyses were conducted in Mr.Bayes 3.2.2 (Huelsenbeck and Ronquist 2001). For each analysis, aligned sequences were partitioned by gene, with each gene assigned the best model of evolution as indicated by jmodeltest 2.1.3 (Posada 2008). For each dataset, between 500,000 and 3 million generations were run depending on the necessary number of generations for standard deviation on split frequencies to equal 0.0000 (Huelsenbeck and Ronquist 2001), the full mitogenome analysis required 3 million generations, the protein coding gene analysis was run for 1 million generations, and both tRNA and rRNA analyses were run for 500,000 generations. All analyses used 4 chains and a burn-in of 30% to allow chain convergence.

In the resulting phylogenetic hypothesis, confidence in each node was expressed in the form of posterior probability values. Distinct taxonomic groupings with relatively high posterior probability support values (>0.90) were recovered by analyses of each dataset. The full mitogenome topology required some re-analysis due to the recovery of several polytomies with low support values within one monophyletic taxonomic grouping (that overall had high support values). Reanalysis conducted for the Bayesian full mitogenome topology involved running analyses for the nine taxa in this group separate from the other sturgeon taxa. Analyses were run for one-million generations, and as with other Baysian analyses in this study, 4 chains and a burn-in of 30% were used. These analyses recovered a highly supported topology. I constrained the full mitogenome dataset including all taxa to include this topology, re-ran Bayesian analyses using a burn-in of 30% and running for one million generations, and recovered a well-supported, polytomy-free topology for all taxa. Relationships among other taxa did not change as a result of this constraint. The resulting consensus tree is presented as my preferred full mitogenome topology.

2.6 Tree Comparisons

To identify which genes most strongly influence the full mitogenome phylogenetic hypothesis, I calculated tree distance metrics between the topology resulting from the full mitogenome Bayesian analysis and those generated through alternate data sets. The Robinson-Foulds metric (Rzhetsky and Nei 1992; similar to the Penny-Hendy distance (Penny and Hendy 1985), which is defined as twice the number of bipartitions that differ between trees, is a tree distance metric used to quantify similarity and differences among trees (Paradis et al. 2004). Duchene *et al.* (2011) used it to identify which mitochondrial genes carry the same phylogenetic signal as the full mitogenome in studies of delphinid and killer whale phylogenies (Duchene *et al.* 2011). Because the Rzhetsky-Nei distance metric is well defined for use only in strictly bifurcating trees and many of my gene trees have polytomies, I employed a version of this metric modified to allow for polytomies, the Rzhetsky-Nei distance (Paradis *et al.* 2004, Rzhetsky and Nei 1992). Rzhetsky-Nei (R-N) distances were calculated between 1) my Bayesian mitogenome phylogeny and each of the gene-family topologies (protein-coding, tRNA and rRNA topologies), 2) each individual gene tree and the Bayesian full mitogenome tree, and3) each individual gene tree against the corresponding gene family tree (each individual protein-coding gene tree was run alongside the protein coding family topology). Calculations for all topology comparisons were done using the dist.topo function in the ape package (version 3.4) in R (Paradis *et al.* 2004, R Core Team 2014). An R-N distance of 0 indicates identical phylogenies. High Rzhetsky-Nei distances indicate divergent trees (Duchene *et al.* 2011). When conducting multiple tree comparisons, tree pairs with the lowest R-N distances are most similar to one another. Statistical significance tests are not available for R-N distances; this is a method for comparison between topologies only.

2.7 Comparisons to Previous Studies

The phylogenetic analyses of sturgeons conducted by Birstein *et al.* (2002) included a taxon composition close to that of the present study, and provides three topologies to compare with my results. They conducted parsimony and maximum likelihood analyses of 22 sturgeon species based on NADH5, cytochrome *b,* 12s, 16s, and D-loop (Figures 1, 2). They also conducted a combined-evidence phylogenetic analysis of 16 sturgeon species, based on the same molecular data and morphological data derived from Mayden and Kuhajda 1996 (Figure 3).

In order to compare Birstein *et al.*'s (2002) parsimony and maximum likelihood trees to my Bayesian, parsimony, maximum likelihood, 12s, 16s, cyt*b,* NADH5, and D-loop topologies, I pruned (using the command drop.tip in the ape package, version 3.4, Paradis *et al.* 2004, for R, R Core Team 2014) outgroups and sturgeon taxa not included in both studies from the trees . I ran Rzhetsky-Nei analyses on the pruned topologies to compare Birstein *et al.*'s (2002) parsimony and maximum likelihood tree with mine. I performed the same analytical methods using my gene trees and Birstein *et al.*'s (2002) parsimony and maximum likelihood trees to determine which genes most strongly influence their trees. I also compared my phylogenies and gene trees to Birstein *et al.*'s (2002) combined-evidence phylogeny. Birstein *et al.*'s combined evidence phylogeny included fewer taxa than their parsimony and maximum likelihood phylogenies, requiring pruning of additional taxa from my phylogenies and gene trees. I calculated R-N distances to determine the level of similarity between the different studies and to identify the genes most strongly represented in the combined evidence phylogeny of Birstein et al. (2002).

3. Results

3.1 Sequences and Alignment

The final, edited, aligned matrix for the full mitogenome was 14,105 base pairs in length (Appendix 1). It included the control region, or D-loop (1,295 base pairs), 13 protein coding genes (9,614 base pairs total), both rRNAs (2,676 bps long), 22 tRNAs (1,693 bps). Gene order was conserved across sturgeon mitogenomes. Substitution models, gene family, and gene length for each gene are presented in Table 2.

3.2 Full mitogenome results

3.2.1 Phylogenetic Hypothesis Support Values

The full mitogenome Bayesian phylogenetic hypothesis recovered the sturgeons as a clade with posterior probability (PP) support of 1.0 (Figure 4). Maximum likelihood recovered sturgeons as monophyletic with a bootstrap (BS) value of 100 (Figure 5), while parsimony recovered sturgeons as monophyletic, though with a BS value of only 50 (Figure 6).

3.2.2 Full Mitogenome Hypotheses

The Bayesian, maximum likelihood, and parsimony full mitogenome phylogenetic hypotheses had similar topologies, with only a few differences. For ease of discussion, these hypotheses will subsequently be referred to as the Bayesian, maximum likelihood, and parsimony topologies. All three full mitogenome topologies recovered *Scaphirynchus* as a clade in which *S. platorynchus* and *S. albus* are sister [*(S. suttkusi, (S. platorynchus, S. albus))*] (Bayesian PP= (1.0,(.78)), parsimony BS=(100,(88)), maximum likelihood BS=(1.0,(.95))). *Acipenser sturio* was recovered in a clade with the sister taxa *A. oxyrinchus* and *A. o. desotoi* [*(A. sturio, (A. oxyrinchus, A. o. desotoi))*] in all three topologies (support values of 100 in all). A clade with two sets of sister species was recovered in all the three analyses: *A. dabryanus* as sister to *A. sinensis* and *A. transmontanus* as sister to *A. schrenckii* [*((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii))*] (BS=100, PP=1.0). *Huso dauricus* was consistently recovered in a clade with the sister taxa *A. medirostris* and *A. mikadoi* [*(H. dauricus, (A. medirostris, A. mikadoi))*] (BS=100, PP=1.0). *Huso huso* was basal in a clade that contained *P. kaufmanni* as sister to *A. stellatus* [*(A. stellatus, P. kaufmanni)*] and also included [*(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii)))))*]. It was in this clade that some differences across topologies were found. Other differences were found in the relationships amongst the clades.

For ease of discussion among topologies, these clades (Table 4) will subsequently be referred to as:

Clade 1: (*H. huso,* Subclade 1.1, Subclade 1.2)
    Subclade 1.1: (*A. stellatus, P. kaufmanni)*
    Sublade 1.2: *(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii)))))*
Clade 2: *(H. dauricus, (A. medirostris, A. mikadoi))*
Clade 3: *((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii))*
Clade 4: *(A. sturio, (A. oxyrinchux, A. o. desotoi))*
Clade 5:(*(S. suttkusi, (S. platorynchus, A. albus))*).

Although the three full mitogenome analyses (Bayesian, Parsimony, and Maximum Lieklihood, Figures 4-6) all recovered Clade 1, the position of this clade varied among phyogenies. All three topologies recovered Subclade 1.1, *P. kaufmanni* and *A. stellatus* as sister, and Subclade 1.2, *(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))).* The Bayesian and maximum likelihood analyses recovered Clade 1 (within which *H. huso* was basal) as sister to Clade 2. The maximum parsimony analysis recovered Subclade 1.1, Subclade 1.2, and *H. huso* to be in a polytomy.

Clades 2 and 3 were consistently recovered as sister to each other in all analyses. The Bayesian analysis recovered Clade 4 as sister to this group and Clade 1 as sister to the Clade 2, 3, and 4 grouping. Parsimony analyses recovered an opposing topology, with Clade 1 as sister to Clades 2 and 3 and Clade 4 as sister the the three clade grouping. Both Bayesian and parsimony phylogenies recovered Clade 5 *((S. suttkusi, (S. platorynchus, S. albus)))* as sister to the other sturgeons. In contrast, the maximum likelihood analysis recovered Clade 5 as sister to Clade 1. Clades 2 and 3 were recovered as sister to Clade 1 + 5 group, with Clade 4 as sister to all other sturgeons.

I favor the topology resulting from Bayesian analyses for comparison to individual gene trees for several reasons. First, this method allowed me to specify the best model of nucleotide evolution for the full mitogenome and the individual gene analyses. Second, it showed slightly more phylogenetic structure than the parsimony phylogeny (which contained a polytomy). Finally, multiple previous datasets, including molecular, morphological, and combined evidence data, support a *Scaphirhynchus* as the sister lineage to all other sturgeons, as recovered in this topology.

3.3 Gene Tree Analysis and Comparison to the Full Mitogenome

The Rzhetsky-Nei distances between the Bayesian full mitogenome topology and the Bayesian protein-coding, tRNA, and rRNA topologies were 1, 10, and 12 respectively (Table 3). This indicates that the protein-coding gene family produced a phylogenetic hypothesis most similar, but not identical, to that of the full mitogenome. The tRNAs had the second most similar topology. The rRNA gene family tree was the least similar to the full mitogenome phylogeny.

3.3.1 tRNA Gene Family Phylogeny

Aligned and edited sturgeon tRNAs ranged from 67 to 165 base pairs in length (Table 2), and the complete tRNA dataset was 1,693 base pairs long. The Bayesian tRNA gene family phylogeny (Figure 7) recovered the same composition and structure within Clades 2, 4, and 5 (Table 4) as the full mitogenome Bayesian analysis. Clade 3 was found to be monophyletic, but the sister grouping of *A. transmontanus* and *A. schrenckii* was dissolved, so that these two species formed a polytomy with *(A. dabryanus, A. sinensis).* As in the full mitogenome topology, Clades 2 and 3 were recovered as sister, and Clade 5 (*Scaphirynchus)* was sister to all other extant sturgeon taxon.

In the tRNA phylogeny, Clade 1 was nonmonophyletic. Subclade 1.1 was also nonmonophyletic; *P. kaufmanni* was recovered as sister to *H. huso.* Four species of Subclade 1.2 grouped together, with a structure similar to that recovered in the full mitogenome analysis: *(A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii)))*; these four species formed one branch of a five branch polytomy. The other branches are: *A. stellatus; A. fulvescens,* Clades 2 and 3 as sister; and *H. huso* and *P. kaufmanni* as sister. *Acipenser ruthenus* (full mitogenome Subclade 1.2) was recovered as sister to this polytomy. Clade 4 was recovered as sister to all other sturgeons except *Scaphirynchus.*

3.3.2 rRNA Gene Family Phylogeny

The aligned, edited rRNA sequence was 2,676 base pairs long, with a 960 bp 12s and a 1716 bp 16s. The Bayesian rRNA gene family phylogeny (Figure 8) recovered the same composition and structure within Clades 2-5 (Table 4) as the full mitogenome Bayesian analysis. Clade 1 was monophyletic, but the structure was different from that recovered by the full mitogenome. Clades 2 and 3 were recovered as sister taxa and *Scaphirynchus* (Clade 5) was found to be sister to all other extant sturgeons. Clades 2 and 3 were sister to Clade 1 (these were recovered as sister to Clade 4 in analysis of the full mitogenome).

Within Clade 1, Clades 1.1 and 1.2 were found to be nonmonphyletic in the rRNA phylogeny. *Acipenser brevirostrum* and *P. kaufmanni* were recovered as sister taxa in Clade 1. Rather than being found sister to other Clade 1 taxa, as in the full mitogenome phylogeny, *H. huso* was recovered as sister to *A. ruthenus*, with *A. stellatus* as sister to them. Sister to this grouping was the clade *(A. fulvescens, (A. baerii, (A. naccarii, A. gueldenstaedtii))).* Although the relationships among *A. baerii, A. naccarii,* and *A. gueldenstaedtii* match those recovered in the full mitogenome analysis, a close relationship of *A. fulvescens* with these taxa does not.

3.3.3 Protein-coding Gene Family Phylogeny

Lengths of the edited, aligned gene sequences varied from 169 to 1422 base pairs, with the protein coding gene family alignment having a total of 9,614 base pairs. The resulting Bayesian protein-coding gene family phylogeny (Figure 9) recovered Clades 1-4 from the full mitogenome topology, with the same within-clade structure (Table 4). A monophyletic *Scaphirynchus* was recovered, although relationships among the species were unresolved. Clades 2 and 3 were recovered as sister taxa, as in the other topologies. Clade 1, however, was recovered as sister to Clade 5. Further diverging from the full mitogenome topology, analysis of the protein-coding genes recovered Clade 4 as sister to all other extant sturgeons.

3.4 Tree Comparisons: Influence of Individual Gene Regions on Gene Tree topology

Among tRNAs, the lowest R-N distance (indicating the most similar topology) between a tRNA gene tree and the tRNA gene family phylogeny was that of tRNA Asp (R-N=11). The tRNAs Cys, Phe, and Ser (segment 1) all had the next shortest distances (R-N= 12). Of the two rRNAs, 16s produced a tree most similar to that of the rRNA phylogeny, with a R-N distance of 11; that of 12s (R-N= 12) was slightly higher.

Among the protein coding genes, ATP6 had the shortest R-N distance (1) to the full protein coding phylogeny. NADH5 had the second shortest distance of 5. COIII, NADH1, and NADH4L all had distances of 8, and NADH2 had a distance of 9. Both NADH3 and NADH4 had R-N distances of 11 to the protein coding family phylogeny, and ATP8 and COII were each a R-N distance of 12 to the protein-coding phylogeny.

3.5 Gene Influence on Bayesian Full Mitogenome Tree

3.5.1 Gene Families

The R-N distances suggest that the protein coding genes as a group likely influence the topology of the full mitogenome phylogenetic hypothesis most strongly, followed by the tRNAs, with the rRNA gene family being least influential (Table 3). As a unit, the protein coding genes recovered, with strong support (posterior probability values >0.98), all five clades that were also recovered by the full mitogenome. The phylogenetic structure within all but one of these groups was identical to that recovered by the full mitogenome, although the structure among the clades differed greatly from that of the full mitogenome. For example, the protein-coding genes recover *Scaphirhynchus* nested within all other sturgeons, whereas Bayesian analysis of the full mitogenome found it to be the sister group of all other sturgeons.

In contrast, the tRNA analysis recovered *Scaphirhynchus* as sister to the other sturgeons. This gene family, however, only recovered four of the five monophyletic groupings (Clades 2-5) found by the full mitogenome. Structure matching that of the full mitogenome was recovered with strong support in Clades 2 and 4 (PP >.99) and with weak support in Clade 5 (*Scaphirhynchus*, PP= 0.78). Only partial structure (one of two sister taxa) was recovered in Clade 3.

The rRNA data recovered all five monophyletic groupings that were recovered by the Bayesian analysis of the full mitogenome, and identical structure to the full mitogenome in four of these (PP >0.95) (Clades 2-5). Clade 1 was recovered as monophyletic but the structure within this grouping does not match that found in analysis of the full mitogenome and support within Clade 1 was low (average posterior probability <0.80 for nodes within this clade).

3.5.2 Individual Genes

Comparisons between individual gene topologies and the full mitogenome hypothesis found that the 25% best-scoring genes (i.e., most-similar topologies) were 16s rRNA and eight protein-coding genes (ATP6, NADH5, CO3, NADH1, NADH4L, NADH2, NADH4, and NADH3) (Table 5). These genes exhibit strong phylogenetic signal matching that of the full mitogenome.

Protein coding genes made up 13 of the 14 genes recovering topologies most similar to that of the full mitogenome. ATP6 and NADH5 were the two genes with topologies most similar to that of the full mitogenome, having Rzhetsky-Nei scores of 2 and 6, respectively. COIII, NADH1, NADH4L , and 16s, had Rzhetsky-Nei scores of 9. The tRNAs Asp, Cys, Phe, Ser1, and His fell within the 50% best scoring genes, indicating that they had the most similar topologies to the full mitogenome; Asp had the shortest Rzhetsky-Nei distance (15) . D-loop had the greatest Rzhetsky-Nei distance (33) from the full mitogenome phylogenetic hypothesis, indicating a lack of similarity between the topologies. This non-coding region recovered only one clade also found by the full mitogenome phylogenetic hypothesis. Additional structure was present in this phylogeny, but this structure is in conflict with that of the full mitogenome and had poor support values (PP <0.90 for more than half the clades).

Evolutionary rate of gene regions does not seem to correlate with the recovery of the same topology as the full mitogenome. Four models of nucleotide rate substitution (Jukes-Cantor, General Time Reversible, Hasegawa-Kishino-Yano, and Kimura 2-parameter) and four models of rate heterogeneity (equal, gamma, some sites invariable, and some sites invariable with the rest drawn from gamma-distributionTable 2) were represented across all genes. Both the most similar (ATP6) and least similar (D-loop) gene trees to the full mitogenome had the same models of nucleotide substitution.

A Pearson's product-moment correlation test between Rzhetsky-Nei distance to the full mitogenome phylogenetic hypothesis and the length (in base pairs) of each gene revealed a significant linear relationship ($r=0.59$, $p<0.05$). This test indicates a relationship between the similarity of a single gene tree to the full mitogenome topology and the length of the gene sequence, with longer genes carrying more phylogenetic signal (similar to that of the full mitogenome) than shorter genes. This contradicts previous studies, where gene length was considered to be a poor estimator of phylogenetic signal (Zardoya and Meyer 1996).

3.6 Comparison to Birstein et al. (2002)

3.6.1 Parsimony

The molecular phylogenetic hypothesis of Birstein *et al.* (2002) based on a maximum parsimony analysis contains multiple polytomies made up of 9 clades, named here as Polytomy A, B, and C and Clades 1-9 (Figure 1, Table 6). The most derived polytomy, Polytomy A, comprises five clades: A1) *(A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. persicus, A. naccarii))))*, A2) *(A. stellatus, Pseudoscaphirynchus)*, A3) *(H. dauricus, (A. ruthenus, H. huso)*, A4) *A. fulvescens,* and A5) *A. nudiventris.* Sister to this polytomy is *A. mikadoi.* This grouping is in another polytomy, Polytomy B, with two additional clades, B1) *(A. schrenckii, (A. transmontanus, A. medirostris))* andB2) *A. sinensis.* A third polytomy, Polytomy C, comprises Polytomy B and two more clades, C1) *(A. sturio, A. oxyrinchus)* and C2) *Scaphirynchus.* The clade containsing all three species *Scaphirynchus* species (*S. platorynchus, S. albus,* and *S. suttkusi)* recovered these species in a polytomy.

Comparing Birstein *et al.*'s (2002) parsimony topology with the results of my parsimony analysis reveals broad similarity in the composition of monophyletic groupings (not considering taxa that are missing from either topology). With the exception of *H. dauricus,* all of the taxa in Birstein *et al.*'s (2002) Polytomy A were found to form a clade in my topology. All of the taxa included in Polytomy B were also found to be monophyletic in my topology, although *H. dauricus* was recovered with *A. sinensis, A. transmontanus, A. schrenckii, A. medirostris,* and *A. mikadoi* instead of with Polytomy A as in Birstein *et al.* (2002).

Reviewing the topologies clade-by-clade reveals, on a finer scale, more similarities and several differences. Birstein *et al.*'s (2002) Clade A1 is nearly identical to Clade 1 in my topology both in composition and structure, except that I did notinclude *A. persicus* and *A. nudiventris* in my study. Birstein *et al.*'s (2002) Clade A2, the sister-species relationship between *A. stellatus* and *Pseudoscaphirynchus,* was also recovered in my topology, although Birstein *et al.* (2002) included both *P. hermanni* and *P. kaufmanni* in their study, and I only included *P. kaufmanni.* As in Birstein *et al.*'s study, my results recovered these two clades in a broader monophyletic grouping including *H. huso* and *A. ruthenus.* Unlike the topology recovered by Birstein *et al.* (2002), in my analysis *A. fulvescens* (CladeA4) was found to be sister to Clade A1, with *H. huso* sister to that grouping. *Huso dauricus* was not monophyletic with *H. huso* and *A. ruthenus* (Clade A3). Birstein *et al.*'s (2002) Clade B1 and B2 were not recovered by my analysis. Instead, *A. mikadoi* (Clade C2), *A. medirostris* (Clade B2), and *H. dauricus* (Clade 3) were found to be monophyletic and sister to *A. dabryanus* (not included by Birstein *et al.* (2002), *A. sinensis* (Clade C1), *A.transmontanus* (Clade B2), and *A. schrenckii* (Clade B2).

Birstein *et al.*'s (2002) monophyletic group *A. sturio* + *A. oxyrinchus* was reflected in my parsimony topology, which included both subspecies of *A. oxyrinchus.* Similarly, Birstein *et al.*'s (2002) monophyletic *Scaphirynchus* was included in my topology, although they recovered these taxa in a polytomy, whereas my hypothesis found *S. platorynchus* and *S. albus* to be sister taxa. Birstein *et al.* (2002) based their parsimony analysis on five genes: 12s, 16s, D-loop, cyt*b,* and NADH5. Comparison of the single gene-trees I generated indicated that four of these genes 12s, D-loop, cyt*b,* and NADH5 also recovered the three *Scaphirhynchus* species in a monophyletic polytomy. A lack of structure among the taxa composing Polytomy A was reflected in the 12s, D-loop, and 16s gene trees. Rzhetsky-Nei distances indicate that of the five individual genes included in Birstein *et al.*'s analysis, the NADH5 gene tree has the shortest distance from, and is therefore most similar to, Birstein *et al.*'s (2002) parsimony topology (R-N = 17), followed by D-loop and 16s (both have R-N = 20). These genes likely influence the topology recovered by Birstein *et al.* most strongly.

The differences of topology and resolution between my parsimony tree and Birstein *et al.*'s parsimony Tree may be related to the amount of data used. Of the five genes Birstein *et al.*(2002) used in their phylogeny, I found two to be inconsistent with the full mitogenomic phylogeny (12s and 16s, discussed above). Rzhetsky-Nei analyses indicate that of the thirteen genes that most strongly influence the results of my parsimony analysis, eleven were found to produce topologies similar to that of the full mitogenome. These results support the finding of previous studies that an increase in data improves resolution in phylogenetic analyses (Hajibabei *et al.* 2007).

3.6.2 Maximum Likelihood

Birstein *et al.*'s (2002) maximum likelihood topology recovered several clades referred here as clades ML1-6 (Figure 2, Table 7). Solely *Acipenser* species made up Clade ML1: *(A. fulvescens, (A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. naccarii, A. persicus))))).* Clade ML1 was recovered as sister to Clade ML2: *(A. stellatus, (P. kaufmanni, P. hermanni)),* with *A. nudiventris* as sister to this group. Clade ML3: *(H. dauricus, (H. huso, A. ruthenus))* was found to be sister to Clades ML1, ML2, and *A. nudiventris*; the three species of *Scaphirynchus* form a polytomy,Clade ML4, which was found to be the sister group of Clades ML1, ML2 and ML3. *A. mikadoi* is sister group of to Clades ML1-ML4. Four *Acipenser* species form Clade ML5: *(A. medirostris, (A. sinensis, (A. transmontanus, A. schrenckii))).* CladeML 6, *(A. oxyrinchus, A. sturio),* was found to be the sister group of all other sturgeons.

Comparing Birstein *et al.*'s (2002) maximum likelihood topology to my full mitogenome maximum likelihood topology reveals a few similarities and many differences, with an R-N distance of 27. Clade ML1 is identical to my Subclade 1.2, except that *A. persicus* was not included in my analyses. Likewise, Clade ML2 is identical to the sister grouping my analyses recover between *Pseudoscaphirynchus* and *A. stellatus,* except that my study did not include *P. hermanni.* The species in Clade ML3 of Birstein *et al.*'s (2002) phylogeny are not recovered in a clade in my analyses. Instead, in my topology, *A. ruthenus* was recovered as sister to Clade ML1, and *H. huso* was sister to Clades ML1 + ML2. *Huso dauricus* was found to be sister to *A. mikadoi* and *A. medirostris* (which was recovered by Birstein *et al.* (2002) in Clade ML5). Like Birstein *et al.*'s (2002) maximum likelihood topology, my analyses recovered a monophyletic *Scaphirynchus* as sister to the taxa in Clades ML1, ML2, and ML3 (except for *H. dauricus). Scaphirynchus albus* and *S. platorynchus* were recovered as sister species within this clade in my topology. The taxa Birstein *et al.* (2002) recovered in Clade 5 were not recovered as monophyletic in my analyses. In my topology, *A. medirostris* was recovered as sister to *A. mikadoi* in a clade also containing *H. dauricus* (Clade ML3). This clade was recovered as sister to a monophyletic group containing three of Birstein *et al.*'s (2002) Clade ML5 species (*A. transmontanus, A. schrenckii,* and *A. sinensis)* along with *A. dabryanus,* which was not included in Birstein *et al.*'s (2002) study. As in Birstein *et al.*'s (2002) topology, though, my phylogenetic hypothesis recovered *A. sturio* + *A. oxyrinchus* as the sister group of all other sturgeons.

Of the genes included in Birstein *et al.*'s (2002) analyses, NADH 5 appears to be the most influential in the maximum likelihood topology (R-N = 17), followed by 16s (R-N= 18), cyt*b* (R-N= 21), D-loop (R-N= 22), and finally 12s (R-N= 23). D-loop is the only one of these genes that recovered *Scaphirhynchus* within *Acipenser.* I consider the recovery of *Scaphirhynchus* within *Acipenser* to be a significant difference between these two topologies and the other gene trees.

3.6.3 Combined Evidence

Birstein *et al.*'s (2002) combined evidence topology, based on both molecular and morphological data, recovered five distinct clades (Table 8). Clade CE1 included the species *(A. baerii, A. brevirostrum)* with Clade CE2 *(A. stellatus, (P. hermanni, P. kaufmanni)* and Clade CE3 (*H. dauricus, (A. ruthenus, H. huso))* as sequential sister groups. *Acipenser fulvescens* was found to be sister to (Clade CE3, (Clade CE2, Clade CE1)). *Acipenser sinensis* + *A. transmontanus* (Clade CE4) formed the sister group of the previously described clades. *Acipenser oxyrinchus* was sister to Clades CE1-4. *Scaphirynchus* (Clade CE5), with *S. albus* and *S. platorynchus* as sister taxa, was the sister group of all other sturgeons.

Based on Rzhetsky-Nei analyses, the Birstein *et al.* (2002) combined evidence phylogeny (molecular and morphological data), is most similar to the topology based on my parsimony analysis (R-N= 9), than to the results of my Bayesian analysis (R-N= 12), and least similar to my maximum likelihood results. The results of my parsimony analysis suggests that *A. oxyrinchus* and *A. sturio* form the sister group of all sturgeons other than *Scaphirhynchus*, as does the Birstein *et al.* combined evidence phylogeny. This similarity is absent from my Bayesian results. This illustrates that the algorithm selected for phylogenetic analysis dictates a role in the relationships recovered.

Rzhetsky-Nei analyses suggest that, of the genes considered, Birstein *et al.*'s (2002) combined evidence phylogeny is most strongly influenced by NADH5 (R-N=10), then by 16s (R-N=11), followed by 12s and cyt*b* (R-N=16), with D-loop being the least influential (R-N=28).

4. Discussion

4.1 Phylogenetic structure within Acipenseridae

Several studies have proposed varying relationships regarding *Huso huso.* Based on skeletal characters, Findeis (1997) recognized a subfamily Husinae, containing the two species in the genus *Huso* (*H. huso* and *H. dauricus*), with this subfamily as sister to all other Acipenseridae. *Huso* since has been found to be nonmonophyletic (Krieger *et al.* 2008), with *H. huso* included in *Acipenser* (Vasil'eva *et al.* 2009). Relationships between *H. huso* and various species or species groups have been recovered in past analyses, including those similar to the monophyletic grouping I recovered here and discussed above. For example, based on morphological data, Hilton *et al.* (2011) defined the subfamily Husinae as including only *H. huso* and *A. ruthenus,* with *A. baerii* as its sister-group. Although I recover *H. huso* in a clade with these species, these data and analyses do not support Husinae *sensu* Hilton et al. (2011).

My recovery of *(H. huso, ((Acipenser stellatus, Pseudoscaphirynchus kaufmanni), (A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii))))))* illustrates a potential new sturgeon clade containing significant phylogenetic structure. This result supports the previous recovery of *Pseudoscaphirynchus* and *A. stellatus* as a clade in both morphological (Hilton *et al.* 2011) and molecular studies (Birstein *et al.* 2002, Dillman *et al.* 2007, Krieger *et al.* 2008). My results support the recognition of Pseudoscaphirhynchinae, proposed by Hilton *et al.* 2011, that includes *Pseudoscaphirynchus* and *A. stellatus.*

Recovery of *H. huso* within a group of *Acipenser* species is not surprising, as *H. huso* has been frequently recovered within *Acipenser* (e.g. Ludwig *et al.* 2001, Zhang *et al.* 200, Fontana *et al.* 2001, Hilton *et al.* 2011, Dillman *et al.* 2007, Krieger *et al.* 2008). Krieger *et al.* noted that two positions of *Huso* within *Acipenser* have been proposed: 1) *Huso* clustered with *A. ruthenus;* and

2) *Huso* basal to other *Acipenser* species. Krieger *et al.* (2008) first recovered *H. huso* clustered with *A. ruthenus*. After examination of sequences used in analyses, Krieger *et al.* (2008) suggested that this position could be a result of past introgression. When Krieger *et al.* (2008) excluded *A. ruthenus* from analyses, *H. huso* was found as sister to other *Acipenser* species. Krieger *et al.* (2008) analyzed sequences of specific gene segments used in their and Birstein *et al's.* (2002) analyses, and found that the 12s sequences of the two taxa had only three differences. In my analyses *A. ruthenus* was recovered in the same clade as *H. huso* by full mitogenome and protein-coding data, but *H. huso* was found to be basal to the other taxa within this clade, with *A. ruthenus* more deeply nested in this clade. My rRNA topology, however, recovered *H. huso* and *A. ruthenus* as sister taxa within a broader *Acipenser* clade. My 12s topology recovered *H. huso* clustered with *A. ruthenus* and *A. stellatus,* whereas 16s recovered *H. huso* as sister to an *Acipenser* species clade. None of the other gene trees recovered a close relationship between *H. huso* and *A. ruthenus*. Although further investigation might reveal more regarding the position of *Huso,* my study suggests that a close or sister relationship with *A. ruthenus* is likely influenced by 12s. My analyses ultimately support an *Acipenser* species clade in which *H. huso* occupies a sister position.

My full mitogenome and protein coding data configurations recovered a core group of taxa clustered together: *Acipenser gueldenstaedtii, A. naccarii, A. baerii, A. brevirostrum,* and *A. fulvescens.* These species have been recovered together by several previous studies. Ludwig *et al.* (2001) and Fontanta *et al.* (2001) used molecular analyses and recovered these taxa (along with *A. persicus*) with the same structure as my full mitogenome and protein coding phylogenies. This clade is not novel, but my data support previous phylogenetic hypotheses including this configuration of taxa.

The grouping of *A. medirostris, A. mikadoi*, and *H. dauricus* was recovered here and previously in analyses by Krieger *et al.* (2008) and Ludwig *et al.* (2001). The separation between *H. dauricus* and *H. huso* provides further support for the dissolution of a monophyletic *Huso*.

I recovered a monophyletic pair of sister species: *((A. dabryanus, A. sinensis), (A. schrenckii, A. transmontanus))*. A close relationship among *A. schrenckii, A. sinensis*, and *A. transmontanus* was described by Birstein and DeSalle (1998) and Krieger *et al.* (2008). Clades including three of these taxa were recovered by Ludwig *et al.* (2008) and Fontana *et al.* (2008). A sister relationship between *A. sinensis* and *A. dabryanus* has been supported previously by morphological data, protein coding genes, and biogeographic analyses (Artyukhin 1995, 2006, Choudhury and Dick 1998, Zhang *et al.* 2000).

The two subspecies of *A. oxyrinchus (A. o. oxyrinchus* and *A. o. desotoi*) and *A. sturio* form Clade 4. Historically *A.oxyrinchus* had been considered a subspecies of *A.sturio* (Birstein and DeSalle 1998, Ludwig *et al.* 2001, Fontana *et al.* 2001, Krieger *et al.* 2008, Choudhury and Dick 1998, Artyukhin 2006). The recovery of these two in a clade together is therefore not surprising and supports previously hypothesized sturgeon phylogenies.

All four of the main data configurations recovered *Scaphirhynchus* as a monophyletic genus including *S. albus, S. platorhynchus,* and *S. suttkusi.* My recovery of *Scaphirhynchus* as the sister group of all other sturgeons (by all data configurations except for the protein-coding sequences) supports the dissolution of the tribe Scaphirhynchini (sensu Findeis, 1997), as previously proposed (Birstein *et al.* 2002). *Scaphirhynchus* has been recovered as basal to other sturgeons by molecular analyses, morphological studies, and studies combining both types of data. This topology was recovered in a strictly molecular analysis based on cyt*b,* 12s, and 16s, and in a combined-data study using these genes, NADH5, D-loop, and morphological data. A morphological study including extant and extinct sturgeons recovered two equally most parsimonious trees, one with each topology (Hilton *et al.* 2011). The authors favored the topology recovering *Scaphirhynchus* as sister to all other extant sturgeons. Some previous studies have instead considered *Scaphirhynchus* to be within*Acipenser;* these include studies based on cyt*b* and studies using cytochrome *b with* cytochrome oxidase 2, NADH5, D-loop, both rRNAs, and the tRNAs Phe and Asp (Fontana *et al.* 2001, Krieger *et al.* 2008).

Most of my analyses recover *S. albus* and *S. platorhynchus* as sister taxa within the monophyletic genus *Scaphirhynchus*. Support for this relationship is especially strong (e.g., PP=1) in the full mitogenome and rRNA analyses. This is consistent with previous studies combining morphological and molecular data (Birstein *et al.* 2002), but conflicts with studies relying exclusively on protein-coding genes or morphological data. Based on morphological characters, Mayden and Kuhajda (1996) recovered *S. suttkusi* and *S. platorhynchus* as sister taxa relative to *S. albus.* Ludwig *et al.* (2001) recovered that same sister relationship using exclusively cyt*b*, but Fontana *et al.* (2001), who also used cyt*b* instead found *S. suttkusi* to be sister to *S. albus* relative to *S. platorhynchus.* My data support a monophyletic *Scaphirhynchus* sister to other extant sturgeons, with a sister relationship between *S. platorhynchus* and *S. albus*, as previously recovered by Birstein *et al.* (2002).

My full mitogenome analysis recovered a large clade containing *A. oxyrinchus, A. sturio, A. schrenckii, A. transmontanus, A. sinensis, A. dabryanus, A. mikadoi, A. medirostris,* and *H. dauricus.* Within this group are three well-supported, previously defined clades. Two of these clades, (*H. dauricus, (A. medirostris, A. mikadoi*) and ((*A. dabryanus, A. sinensis), (A. transmontanus, A. schrecknii)),* were recovered as sister by all of my major data configurations. This sister relationship has been previously recovered by Birstein and DeSalle (1998) based on protein coding and rRNA sequences. A sister relationship between these grouping (excluding *A. dabryanus)* was found by Ludwig *et al.* (2001).

4.2 Influence of Individual Genes on Phylogeny

All of the clades recovered by the full mitogenome phylogenetic hypothesis were also recovered by ATP6, and all but one had the same structure exhibited in the full mitogenome phylogenetic hypothesis. It is unsurprising that NADH5 also yielded one of the most similar gene trees to the full mitogenome phylogenetic hypothesis; previous studies have found this gene to recover mammal, dolphin, orca, teleostean, and sturgeon phylogenies similar to that recovered by the full mitogenome (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). COIII, NADHI, and NADH4L recovered topologies similar to that found by the full mitogenome phylogenetic hypothesis. CO3 was previously found to recover teleostean phylogenies similar to that of the full mitogenome (Miya and Nishida 2000) and NADH1 was useful in testing the phylogeny of Delphinae (Duchene *et al.* 2011). Previously, Rastorguev *et al.* (2008) found that NADH4L did not recover a phylogeny very similar to that of the full mitogenome topology of Russian sturgeons. The relatively high (compared to other protein trees) tree distance results from this study for ATP8, COII, NADH2, NADH3, and NADH6 (R-N = 10 to 14) correspond with conclusions of similar studies that although these genes carry similar signal to that ofthe full mitogenome in various taxa, the signal carried by other protein-coding genes is more similar (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). Surprisingly, I obtained relatively high Rzhetsky-Nei distances between the full mitogenome and both cyt*b* and COI. These genes have been consistently found by previous studies to recover topologies similar to that of the full mitogenome (in studies of sturgeons, teleosts, dolphins, and terrestrial mammals) (Zardoya and Meyer 1996, Miya and Nishida 2000, Rastorguev *et al.* 2008, Duchene *et al.* 2011). Although cyt*b* and COI did recover similar topologies to the full mitogenome in this study, the topologies recovered by 16s and the other protein coding genes were more similar to that of the mitogenome.

Results of my tree comparison indicate that individual mitogenome segments with varying functions strongly influence the topology of the full mitogenome phylogenetic hypothesis. Further, segments in the same gene family (rRNA, tRNA, protein-coding) do not necessarily carry the same phylogenetic signal. Therefore, I cannot recommend restricting mitochondrial phylogenetic studies by gene family. Instead, I underscore conclusions of previous studies indicating that as a group, the protein coding genes carry similar signal to the full mitogenome, but this signal is not identical to that of the full mitogenome (when all genes are considered). In addition, 16s rRNA recovered a topology similar to that of the full mitogenome. Further studies, however, are necessary to determine the relative influence of signal carried by 16s rRNA, as it relate to the full mitogenomic signal, beyond this family.

4.3 Taxon coverage

This study represents almost complete taxon coverage for extant sturgeons. The three extant (*P. fedtschenkoi* is presumed extinct) sturgeons missing from this analysis are *P. hermanni, A. persicus*, and *A. nudiventris* (IUCN 2014). While the absence of species from a topology represents missing phylogenetic signal that, if present, might change the topology, the omission of *P. hermanni* and *A. persicus* from my analyses might be of little consequence. *Pseudoscaphirhynchus hermanni* has been recovered, by both morphological (Hilton *et al.* 2011) and molecular datasets (e.g. Dillman *et al.* 2007, Birstein *et al.* 2002), as sister to *P. kaufmanni,* which was included in my study. Molecular characters recovering these taxa as sister taxa include three of the genes I found to carry similar phylogenetic signal to the full mitogenome phylogenetic hypothesis (NADH5, 16s, and cyt*b;* Birstein *et al.* 2002). Because previous studies based on protein coding genes have recovered *P. hermanni* and *P. kaufmanni* as sister, I predict that if *P. hermanni* had been included in my study the same sister relationship would have been recovered. This would likely not change the rest of the topology.

Also missing from my analyses is *A. persicus*, which has been regarded as a subspecies of *A. gueldenstaedtii* rather than a valid species (Birstein *et al.* 2005). A recent study focused on the taxonomic status of *A. persicus* employed morphological data and two mitochondrial genes, cyt*b* and D-loop, and concluded that *A. persicus* is not a valid species (Ruban *et al.* 2008). The authors of the study (Ruban *et al.* 2011) argued, based on both morphological and molecular data, that these are conspecific. Molecular studies based on cyt*b*, 12s, 16s, NADH5, and D-loop failed to recover *A. persicus* as a monophyletic species separate from *A. gueldenstaedtii* (Birstein and DeSalle 1998; Birstein *et al.* 2000; Birstein and Doukakis 2001, and Birstein *et al.* 2005). My gene tree comparison analyses found that three of the genes supporting this conclusion (NADH5, 16s, and cyt*b*) reflect the full mitogenome phylogenetic hypothesis. I hypothesize that if included in my study, *A. persicus* would fall as a sister species to *A. gueldenstaedtii*, and would not significantly alter the overall topology.

Acipenser nudiventris has been recovered as sister to various taxa across phylogenetic studies. Ludwig *et al.* (2001) and Fontana *et al.* (2001) both recovered it as sister to *A. ruthenus* based on cyt*b* data. Using 12s and 16s sequences, Birstein and DeSalle (1998) found *A. nudiventris* to be sister to *A. dabryanus*. Employing more sequence data, however, both Krieger *et al.* (2008; using cyt*b*, 12s, cytochrome c, the tRNAs Phe and Asp, NADH5, and D-loop) and Birstein *et al.* (2002; in their combined data set) found *A. nudiventris* to be sister to *(P. kaufmanni, A. stellatus)*. Both studies used cyt*b*, which I found to carry a signal similar to that of the full mitogenome. Birstein *et al.* (2002) also employed NADH5, the second best scoring gene in my study for recovering full-mitogenome topology, and 16s, which also recovered a topology similar to that of the full mitogenome analysis. I consider it likely that had I been able to include

A. nudiventris in my analyses, it would have been recovered as closely related to both *Pseudoscaphirhynchus* and *A. stellatus*. I recommend that future studies include *A. nudiventris*, ideally with full mitogenomic data or, at least, with the full protein-coding family and 16s data, to resolve its relationship.

4.4 Mitogenome Coverage

Although two other studies included more sturgeon species than mine (Birstein *et al.* 2002, Krieger *et al.* 2008), I included the greatest amount of molecular data applied to sturgeon phylogeny to date, over 14,000 base pairs. More data may not always be better, but various studies have shown that longer mitochondrial sequences produce higher resolution phylogenies (Vilstrup *et al.* 2011).

4.5 Use of Mitochondrial Sequences: Benefits and Drawbacks

Most concerns regarding the use of mitochondrial DNA sequences in phylogenetic reconstruction have been addressed in this study. Although the high substitution rate is of concern due to potential saturation and subsequent loss of lineage information, genes with high substitution rates are particularly useful in determining relationships between genetically similar species (Moore 1995). Nuclear genes are also at risk for loss of evolutionary information, though; recombination can lead to loss of the phylogenetic signal reflecting true evolutionary history (Rautenberg *et al.* 2008). The issue of recovering poorly resolved phylogenies due to use of short sequence segments is addressed easily. Addition of more sequences, for example the use here of the full mitogenome, leads to higher resolution in the phylogenies produced. This has been illustrated in studies across taxa, and is illustrated here by the number of polytomies present in single gene trees compared to the lack of polytomies in the phylogenetic hypothesis presented here (Vilstrup *et al.* 2011). The concern that mitochondrial DNA is linked and therefore only provides a single estimate of phylogeny, making introgression and hybridization undetectable, is valid. However, mitochondrial gene regions evolve at different rates and carry different phylogenetic signals, illustrated by the difference in topology among gene trees. Therefore, including the full mitogenome in this study ensures that various signals are considered in the analysis, even if they are from linked genes. Additionally, various studies have shown that the full mitogenome phylogenetic hypothesis is more likely to reflect the actual presumed species phylogeny than one nuclear gene alone (Moore 1995).

I am unable to account for potential hybridization and introgression in sturgeons, because nuclear sequences are not available for phylogenetic analysis and comparison to my phylogeny. Sturgeon are polyploid, making efforts to sequence nuclear DNA for phylogenetic studies extremely difficult (Krieger *et al.* 2006).This is a particular concern because both historic and recent hybridization is known in sturgeons (Peng 2007). To address this, future work should focus on sequencing nuclear DNA and producing sturgeon phylogenies for comparison to my mitochondrial phylogeny.

5. Concluding Remarks

Based on more data than has been previously included in any phylogenetic analysis of Acipenseridae, the full mitogenome of 22 sturgeon species (plus one subspecies), I propose a new phylogeny for the Acipenseridae. This topology, like many recent studies, recovers *Scaphirhynchus* as sister to all other sturgeons. These data support three previously described clades: *(A. sturio, (A. oxyrinchus, A. o. desotoi)); ((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii));* and (*H. dauricus*, (*A. medirostris, A. mikadoi*)). Pseudoscaphirhynchinae *sensu* Hilton *et al.* (2011) is supported, but falls within a broader clade containing additional *Acipenser* species and *H. huso.*

The monophyly of Husinae *sensu* Hilton *et al.* (2011) is not supported. Husinae, if restricted to only *H. huso*, is included in a clade with Pseudoscaphirhynchinae and a clade comprising *A. gueldenstaedtii, A. naccarii, A. ruthenus, A. baerii, A. fulvescens*, and *A. brevirostrum*; additional data are required to determine whether *A. persicus* and *A. nudiventris* (not included in this study) belong within this group.

Past molecular studies of sturgeon phylogeny have frequently employed a single or few mitochondrial genes. There is strong support for the ability of protein coding genes, individually and as a unit, to recover the same monophyletic groupings, and structure within these groupings, as the full mitogenome. My results suggest that the protein coding gene family along with the rRNA 16s may reflect a topology more similar to that of the full mitogenome than any other mitochondrial gene grouping. My results reinforce the conclusion drawn by previous authors that evolutionary rate does not account for gene performance in topology recovery, but contradict the assertion that gene length cannot be used to infer utility of a gene to recover the full mitogenome phylogenetic hypothesis. Although past studies have not found a trend between gene length and reliability in recovering phylogeny, this study did recover such a trend. I suggest that future studies consider addressing this further. The use of the full mitogenome along with other characters (nuclear genes, morphological data) may recover differing or better supported topologies.

Acknowledgements

I would like to acknowledge Dr. Masaki Miya for his tremendous mentorship, support, and input during all aspects of this research. Also, many thank to Mr. Tetsuya Sado for his assistance and kindness, and to everyone at the Natural History Museum and Institute in Chiba, Japan for welcoming me during my time there. Dr. Jan McDowell provided much appreciated assistance and advice, and allowed me space in her and Dr. John Graves' lab. I'd like to also thank Dr. John Graves for his mentorship and support, and comments on early drafts of this chapter. Thanks also to Casey Dillman for help in development of the chapter and in the lab, and for comments on early drafts. I acknowledge all of my committee members for support and input during development of this Chapter. Also, thanks to all of my Molecular Labmates for general support and inspiration- Nadya Mamoozedeh, Brendan Turley, Hamish Small, and Heidi Brightman. Funding sources included National Science Foundation (DEB 0841691), the NSF GK-12 Program, the NSF East Asia and Pacific Summer Institute Program, the Japan Society for Promotion of Science Summer Program, and the Virginia Institute of Marine Science Dean and Director's Office and Fisheries Department.

References

Ahn S., Kim T., Lee S., Kim D., Ghang H., Kim D., Kim B., Kim S., Kim W., Kim C., Park D., Lee Y.S., Kim S., Reja R., Jho S., Kim C.G., Cha J., Kim K., Lee B., Bhak J., and Kim S. 2009. The first Korean genome sequence and analysis: Full genome sequencing for a socio-ethnic group. Genome Research, 19:1622-1629.

Arnason U., Gullberg A., Janke A., Joss J. and Elmerot C. 2004. Mitogenomic analyses of deep gnathostome divergences: a fish is a fish. Gene, 333: 61-70.

Artyukhin E. N. 1995. On biogeography and relationships within the genus *Acipenser*. Sturgeon Quarterly, 3(2): 6–8.Artyukhin, E. N. 2006. Morphological Phylogeny of the Order Acipenseriformes. Journal ofApplied Ichthyology, 22: 66–69.

Avise J.C. 1998. The history and purview of phylogeography: a personal reflection. Molecular Ecology, 7(4): 371-379.

Avise J.C. 2010. Perspective: conservation genetics enters the genomics era. Conservation Genetics, 11(2):665-669.

Baack E.J. and Rieseberg L.H. 2007. A genomic view of introgression and hybrid speciation.Current Opinion in Genetics & Development, 17(6): 513-518.

Bemis W.E. and Kynard B. 1997. Sturgeon rivers: an introduction to acipenseriform biogeography and life history. Environmental Biology of Fishes, 48(1): 167-183.

Berg L. S. 1904. Zur Systematic der Acipenseriden. Zoologischer Anzeiger, 27: 665–667.

Birstein V.J. & DeSalle R. 1998. Molecular phylogeny of Acipenserinae. Molecular Phylogeneticsand Evolution, 9(1): 141-155.

Birstein V. J., Doukakis P., DeSalle R., 2002. Molecular phylogeny of Acipenseridae: Nonmonophyly of Scaphirhynchinae. Copeia, 2002(2): 287–301.

Botero-Castro F., Tilak M., Catzeflis F., Delsuc F., Douzery E.J.P. 2013. Next-generation sequencing and phylogenetic signal of complete mitochondrial genomes for resolving the evolutionary history of leaf-nosed bats (Phyllostomidae). Molecular Phylogenetics and Evolution, 69: 728-39.

Brown J.R., Beckenbach K., Beckenbach A.T., Smith M.J. 1996. Length Variation, Heteroplasmy and Sequence Divergence in the Mitochondrial DNA of Four Species of Sturgeon (*Acipenser*). Genetics, 142(2): 525–535.

Caravas J. 2012. Phylogenetic Utility of Mitochondrial and Nuclear Genes: A Case Study in theDiptera (True Flies). Wayne State University Dissertations. Available at:DigitalCOmmons@WayneState.

Chen X.W., Jiang S., Shi Z.Y., Li Q., Xun X.R. and Guo da Q. 2012. Mitochondrial genome of the Siberian sturgeon *Acipenser baerii*. Mitochondrial DNA, 23(2), 120-122.

Chen W.Y. and Lin S.M. 2013. Mitochondrial phylogenomics of an ancient ray-finned fish family Polypteridae. Unpublished, NCBI Genome Project Direct Submission.

Cheng S, Chang SY, Gravitt P, Respess R. 1994. Long PCR. Nature, 369:684–685.

Choudhury A. and Dick T.A. 1998. Special Paper: The Historical Biogeography of Sturgeons (Osteichthyes: Acipenseridae): A Synthesis of Phylogenetics, Palaeontology andPalaeogeography. Journal of Biogeography,25(4): 623-640.

DeFilippis V.R. and Moore W.S. 2000. Resolution of Phylogenetic Relationships among Recently

Evolved Species as a Function of Amount of DNA Sequence: An Empirical Study Based on Woodpeckers (Aves: Picidae). Molecular Phylogenetics and Evolution, 16(1):143-160.

Dillman C. B., Wood R. M., Kuhajda B. R., Ray J. M., Salnikov V. B., Mayden R. L. 2007. Molecular systematics of Scaphirhynchinae: an assessment of North American and Central AsianFreshwater Sturgeon Species. Journal of Applied Ichthyology, 23: 290–296.

Duchene S, Archer FI, Vilstrup J, Caballero S, Morin PA (2011) Mitogenome Phylogenetics: The Impact of Using Single Regions and Partitioning Schemes on Topology, Substitution Rateand Divergence Time Estimation. PLoS ONE, 6(11): e27138.doi:10.1371/journal.pone.0027138

Faith D.P. 1992. Systematics and Conservation: On Predicting the Featuure Diversity of Subsets of Taxa. Cladistics, 8(4): 361-373.

Feldmeyer B., Hoffmeier K., and Pfenninger M. 2010. The complete mitochondrial genome of *Radix balthica* (Pulmonata, Basommatophora), obtained by low coverage shot gun next generation sequencing. Molecular Phylogenetics and Evolution, 57(3): 1329-1333.

Findeis, E.K. 1997. Osteology and phylogenetic interrelationships of sturgeons (Acipenseridae).Environmental Biology of Fishes, 48(1-4): 73-126

Fontana F., Tagliavini J., Congui L. 2001. Sturgeon genetics and cytogenetics: Recentadvancements and perspectives. Genetica, 111: 359–373.

Forest F., Grenyer R., Rouget M., Davies T.J., Cowling R.M., Faith D.P., Balmford A., Manningg J.C., Proche S., van der Bank M., Reeves G., Hedderson T.A.J., and Savolainen V. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature, 445: 757-760.

Grande L. and Bemis W.E. 1996. Interrelationships of Acipenseriformes with comments on "Chondrostei". pages 85-115. In: M.L.J. Stiassny, L.R. Parenti & G.D. Johnson (eds.) Interrelationships of Fishes. Academic Press, New York.

Grande L., Hilton E.J. 2006. An exquisitely preserved skeleton representing a primitive sturgeonfrom the Upper Cretaceous Judith River Formation of Montana (Acipenseriformes: Acipenseridae: n. gen. and sp.). Journal of Paleontology, Memoir, 65: 1–39.

Grande L., Jin F., Yabumoto Y., and Bemis W.E. 2002. {Protopsephurus liui, a well-preserved primitive paddlefish (Acipenseriformes: Polyodontidae) from the Early Cretaceous of China. Journal of Vertebrate Paleontology, 22: 209–237.

Gredilla R., Bohr V.A., and Stevnsner T. 2010. Mitochondrial DNA repair and association with aging – An update. Experimental Geology, 45(7-8): 478-488.

Groenenberg D.S.J., Pirovano W., Gittenberger E. and Schilthuizen M. 2012. The complete mitogenome of *Cylindrus obtusus* (Helicidae, Ariantinae) using Illumina next generation sequencing. BMC Genomics, 13:114.

Guiterrez V., Rego N., Naya H., and Garcia G. 2015. First complete mitochondrial genome of the South American annual fish *Austrolebias charrua* (Cyprinodontiformes: Rivulidae): peculiar features among cyprinodontiform mitogenomes. BMC Genomics, 16(1):879.

Hardy O.J. and Pavoine S. 2012. Assessing phylogenetic signal with measurement error: Acomparison of Mantel Tests, Blomberg *et al.*'s K, and Phylogenetic Distograms. Evolution, 66(8):2614-2621.

Hilton E. J. and Grande L. 2006. Review of the fossil record of sturgeons, family Acipenseridae (Actinopterygii: Acipenseriformes), from North America. Journal of Paleontology, 80: 672-683.

Hilton E.J., Grande L., Bemis W.E. 2011. Skeletal anatomy of the shortnose sturgeon *Acipenserbrevirostrum* Lesueur, 1818, and the systematics of sturgeons (Acipenseriformes, Acipenseridae). Fieldiana Life and Earth Sciences, 3: 1-168.

Hajibabaei M., Singer G.A.C., Hebert P.D.N., Hickey D.A. 2007. DNA barcoding: how itcomplements taxonomy, molecular phylogenetics and population genetics. TRENDS in Genetics, 23(4): 167-172.

Hogg I.D., Smith B.J., Banks J.C., Dewaard J.R., Hebert P.D.N. 2009. Testing use of mitochondrial COI sequences for the identification and phylogenetic analysis of New Zealand caddisflies (Trichoptera). New Zealand Journal of Marine and Freshwater Research, 43(5): 1137-1146.

Holzkamm, T.E. and L.G. Waisberg. 2004. Native American utilization of sturgeon. In G.T.O.LeBreton, F.W.H. Beamish and R.S. McKinley (eds.). Sturgeon and Paddlefish of NorthAmerica. Kluwer Academic Publishers, Dordreht. Fish and Fisheries Series, 27: 22–39.

Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny.Bioinformatics, 17:754-755.

ICES: International Council for the Exploration of the Sea. Resolution Conf. 12.7 (Rev. CoP16).Conservation of and trade in sturgeons and paddlefish. Available at: <http://cites.orhttpg/eng/res/12/12-07R16.php>

illumina. 2012. Nextera ® XT DNA Sample Preparation Guide. ILLUMINA PROPRIETARY Part #15031942 Rev. C. Available at: <http://www.liai.org/files/nextera_xt_sample_preparation_guide_15031942_c.pdf>.

illumina. 2015. Nextera XT DNA Library Preparation Guide. San Diegom California.

Inoue J.G., Miya M., Tsukamoto K. & Nishida M. 2003. Basal actinopterygian relationships: amitogenomic perspective on the phylogeny of the "ancient fish". MolecularPhylogenetics and Evolution, 26: 110–120.

Inoue J.G., Miya M., Lam K., Tay B., Danks J.A., Bell J., Walker T.I., and Venkatesh B. 2010. Evolutionary Origin and Phylogeny of the Modern Holocephalans (Chondrichthyes: Chimaeriformes): A Mitogenomic Perspective. Molecular Biology and Evolution, 27(11): 2576-2586.

Iorizzo M., Senalik D., Szklarczyk M., Grzebelus D., Spooner D., and Simon P. 2012. BMC Plant Biology, 12: 61.

Isaac N.J.B., Turvey S.T., Collen B., Waterman C. Bailllie J.E.M. 2007. Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. PLOS One, 2(3): e296. doi:10.1371/journal.pone.0000296

Ishiguro N.B., Miya M., and Nishida M. 2003. Basal euteleostean relationships: a mitogenomic perspective on the phylogenetic reality of the "Protacanthopterygii". Molecular Phylogenetics and Evolution, 27(3): 476-488.

(IUCN) The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>.Downloaded on 20 May 2015.

Iwasaki W., Fukunaga T., Isagozawa R., Yamada K., Maeda Y., Satoh T.P., Sado T., Mabuchi K.,Takeshima H., Miya M., Nishida M. 2013. MitoFish and MitoAnnotator: a mitochondrialgenome database of fish with an accurate and automatic annotation. PipelineMolecular Biology and Evolution, 30:2531–2540.

Jacobsen M .W., Hansen M.M., Orlando L., Bekkevold D., Bernatchez L., Willerslev E., GilbertM.T.P. 2012. Mitogenome sequencing reveals shallow evolutionary histories and recentdivergence.time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.). Molecular Ecology, 21:2727–2742.

Jex A.R., Littlewood D.T.J., and Gasser R.B. 2010. Toward next-generation sequencing of mitochondrial genomes — Focus on parasitic worms of animals and biotechnological implications. Biotechnology Advances, 28(1): 151-159.

Krieger J., Fuerst P.A., Cavender T.M. 2000. Phylogenetic relationships of the North Americansturgeons (Order Acipenseriformes) based on mitochondrial DNA sequences. MolecularPhylogenetics and Evolution, 16(1): 64-72.

Krieger J., Hett A.K., Fuerst P.A., Birstein V.J., and Ludwig A. 2006. Unusual intraindividual variation of the nuclear 18srRNA gene is widespread within the Acipenseridae. Journal of Heredity, 96(3): 218-225.

Krieger J., Hett A. K., Fuerst P. A., Artyukhin E., Ludwig A. 2008. The molecular phylogeny of theorder Acipenseriformes revisited. Journal of Applied Ichthyology, 24(supplement 1): 36–45.

Lavoué S., Miya M., Saitoh K., Ishiguro N.B., and Nishida M. 2007. Phylogenetic relationships among anchovies, sardines, herrings and their relatives (Clupeiformes), inferred from whole mitogenome sequences. Molecular Phylogenetics and Evolution, 43(3): 1096-1105.

Lefcheck J.S., Buchheister A., Laumann K.M., Stratton M.A., Sobocinski K.L., Chak S.T.C., Clardy T.R., Reynolds P.L., Latour R.J., Duffy J.E.. 2014. Dimensions of biodiversity in Chesapeake Bay demersal fishes: patterns and drivers through space and time. Ecosphere, 5(2): 1-48.

Li C., Cheng L., Li J.T., Lu C.Y., Wang Y. and Sun X.W. 2013. Complete mitochondrial genome of sterlet (*Acipenser ruthenus*). Mitochondrial DNA, 26(2): 259-260.

Lindqvist C., Schuster S.C., Sun Y., Talbot S.L., Qi J., Ratan A., Tomsho L.P., Kasson L., Zeyl E., Aars J., Miller W., Ingólfsson O., Bachmann L., and Wiig O. 2010. Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. PNAS, 107(11): 5053–5057.

López-Villavicencio M., Aguileta G., Giraud T., de Vienne D.M., Lacoste S., Couloux A., Dupont J.2010. Sex in Penicillium: Combined phylogenetic and experimental approaches. Fungal Genetics and Biology, 47: 693–706.

Ludwig A., Belfiore N. M., Pitra C., Svirsky V., Jenneckens II I. 2001. Genome duplication eventsand functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*. Genetics, 158: 1203–1215.

Mardis E.R. 2008. Next-Generation DNA Sequencing Methods. Annual Review of Genomics and Human Genetics, 9:387-402.

Mayden R. L., Kuhajda B. R. 1996. Systematics, taxonomy, and conservation status of theendangered Alabama sturgeon, *Scaphirhynchus suttkusi* Williams and Clemmer (Actinopterygii, Acipenseridae). Copeia, 1996(2): 241–273.

Meiklejohn K.A., Wallman J.F., and Dowton M. 2013. DNA Barcoding Identifies all Immature Life Stages of a Forensically Important Flesh Fly (Diptera: Sarcophagidae). Journal of Forensic Sciences, 58(1):184-187.Miya M., Kawaguchi A., and Nishida M. 2001. Mitogenomic Exploration of Higher Teleostean Phylogenies: A Case Study for Moderate-Scale Evolutionary Genomics with 38 Newly Determined Complete Mitochondrial DNA Sequences. Molecular Biology and Evolution, 18(11): 1993-2009.

Miller M.A., Pfeiffer W., and Schwartz T. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses. Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery, Article No. 41.

Miya M. and Nishida M. 1996. Molecular phylogenetic perspective on the evolution of the deepsea fish genus Cyclothone (Stomiiformes: Gonostomatidae). Ichthyological Research, 43:375–398.

Miya M. and Nishida. M. 1999. Organization of the Mitochondrial Genome of a Deep-Sea Fish, *Gonostoma gracile* (Teleostei: Stomiiformes): First Example of Transfer RNA GeneRearrangements in Bony Fishes. Marine Biotechnology, 1:416-426.

Miya M. and Nishida M. 2000. Use of mitogenomic information in teleostean molecularphylogenetics: a tree-based exploration under the maximum-parsimony optimalitycriterion. Molecular Phylogenetics and Evolution, 17(3): 437-455.

Miya M. and Nishida M. 2015. The mitogenomic contributions to molecular phylogenetics andevolution of fishes: a 15-year retrospect. Ichthyological Research, 62:29-71.

Moore W. 1995. Inferring Phylogenies from mtDNA Variation: Mitochondrial-Gene Trees VersusNuclear-Gene Trees. Evolution, 49(4): 718-716.

Moritz C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Biology, 3(4): 401-411.

Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in Rlanguage. Bioinformatics, 20: 289-290.

Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in Rlanguage. Bioinformatics, 20: 289-290.

Peng Z., Ludwig A., Wang D., Diogo R., Wei Q. and He S. 2007. Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes). Mol. Phylogenet. Evol. 42(3), 854-862.

Penny D. and Hendy M.D. 1985. The Use of Tree Comparison Metrics. Systematic Zoology, 34(1):75-82.

Posada D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution, 25(7): 1253-1256.

Ptaszyńska A.A., Łętowski J., Gnat S., and Małek W. 2012. Application of COI sequences in studies of phylogenetic relationships among 40 Apionidae species. Journal of Insect Science, 12(1): Article 16.

Qiagen. 2006. DNeasy® Blood & Tissue Handbook. In DNeasy Blood & Tissue Kit. Available at: <https://www.qiagen.com/us/resources/resourcedetail?id=6b09dfb8-6319-464d-996c-79e8c7045a50&lang=en>.

QIAGEN. 2016. Manual for CLC Genomics Workbench 8.5.1 Windows, Mac OS X and Linux.

Rastorguev S., Mugue N., Volkov A., Barmintsev V. 2008. Complete mitochondrial DNA sequenceanalysis of Ponto-Caspian sturgeon species. Journal of Applied Ichthyology, 24(1): 4649.

Rastogi P.A. 1999. MacVector. Bioinformatics Methods and Protocols, 132: 47-69.

R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation forStatistical Computing.

Rehbein H., Gonzales-Sotelo C., Perez-Martin R., Quinteiro J., Rey-Mendez M., Pryde S., Mackie M., and Santos A.T. 1999. Differentiation of sturgeon caviar by single strand conformation polymorphism (PCRSSCP) analysis. Archrv fur Lebensmittelhygiene, 50: 1-24.

Robinson D.F. and Foulds L.R. 1981. Comparison of Phylogenetic Trees. Mathematical Biosciences, 53: 131-141.

Rodrigues R.R., Gandolfi S., Nave A.G., Aronson J., Barretto T.E., Vidal C.Y., and Brancalion P.H.S. 2011. Large-scale ecological restoration of high-diversity tropical forests in SE Brazil. Forest Ecology and Management, 261: 1605-1613.

Rodrigues A.S.L. and Gaston K.J. 2002. Maximising phylogenetic diversity in the selection of networks of conservation areas. Biological Conservation, 105(1): 103-111.

Roos J., Aggarwal R.K., and Janke A. 2007. Extended mitogenomic phylogenetic analyses yield new insight into crocodylian evolution and their survival of the Cretaceous–Tertiary boundary. Molecular Phylogenetics and Evolution, 45(2): 663-673.

Ruban G.I., Kholodova M.V., Kalmykov V.A. and Sorokin P.A. 2011. A review of the taxonomic status of the Persian sturgeon (*Acipenser persicus* Borodin). Journal of Applied Ichthyology, 27: 470-476.

Rzhetsky A. and Nei M. 1992. A simple method for estimating and testing minimum-evolutiontrees. Molecular Biology and Evolution, 9: 945-967.

Shen K., Chen C., Hsiao C., and Durand J. 2015. Next-generation sequencing yields the complete mitochondrial genome of the flathead mullet, *Mugil cephalus* cryptic species in East Australia (Teleostei: Mugilidae). Mitochondrial DNA: The Journal of DNA Mapping, Sequencing, and Analysis, 9:1-2.

Simon C., Fratie F., Bechenbach A., Crespi B., Liu H., Flook P. 1994. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. Annals of the Entomological Society of America, 87(6): 651-701.

Simons A., Wood R. M., Heath L. S., Kuhajda B. R., Mayden R. L. 2001. Phylogenetics of*Scaphirhynchus* based on mitochondrial DNA sequences. Transactions of the AmericanFisheries Society, 130: 359–366.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models", Bioinformatics, 22(21):2688–2690.

Sloan D.B., Oxelman B., Rautenberg A., and Taylor D.R. 2009. Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe *Sileneae*. BMC Evolutionary Biology, 9: 260. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2777880/≥

Song S., Pursell Z.F., Copeland W.C., Longley M.J., Kunkel T.A., Mathews C.K. 2005. DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. Proceedings of the National Academy of Sciences of the United States of America 102:4990-4995.

Varshney R.K., Nayak S.N., May G.D., and Jackson S.A. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology, 27(9): 522-530.

Vasil'eva E.D., Vasil'ev V.P., Shedko S.V., Novomodny G.V. 2009. The Revision of the Validity ofGenus *Huso* (Acipenseridae) Based on Recent Morphological and Genetic Data withParticular Reference to the Kaluga *H. dauricus**. Journal of Ichthyology, 49(10):861-867.

VilstrupJ.T., Ho S.Y.W., Foote A.D., Morin P.A., Kreb D., KrützenM, Parra G.J., Robertson K.M., de Stephanis R., Verborgh P., Willerslev E., Orlando L., and Gilbert M.T.P. 2011. Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae. BMC Evolutionary Biology, 11(65): 1-10. Available at: <http://www.biomedcentral.com/1471-2148/11/65>

Ward R.D. 2009. DNA barcode divergence among species and genera of birds and fishes. Molecular Ecology Resources, 9(4): 1077-1085.

Webb K.M. and Rosenthal B.M. 2011. Next-generation sequencing of the Trichinella murrelli mitochondrial genome allows comprehensive comparison of its divergence from the principal agent of human trichinellosis, *Trichinella spiralis*. Infection, Genetics and Evolution, 11(1): 116-123.

Wilimovsky N.J. 1956. *Protoscaphirhynchus squamosus*, a new sturgeon from the UpperCretaceous of Montana. Journal of Paleontology, 30: 1205–1208.

Wolf C., Burgener M., Luthy J. 2000. PCR-RFLP Analysis of Mitochondrial DNA: Differentiation ofFish Species. Journal of the Swiss Society of Food Science and Technology, 33: 144-150.

Wolf C., Rentsch J. Hubner P. 1999. PCR - RFLP Analysis of Mitochondrial DNA: A ReliableMethod for Species Identification. Journal of Agricultural and Food Chemistry, 47: 13501355.

Xia X. 1998. The Rate Heterogeneity of Nonsynonymous Substitutions in Mammalian Mitochondrial Genes. Molecular Biology and Evolution, 15(3): 336-344.

Xia X. and Lemey P. 2009. Assessing substitution saturation with DAMBE. In P. Lemey, M. Salemi, and A-M. Vandamme. (eds.). The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press: 611-626.

Zardoya R.; Meyer A. 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. 13, 933–942.

Zhang D. and Hewitt G.M. 1997. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochemical Systematics and Ecology, 25(2): 99-120.

Zhang P., Liang D., Mao R., Hillis D.M., Wake D.B. and Cannatella D.C. 2013. Efficient Sequencing of Anuran mtDNAs and a Mitogenomic Exploration of the Phylogeny and Evolution of Frogs. Molecular Biology and Evolution, 30(8): 1899-1915.

Zhang P. and Wake M.H. 2009. A mitogenomic perspective on the phylogeny and biogeography of living caecilians (Amphibia: Gymnophiona). Molecular Phylogenetics and Evolution, 53(2): 479-491.

Zhang S., Zhang Y., Zheng X., Chen Y., Deng H., Wang D., Wei Q., Zhang Y., Nei L., Wu Q. 2000.Molecular phylogenetic systematics of twelve species of Acipenseriformes based on mtDNA NADH4L-ND4 gene sequence analysis. Science in China Series C Life Sciences,43: 129–137.Zhang P., Zhou H., Chen Y., Liu Y. and Qu L. 2005. Mitogenomic Perspectives on the Origin and Phylogeny of Living Amphibians. Systematic Biology, 53(3): 391-400.

Tables

Table 1. Tissue samples used in this study. GenBank accession numbers of the tissue specimens sequenced are listed (or are pending); collection numbers for voucher specimens are provided (in some cases these are pending cataloging in a home institution).

Species	Source	Voucher	Accession Number
Acipenser baerii (full mitogenome)	GenBank •	n/a	JQ045341.1
Acipenser baerii (16s rRNA, spot checking)	novel sequence	FMNH 117783	pending
Acipenser dabryanus	GenBank •	n/a	AY510085.1
Acipenser ruthenus	GenBank •	n/a	NC_022453.1
Acipenser stellatus (full mitogenome)	GenBank •	VIMS 13552	NC_005795.1
Acipenser stellatus (16s rRNA, spot checking)	novel sequence	pending	pending
Acipenser oxyrinchus desotoi	novel sequence	VIMS uncataloged	pending
Acipenser brevirostrum	novel sequence	VIMS 12073	pending
Acipenser fulvescens	novel sequence	VIMS uncataloged	pending
Acipenser gueldenstaedtii	novel sequence	VIMS uncataloged	pending
Acipenser medirostris	novel sequence	VIMS 17715	pending
Acipenser mikadoi	novel sequence	n/a*	pending
Acipenser naccarii	novel sequence	n/a**	pending
Acipenser oxyrinchus oxyrinchus	novel sequence	VIMS 12207	pending
Acipenser ruthenus	novel sequence	n/a***	pending
Acipenser schrenckii	novel sequence	VIMS uncataloged	pending
Acipenser sinensis	novel sequence	VIMS 13834	pending
Acipenser sturio	novel sequence	VIMS 13756	pending
Acipenser transmontanus	novel sequence	FMNH 117779	pending
Huso dauricus	novel sequence	NSMT P 91146	pending
Huso huso	novel sequence	n/a***	pending
Pseudoscaphirhynchus kaufmanni	novel sequence	PKN-9	pending
Scaphirhynchus albus	novel sequence	VIMS 12193	pending
Scaphirhynchus platorynchus	novel sequence	VIMS 13515	pending
Scaphirhynchus suttkusi	novel sequence	SUT-3	pending
Polyodon spathula	GenBank •	n/a	NC_004419.1
Polypterus delhezi	GenBank •	NTNU LSP136	NC_020652
Polypterus endlicherii	GenBank •	NTNU LSP117	NC_020791
Psephurus gladius	GenBank •	n/a	AY571339.1
Polyodon spathula	novel sequence	n/a	pending

* tissue sent courtesy of Shinji Adachi at Hokkaido University.

** tissue sent courtesy of Dr. Leonardo Congiu, University of Padova.

*** tissue sent courtesy of Dr. Radu Suciu, Danube Delta National Institute, Tulcea, Romania; photo available upon request.

Table 2. Gene Families, Lengths, and Best Models. Mitochondrial genes are listed, with the gene family to which each belongs and the length of each in base pairs. The best evolutionary model for each gene as determined by jModelTest is listed with the abbreviation for that model and the model name spelled out. These evolutionary models were used in all Bayesian analyses.

Gene	Gene Family	Length (in base pairs)	Best Model	
D-loop	control region	348	GTR+G	General time reversible
ATP6	protein coding	674	GTR+I+G	General time reversible
ATP8	protein coding	169	HKY+I+G	General time reversible
COI	protein coding	1422	GTR+I+G	General time reversible
COII	protein coding	691	HKY+I+G	General time reversible
COIII	protein coding	785	GTR+G	General time reversible
Cytb	protein coding	562	GTR+I+G	General time reversible
ND1	protein coding	750	GTR+G	General time reversible
ND2	protein coding	756	GTR+I+G	General time reversible
ND3	protein coding	348	GTR+I+G	General time reversible
ND4	protein coding	1374	GTR+G	General time reversible
ND4L	protein coding	297	GTR+G	General time reversible
ND5	protein coding	1383	GTR+I+G	General time reversible
ND6	protein coding	403	HKY+G	Hasegawa-Kishono-Yano
12s	rRNA	960	GTR+I+G	Hasegawa-Kishono-Yano
16s	rRNA	1716	GTR+I+G	Hasegawa-Kishono-Yano
Ala	tRNA	74	HKY+I	Hasegawa-Kishono-Yano
Arg	tRNA	70	K80	Hasegawa-Kishono-Yano
Asn	tRNA	75	HKY+G	Hasegawa-Kishono-Yano
Asp	tRNA	87	HKY+G	Hasegawa-Kishono-Yano
Cys	tRNA	67	K80+G	Hasegawa-Kishono-Yano
Gln	tRNA	69	HKY+G	Hasegawa-Kishono-Yano
Glu	tRNA	165	HKY+I	Hasegawa-Kishono-Yano
Gly	tRNA	73	HKY	Hasegawa-Kishono-Yano
His	tRNA	69	HKY	Hasegawa-Kishono-Yano
Ile	tRNA	72	K80+G	Jukes-Cantor
Leu	tRNA	75	K80+G	Kimura
Leu2	tRNA	75	HKY+I	Kimura
Lys	tRNA	75	K80+I	Kimura
Met	tRNA	70	JC	Kimura
Phe	tRNA	71	K80	Kimura
Pro	tRNA	71	HKY+G	Kimura
Ser1	tRNA	78	K80+G	Kimura
Ser2	tRNA	70	K80+G	Kimura
Thr	tRNA	72	K80	Kimura
Trp	tRNA	73	K80	Kimura
Tyr	tRNA	71	K80+G	Kimura
Val	tRNA	71	K80	Kimura

Table 3. Rzhetsky-Nei Distances: Bayesian Mitogenome Topology and Gene-Family Topologies. Rzhetsky-Nei Distances, signifying the level of similarity to that of the Bayesian full mitogenome topology, are listed for each gene-family tree. A distance of 0 signals identical topologies.

Gene Family	**Rzhetsky-Nei distance**
Protein-coding	1
tRNA	10
rRNA	12

Table 4. Full Mitogenome Hypothesis Clade Constituents, and Presence in Gene-Family Trees. Names assigned to clades and subclades recovered in the full mitogenome Bayesian phylogeny are listed, with the constituents of each clade and their presence (indicated by a check mark) or absence in the three gene-family topologies.

Full Mitogenome Clade numbers	**Clade Constituents**	**Present in Gene-Family Topologies?**		
		tRNA	**rRNA**	**protein-coding**
Clade 1	*(H. huso(*Subclade 1.1, Subclade 1.2*))*		√	√
Subclade 1.1	*(A. stellatus, P. kaufmanni)*			√
Subclade 1.2	*(A. ruthenus, (A. fulvescens, (A. brevirostrum, (A. baerii, (A. naccarii, A. gueldenstaedtii)))))*			√
Clade 2	*(H. dauricus, (A. medirostris, A. mikadoi))*	√	√	√
Clade 3	*((A. dabryanus, A. sinensis), (A. transmontanus, A. schrenckii))*	√	√	√
Clade 4	*(A. sturio, (A. oxyrinchus, A. o. desotoi))*	√	√	√
Clade 5	*(S. suttkusi, (S. platorynchus, S. albus))*	√	√	√

Table 5. Rzhetsky-Nei Distances between Bayesian Mitogenome Topology and Gene Trees. Rzhetsky-Nei Distances, signifying the level of similarity to that of the Bayesian full mitogenome topology, are listed for each individual gene phylogeny. A distance of 0 signals identical topologies. The most similar 25% and 50% of genes to the full mitogenome are lindicated.

Gene	Rzhetsky-Nei distance	Gene length	Top-scoring	
ATP6	2	674	top-scoring 25%	top-scoring 50%
ND5	6	1383		
16s	9	1716		
CO3	9	785		
ND1	9	750		
ND4L	9	297		
ND2	10	756		
ND4	10	1374		
ND3	12	348		
ATP8	13	169		
CO2	13	691		
CO1	14	1422		
cytb	14	562		
ND6	14	403		
Asp	15	87		
Cys	16	67		
Phe	16	71		
Ser1	16	78		
His	17	69		
12s	18	960		
Ala	18	74		
Gln	18	69		
Lys	18	75		
Pro	18	71		
Thr	18	72		
Val	18	71		
Asn	19	75		
Glu	19	165		
Gly	19	73		
Leu2	19	75		
Met	19	70		
Ser2	19	70		
Trp	19	73		
Arg	20	70		
Ile	20	72		
Leu1	20	75		
Tyr	21	71		
D-loop	28	348		

Table 6. Birstein et al. 2002 Parsimony Clade Constituents. Names assigned to polytomies and clades recovered in Birstein et al.'s (2002) parsimony phylogeny are listed, along with the constituents of each polytomy or clade.

Polytomy Name (or sister species designation)	**Clade Name (or sister species position)**	**Polytomy/Clade Constituents**
A	A1	*(A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. persicus, A. naccarii))))*
	A2	*(A. stellatus, (P. kaufmanni, P. hermanni))*
	A3	*(H. dauricus, (A. ruthenus, H. huso)*
	A4	*A. fulvescens*
	A5	*A. nudiventris*
sister taxon	sister to Polytomy A	*A. mikadoi*
B	Polytomy A	see above
	B1	*(A. schrenckii, (A. transmontanus, A. medirostris))*
	B2	*A. sinensis*
C	Polytomy B	see above
	C1	*(A. sturio, A. oxyrinchus)*
	C2	*(S. platorynchus, S. albus,* and *S. suttkusi)*

Table 7. Birstein et al. 2002 Maximum Likelihood Clade Constituents. Names assigned to clades recovered in Birstein et al.'s (2002) maximum likelihood phylogeny are listed, along with the constituents of each polytomy or clade. Species not belonging to any of the named clades (ML1-6Z) are listed according to their relationship with named clades.

Clade Name (or sister species position)	**Clade Constituents**
ML1	*(A. fulvescens, (A. brevirostrum, (A. baerii, (A. gueldenstaedtii, (A. naccarii, A. persicus))))))*
ML2	*(A. stellatus, (P. kaufmanni, P. hermanni))*
sister to Clades ML1 and 2	*A. nudiventris*
ML3	*(H. dauricus, (H. huso, A. ruthenus))*
ML4	*(S. platorynchus, S. albus,* and *S. suttkusi)*
ML5	*(A. medirostris, (A. sinensis, (A. transmontanus, A. schrenckii)))*
ML6	*(A. oxyrinchus, A. sturio)*

Table 8. Birstein et al. 2002 Combined Evidence Clade Constituents. Names assigned to clades recovered in Birstein et al.'s (2002) combined evidence (molecular and morphological) phylogeny are listed, along with the constituents of each polytomy or clade. Species not belonging to any of the named clades (CE1-5) are listed according to their relationship with named clades.

Clade Name (or sister species position)	**Clade Constituents**
CE1	*(A. baerii, A. brevirostrum)*
CE2	*(A. stellatus, (P. hermanni, P. kaufmanni)*
CE3	*(H. dauricus, (A. ruthenus, H. huso))*
sister to Clades CE1, 2, and 3	*Acipenser fulvescens*
CE4	*(A. sinensis, A. transmontanus)*
sister to Clades CE1-CE4	*Acipenser oxyrinchus*
CE5	*(S. suttkusi, (S. albus, S. platorynchus))*

Figures

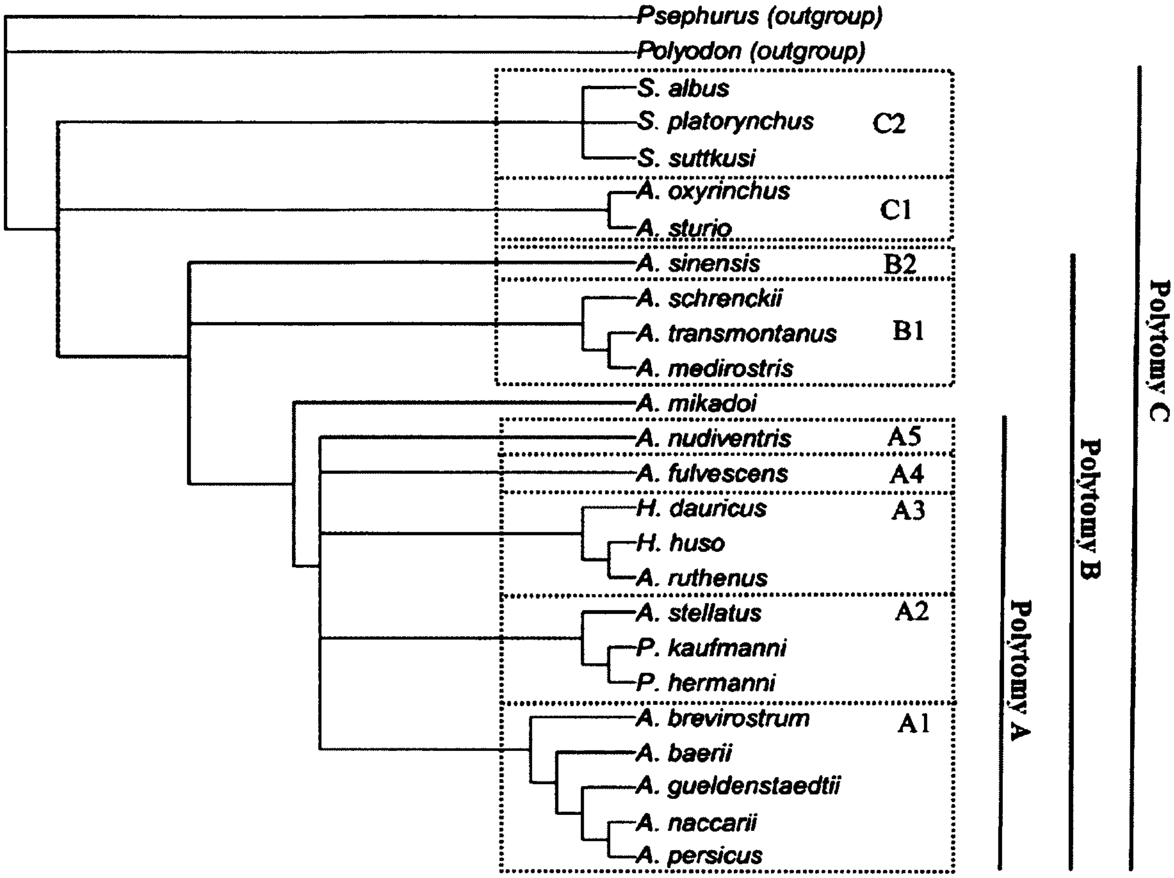


Figure 1: Parsimony tree based on 12s, 16s, cyt*b*, D-loop, and NADH 5; modified from Birstein *et al.* (2002). Clades and clade numbers are in boxes with dashed lines; polytomy constituents and polytomy names are to the right of the phylogeny. Clades and polytomies are listed in Table 6.

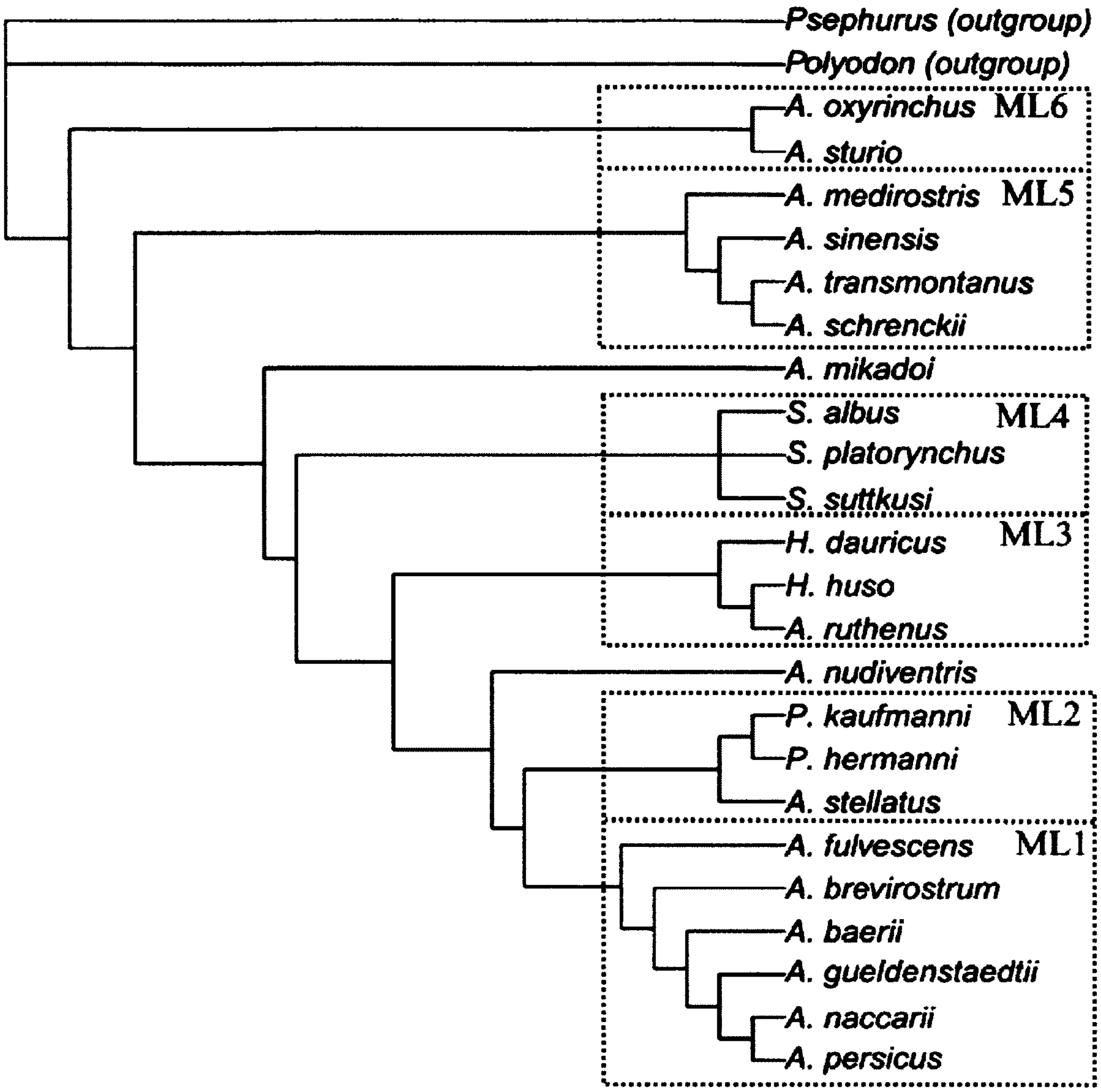


Figure 2: Maximum Likelihood tree based on 12s, 16s, *cytb*, D-loop, and NADH 5; modified from Birstein *et al.* 2002. Clades and clade numbers are in boxes with dashed lines; clades are listed in Table 7.

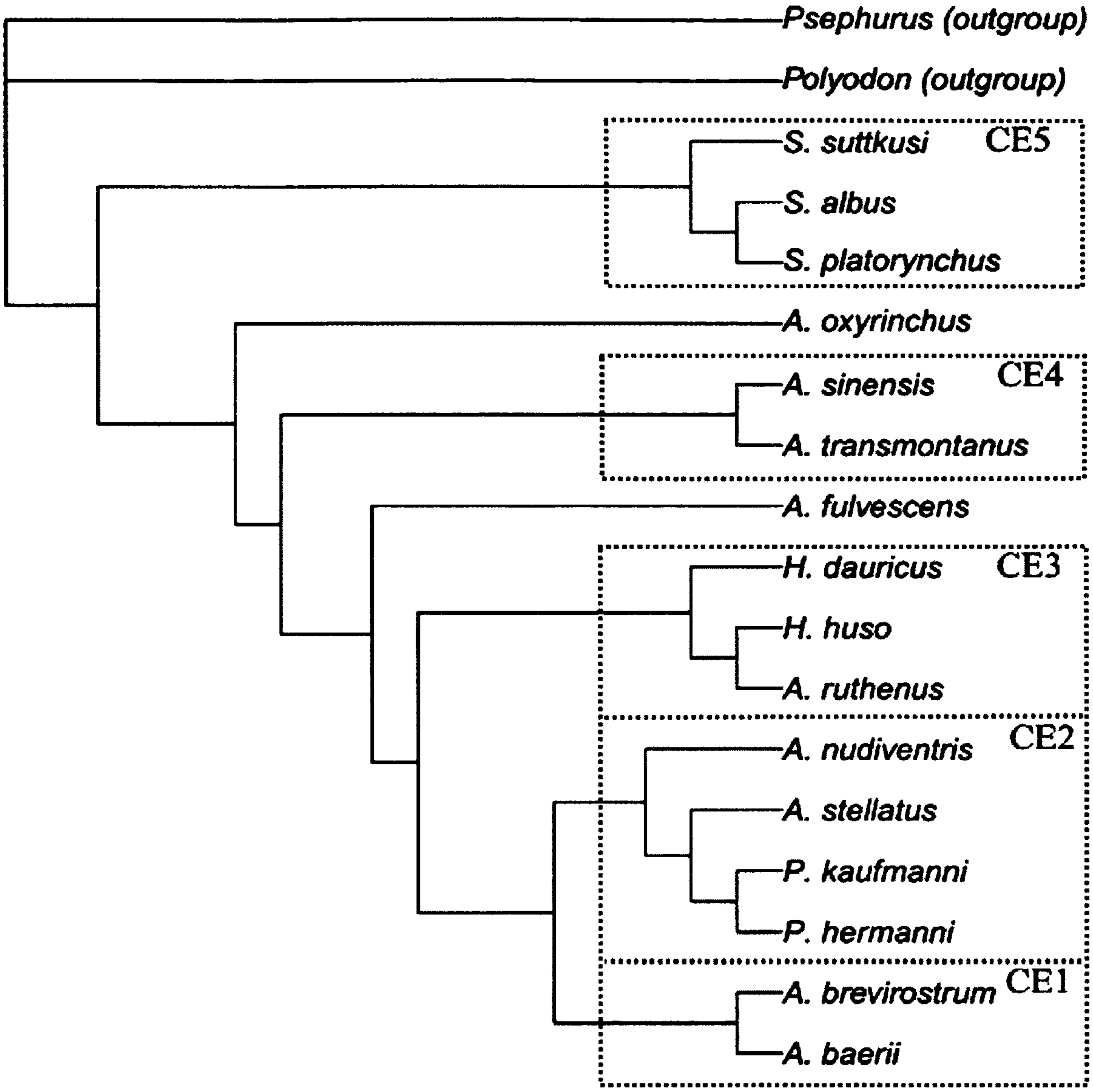


Figure 3: Combined Evidence tree based on 12s, 16s, cyt*b*, D-loop, NADH 5, and morphological data; modified from Birstein *et al.* (2002). Clades and clade numbers are in boxes with dashed lines; clades are listed in Table 8.

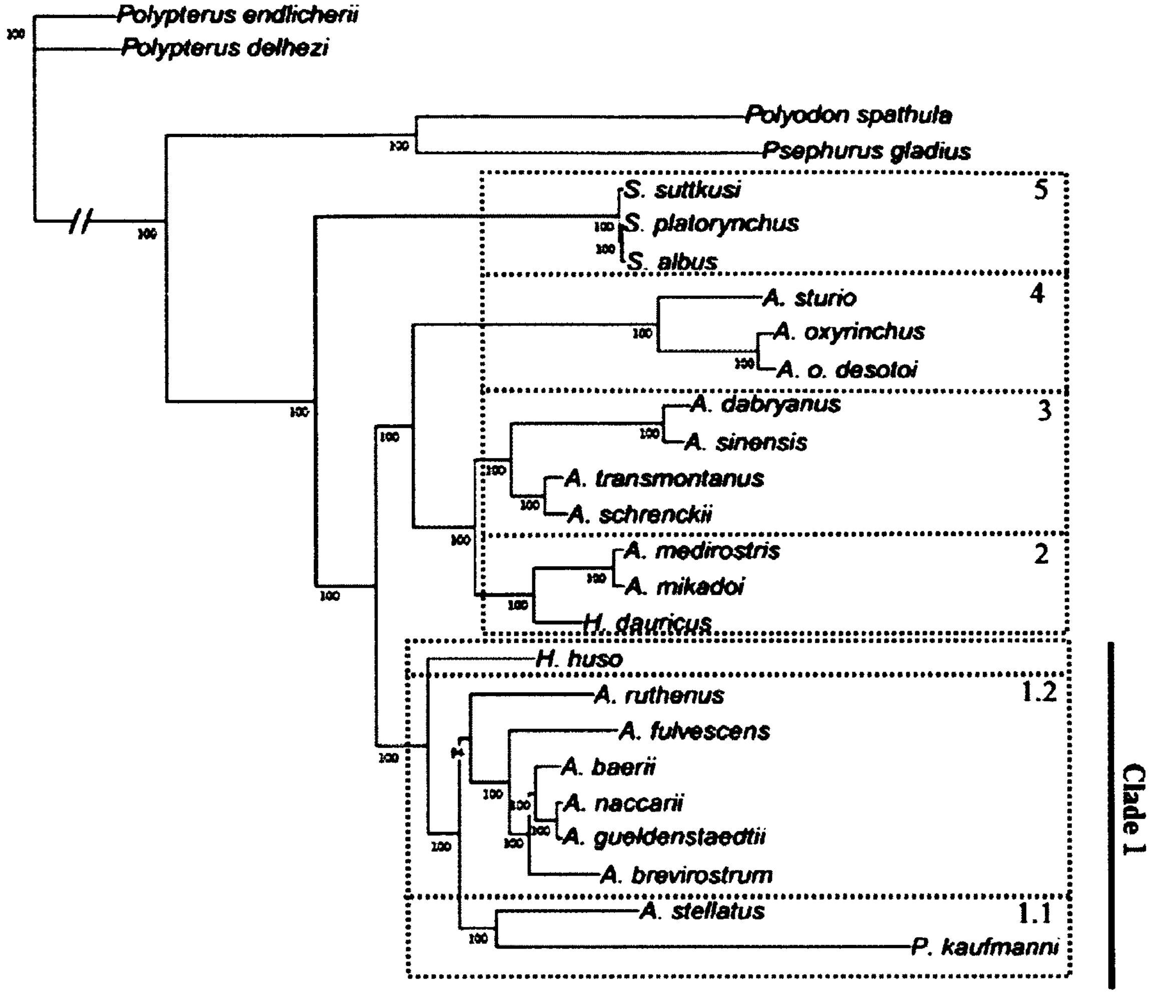


Figure 4: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades and clade numbers are in boxes with dashed lines; Clade 1, which is composed of units 1.1, 1.2, and *H. huso,* is labeled to the right of the phylogeny. Clade constituents are listed in Table 4.

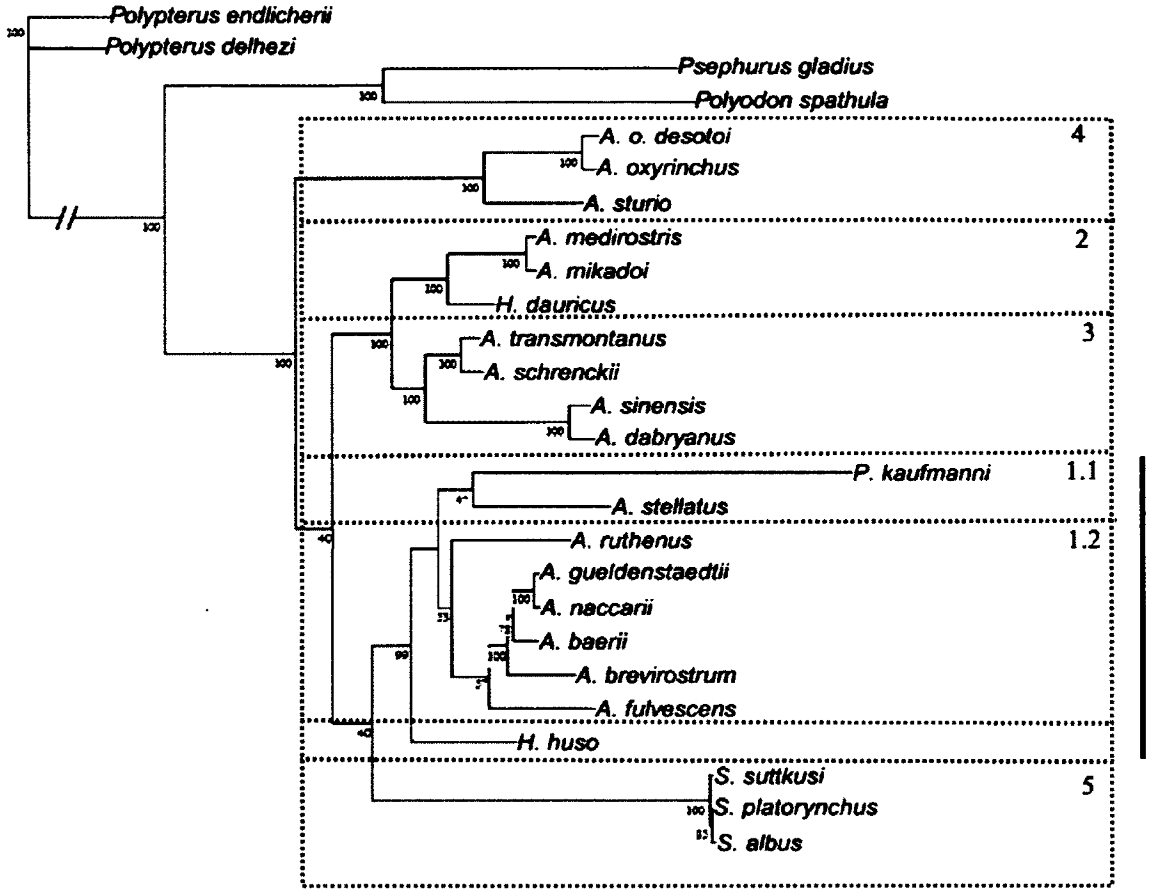


Figure 5: Maximum Likelihood Full mitogenome phylogeny. Bootstrap values are below and to the left of each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades and clade numbers are in boxes with dashed lines; Clade 1, which is composed of units 1.1, 1.2, and *H. huso,* is labeled to the right of the phylogeny. Clade constituents are listed in Table 4.

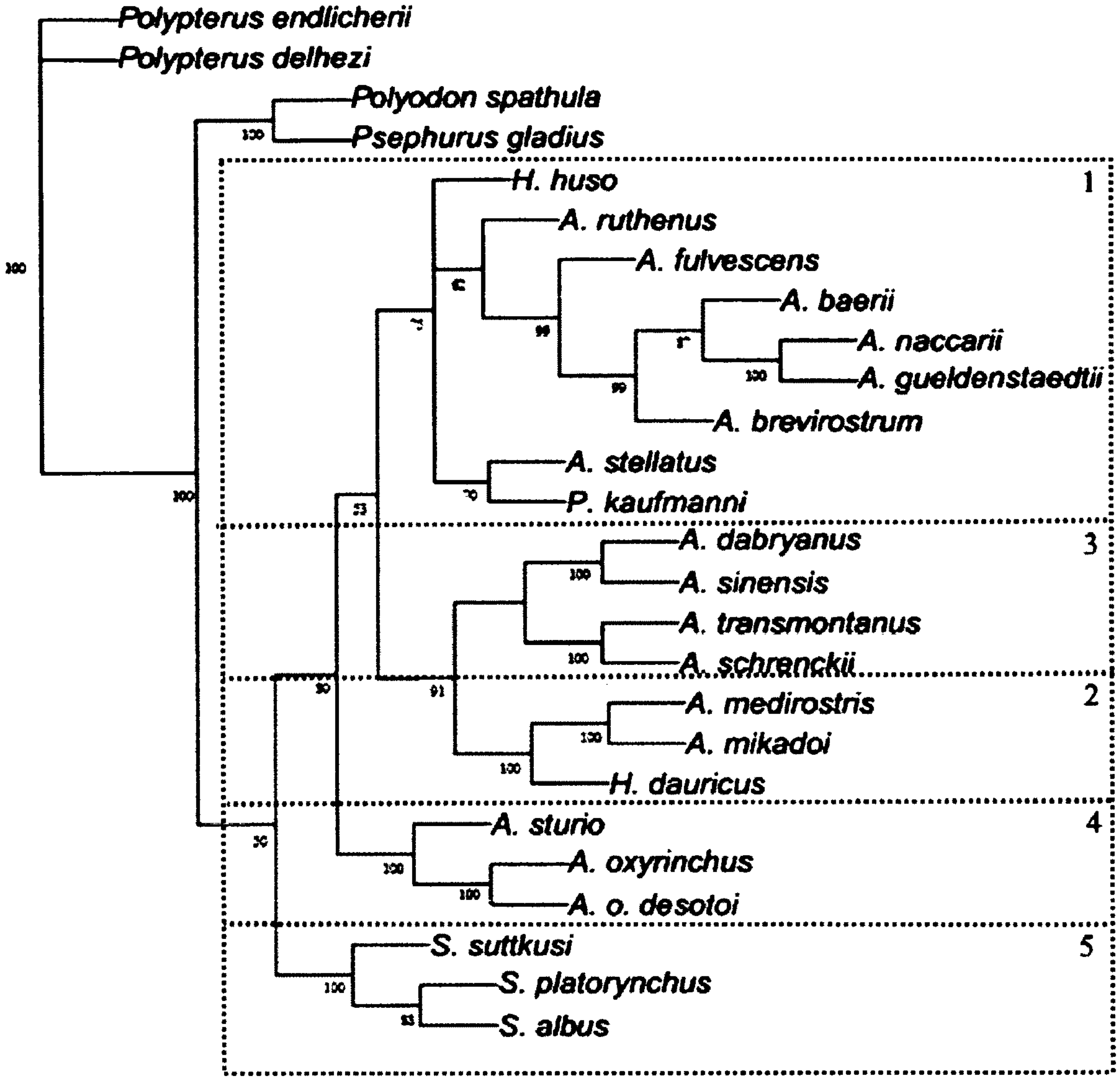


Figure 6: Parsimony full mitogenome phylogeny. Bootstrap values are below and to the left of each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Clades and clade numbers are in boxes with dashed lines; clade constituents are listed in Table 4.

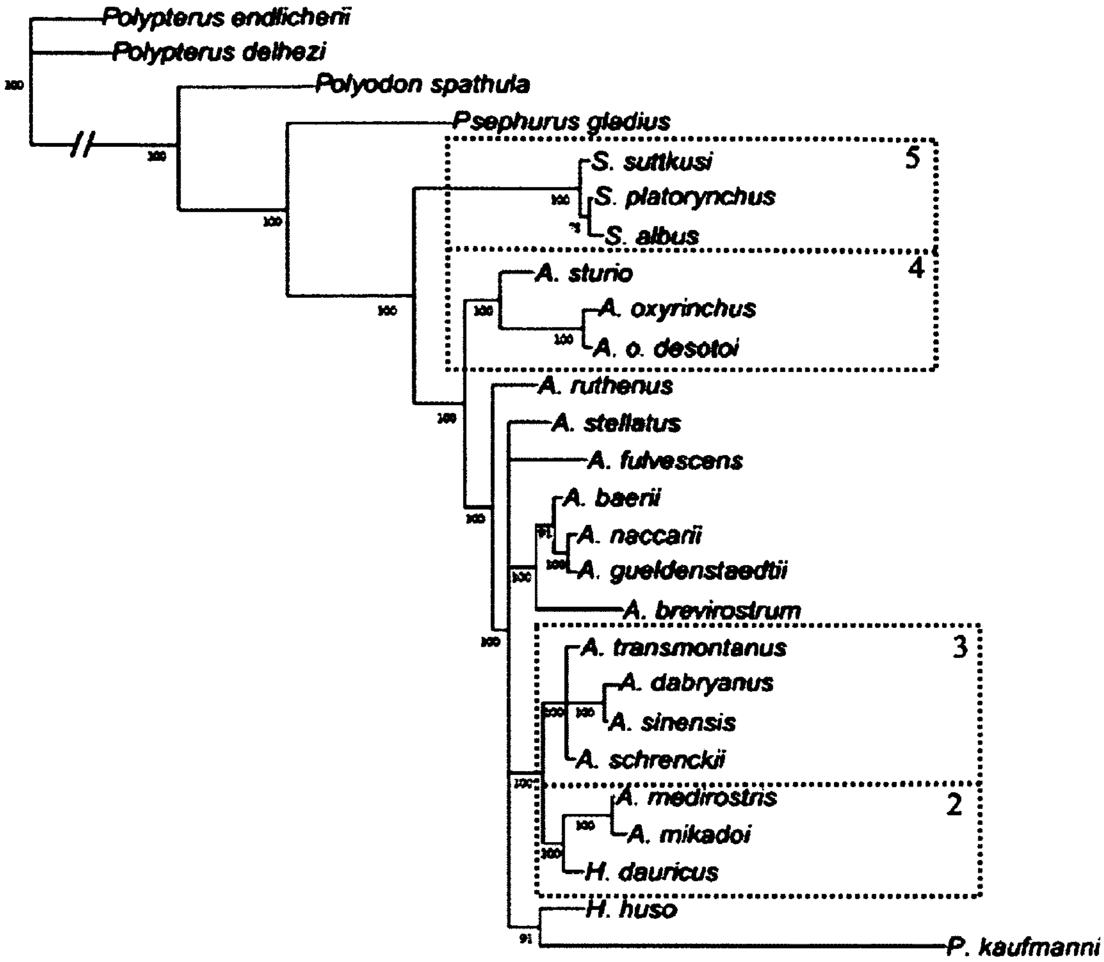


Figure 7: Bayesian phylogenetic hypothesis based on concatenated tRNA sequences. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4.

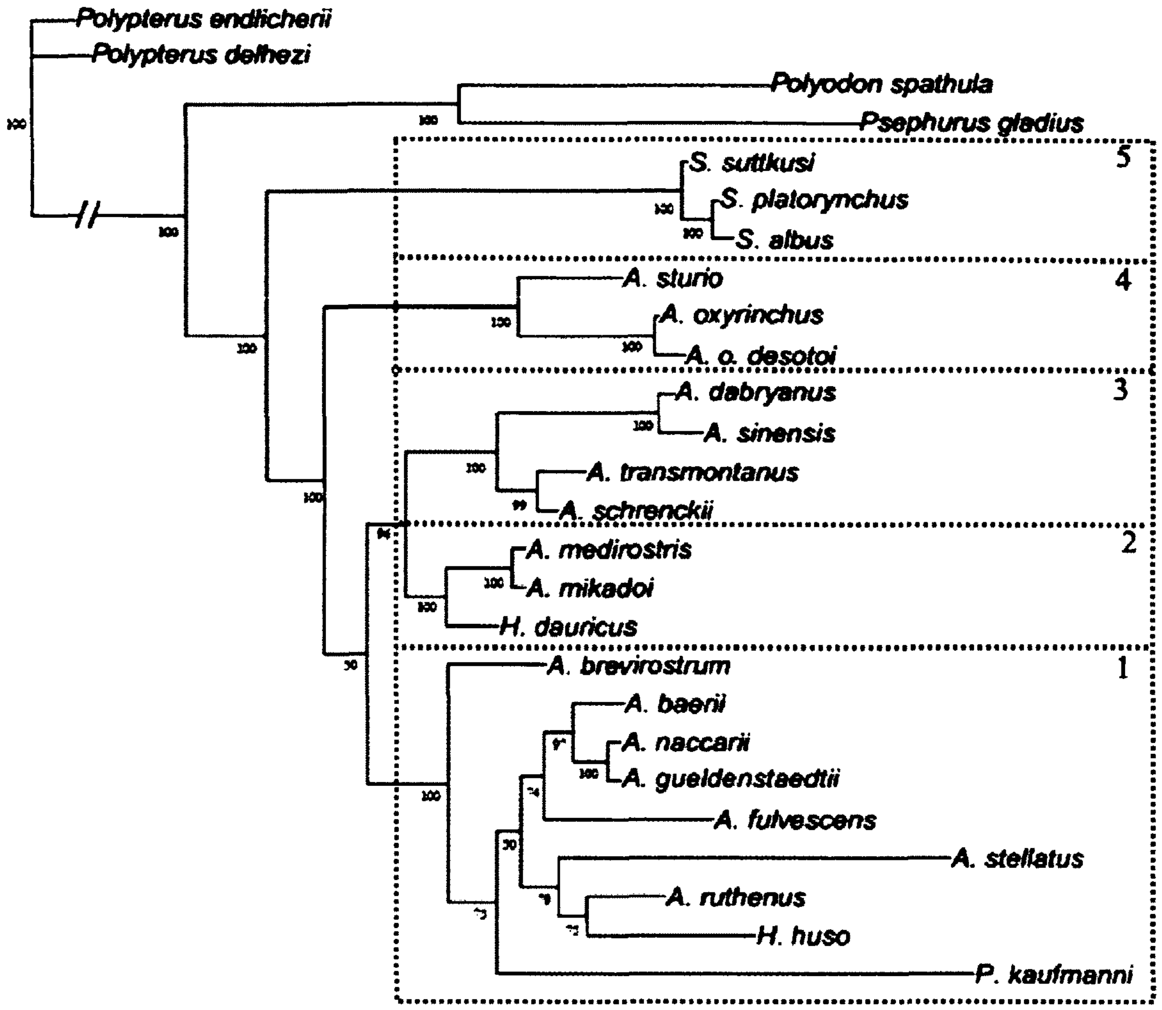


Figure 8: Bayesian phylogenetic hypothesis based on concatenated rRNA sequences. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4.

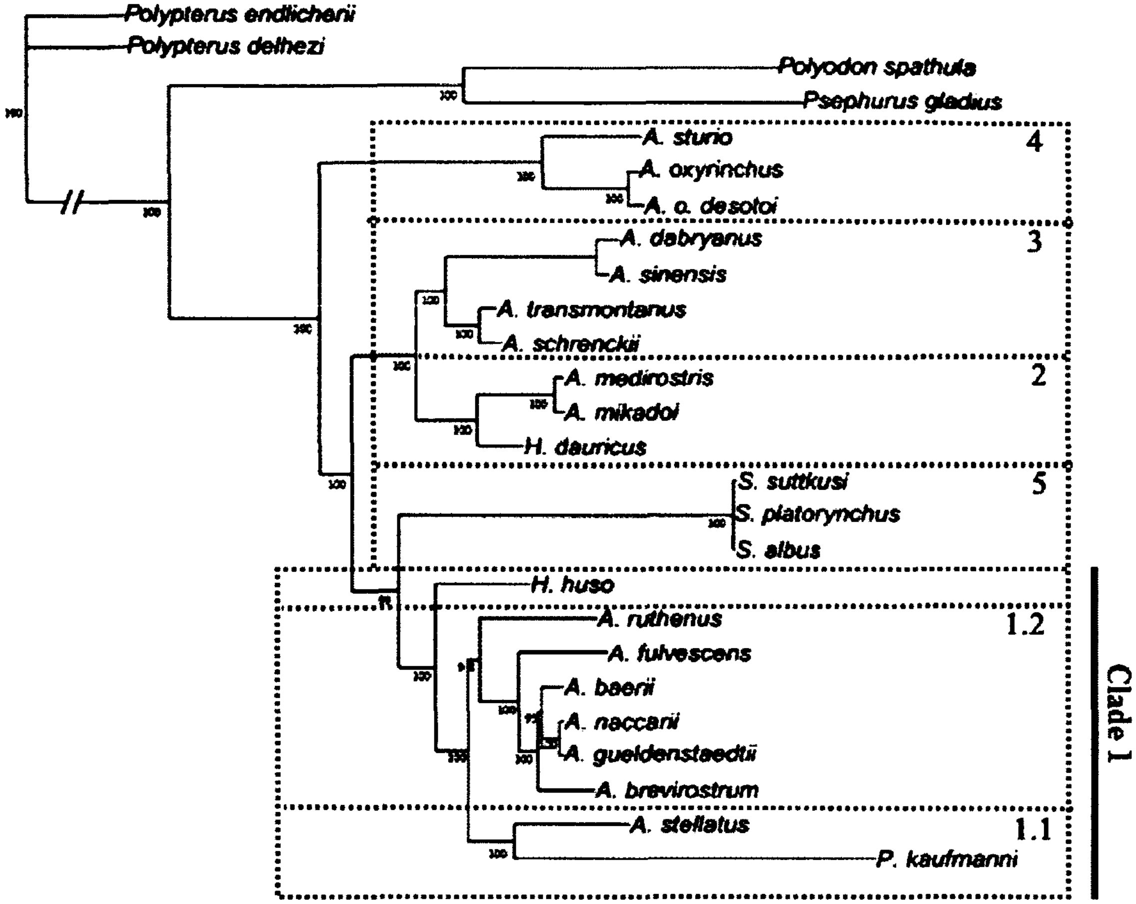


Figure 9: Bayesian phylogenetic hypothesis based on concatenated protein-coding sequences. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Clades (except for Clade 1) that match those recovered by full mitogenome topologies are shown in boxes with dashed lines and labeled with clade numbers listed in Table 4. Clade 1 is made up of groups 1.1, 1.2, and *H. huso,* and is identified to the right of the phylogeny.

CHAPTER 2.

The Biogeography of Extant Sturgeons

Abstract

Previous studies have investigated the biogeography of sturgeons (Acipenseridae), but the presentation of a new, full mitogenome phylogeny for the family presents the opportunity for a revised study. Here, I consider sturgeon phylogeny and their current distributions, along with phylogenetic and distribution information from lampreys and pinnipeds, to explore relationships among geographic areas inhabited by sturgeons. Three-item analysis, a comparative biogeography method, is employed. Eleven areas of endemism are identified, and relationships amongst these areas are proposed based on phylogenetic and species range data. These areas of endemism are recovered in three broad clades illustrating area relationship hypotheses: a Pacific clade, and Atlantic clade, and the Bering Sea. The Atlantic and Pacific clades are recovered as sister, reflecting previous biogeographic hypotheses. Relationships within each clade also follow previously described hypotheses and are explored in the context of geologic history. The Bering Sea area falls outside of this Atlantic-Pacific relationship.

1.0 Introduction

1.1 Sturgeons and Biogeography

Sturgeons (Acipenseridae) are a relatively old (at least 85 million years) group of ray finned fishes, with two known fossil genera: †*Protoscaphirhynchus* and †*Priscosturion* (Grande and Hilton 2006). The 25 broadly recognized extant sturgeon species exhibit a Holoarctic distribution, inhabiting wide and disjunct ranges exclusively in the temperate zones of Europe, Asia, and North America (Bemis and Kynard 1997, Choudhury and Dick 1998, Hilton and Grande 2006). Most extant sturgeons are anadromous, moving between fresh and saltwater to feed (primarily downstream movement) and to reproduce (upstream movement) (Bemis and Kynard 1997, Peng 2007). Sturgeon taxa are distributed among four commonly recognized genera (*Acipenser, Huso, Scaphirhynchus,* and *Pseudoscaphirhynchus*) in the monophyletic family Acipenseridae (Findeis 1997, Bemis *et al.*, 1997). Sturgeons, paddlefishes (Polyodontidae), and their fossil relatives form the order Acipenseriformes, which occupies a basal position among the other ray-finned fishes (Actinopterygii), the clade containing over half of all vertebrate diversity (Hilton *et al.* 2011). Across their ranges, sturgeons face threats such as habitat loss, overfishing, and other fishery interactions (Holzkamm and Waisberg 2004). These threats have led to the decline of sturgeon populations worldwide and the implementation of conservation measures that afford them protection (Holzkamm and Waisberg 2004). The disjunct distributions of sturgeon, their long evolutionary history, and their anadromous behavior have made this group of particular interest for biogeographic studies (Bemis and Kynard 1997).

Biogeography is the study of the relationship between taxa that inhabit the Earth and Earth's geography, with the goal of understanding the shared relationships between taxa and their distributions (Nelson and Platnick 1981, Posadas *et al.* 2006, Parenti and Ebach 2009). Several methods exist for biogeographic studies, and many of them consider phylogeny and use cladistic principles (Parenti and Ebach 2009). Comparative biogeography, a field that combines biogeography with systematics (Nelson and Platnick 1981), is defined by Parenti and Ebach (2009) as the use of hierarchical phylogenetic relationships of taxa to identify area relationships among biogeographic regions. Area relationships may be considered alongside geologic history and compared to historical biogeographic patterns (sets of area relationships shared by more than one taxon) previously revealed. Only after patterns of area relationships are discovered, should the process by which these relationships evolved, be interpreted (Parenti and Ebach 2009).

The first explicitly cladistic study of relationships within Acipenseridae (Findeis 1997) recovered *Huso* as the sister taxon of a clade comprising the other three genera, and recovered *Scaphirhynchus* and *Pseudoscaphirhynchus* as sister taxa forming Scaphirhynchini, which is recognized as Scaphirhynchinae by some authors (Myden and Kuhajda 1996). Many studies since Findeis (1997) have suggested broad changes to this topology. Widely accepted changes indicate that the tribe "Scaphrihynchini" (*sensu* Findeis) is paraphyletic, and suggest nonmonophyly of the genus *Huso* (Dillman *et al.* 2007, Krieger et al 2008, Hilton *et al.* 2011). A phylogenetic analysis of the full mitogenome (see Chapter 1) supports these changes and provides structure within the nonmonophyletic *"Acipenser."*

Many biogeographic studies have examined geographical distributions of sturgeons using the phylogenetic context available at the time. The many recent hypotheses of sturgeon phylogeny encourage a reconsideration of the historical biogeography of sturgeons. For example, Bemis and Kynard (1997) considered Scaphirhynchini to be monophyletic. They also considered *Huso* to be monophyletic and basal to other sturgeons, and noted that were *Huso* nested within *"Acipenser,"* biogeographical hypotheses would likewise change. Choudhury and Dick (1998) also considered Scaphirhynchini to be monophyletic, and based their biogeographic analyses on a modified version of a morphological phylogeny proposed by Artyhukin (1995). Peng *et al.* (2007) considered *Huso* to be nonmonophyletic and nested within *Acipenser*. Many studies (e.g. Vasileva *et al.* 2009) have confirmed the paraphyly of *Huso,* and, subsequently, its classification within *Acipenser*. Following the suggestion of Bemis and Kynard (1997), we propose a new analysis of sturgeon biogeography, based on recent, strongly supported changes to phylogeny.

Traditional biogeographic studies of sturgeons have focused on identifying a center of origin for sturgeons from which dispersal routes are inferred. The Ponto-Caspian Region, which includes the Mediterranean, Black, Caspian, and Aral seas and their tributaries, has the highest diversity of extant sturgeons (Bemis and Kynard 1997, Peng *et al.* 2007). This region has frequently been inferred to be the sturgeon center of origin (Bemis and Kynard 1997, Choudhury and Dick 1998, Peng *et al.* 2007), although some studies have ruled it out (Choudhury and Dick 1998). Conflicting hypotheses regarding the timing of divergences among sturgeon taxa have further led to the suggestion of varying and conflicting mechanisms and paths of dispersal from the Ponto-Caspian (Choudhury and Dick 1998, Peng *et al.* 2007, Bemis and Kynard 1997). Nevertheless, most studies agree that geologic events of the evolution of the Tethys Sea are critical to sturgeon biogeography (Bemis and Kynard 1997, Choudhury and Dick 1998, Peng *et al.* 2007).

The Tethys Sea was one of two large water bodies during the Triassic, and was situated in between Gondwana and Laurasia (Zonenshain and Pichon 1986); the Panthalassic Ocean covered the rest of the non-terrestrial world (Algeo *et al.* 2011). As geologic events shaped the earth, the Tethys Sea began to narrow during the Late Triassic, and eventually became separated from the Paratethys Sea (Bartock 1993, Rogl 1999). The Mediterranean Sea is the last remnant of the Tethys, and the other Ponto-Caspian Seas- the Black, Caspian, and Aral seas-formed from the Paratethys Sea (Rogl 1999), giving rise to the frequently presumed sturgeon center of origin. The proposal of a center of origin of a taxon has been criticized for many reasons, chief among them that the hypothseses are untestable and applied frequently to only one taxon at a time, and hence are not comparative (Croizat *et al.* 1974).

Past biogeographic studies of sturgeons have also focused on the disjunct distributions of closely related species. For example *A. oxyrinchus*, which is sympatric with *A. brevirostrum* inhabits the Atlantic coast and rivers of North America from Canada to Florida, while its sister species *A. sturio*, is distributed in the North Sea, Mediterranean, and Black Sea (Ludwig *et al.* 2002). Although *A. oxyrinchus* and *A. brevirostrum* are sympatric, phylogenetic studies have not found them to be closely related to one another (e.g. Birstein and DeSalle 1998; Laumann *et al.*, unpublished, Chapter 1). Another pair of sympatric sturgeon species, *A. transmontanus* and *A. medirostris*, are distributed along the North American Pacific Coast. Some recent studies have recovered a sister relationship between these two species and found them to be closely related to *A. schrenckii* (e.g., Birstein *et al.* 2002), whereas others have recovered them within a monophyletic group that also includes *A. dabryanus* (e.g., Dillman *et al.* 2007). Still others have not found them to be closely related (e.g., Krieger *et al.*2008). In Chapter 1, I recovered these two species to be within a broad clade also including *A. dabryanus, A. sinensis, A. schrenckii, A. mikadoi,* and *H. dauricus*. Within this clade, *A. medirostris* is closely related to *A. mikadoi* and *H. dauricus*, whereas *A. transmontanus* is sister to *A. schrenckii*, a clade which is sister to *A. sinensis* and *A. dabryanus*. Therefore, based on this and other topologies, there are trans-Pacific relationships reflected within Acipenseridae.

Sturgeons have also been of particular interest for biogeographic analysis due to their anadromous life history (Bemis and Kynard 1997). Diadromous taxa have frequently been the subject of biogeographic studies seeking to identify explanations for their distributions, to identify whether diadromy is a derived or ancient trait, and to identify the ancestral habitat of the taxa involved as freshwater or marine (Parenti 2008). Anadromy in sturgeons has been considered a derived trait with the assumption that the ancestral habitat was freshwater (Sulak and Randall 2002). As stated by Parenti (2008), identifying the ancestral habitat as either freshwater or marine relies on assuming that the habitat has (or has not) changed; this assumption is not necessarily supported (Parenti and Ebach 2009). Any number of events, from stranding (becoming stuck in a newly isolated habitat) to invasion, may be used to explain distributions of anadromous taxa, as they could be used to explain biogeographic distributions of any taxonomic group (Parenti 2008).

In this chapter, I employ comparative biogeographic methods to identify patterns of relationships among regions inhabited by sturgeons. I use a new phylogenetic hypothesis based on full mitogenomic data (Figure 1; Chapter 1) and current sturgeon distributions. Mechanisms leading to current distributions of sturgeons are not discussed. The aim of this study was to identify area relationships (relationships between geographic regions) using comparative biogeographic methods, consider them alongside geologic history, and discuss them in relation to area relationships suggested by other taxa distributed throughout the northern hemisphere. Broad biogeographic patterns that are considered alongside the area relationships hypothesized in this study include those involving the five areas of global distribution identified by Croizat (1958): Atlantic, Indian, Pacific, boreal, and austral (Figure 2). Historically the boreal and austral regions have been associated with the distributions of holoarctic taxa like sturgeons (Nelson and Ladiges 2001, Parenti and Ebach 2009). These regions have been repeatedly linked with the Pacific by biogeographic studies, while the Atlantic has been linked with the Indian regions. Nelson and Ladiges (2001) distilled these frequently proposed area relationships into a "cladistic summary of global patterns" (Figure 3).

1.2 History of Areas Inhabited by Sturgeons

One important aspect of comparative biogeographic analysis is the consideration of the geologic history of areas under consideration, in this case the areas inhabited by extant sturgeon taxa. The geologic history of these areas is discussed below.

1.2.1 Geologic History: Formation of the Atlantic and Pacific Oceans and Ponto-Caspian Region

During the Triassic, Pangea enclosed the Tethys Sea on three sides (Zonenshain and Pichon 1986) and an archipelgo comprising portions of China defined the eastern border of the Tethys (Yang *et al.* 1982). The Panthalassic Ocean covered the rest of the Earth (Algeo *et al.* 2011). The oldest of the world's oceans, the Pacific arose from the Panthalassic Ocean at least 167 million years ago (Neall and Trewick 2008). In the Late Triassic and Early Jurassic, Pangea began breaking into two super-continents, Laurasia and Gondwana (Golonka 2007). Laurasia, to the north, comprised present day North America and Eurasia (excluding the Indian subcontinent). Gondwana, the southern supercontinent, was composed of South America, Africa, Australia, India, and the Arabian Peninsula. The shift of Laurasia away from Gondwana opened the North Atlantic, then an arm of the Tethys Sea (McHone 2000, Thiede 1979). The rifting of Pangea also resulted in the formation of the Gulf of Mexico and the movement of Eurasia southwards, narrowing the Tethys Sea (Bartock 1993). Uplift of Great Britain and Ireland began, as did a cycle of marine incursions, flooding what is now the North Sea (Ziegler1975, Torsvik *et al.* 2002).

Rising sea levels inundated much of Laurasia so that large portions of North America and Europe were submerged by 100 million years ago (Cretaceous) (Schlee 1999). During the Cretaceous, the South Atlantic opened and the Atlantic Ocean as a whole continued to widen (Thiede 1979, Fairhead 1988). Connections between the narrowing Tethys and the Atlantic closed, with the Tethys separating Laurasia from Gondwana (Rogl 1999). Modern day continents that formed Laurasia began to separate, as did Gondwana (Thiede 1979). During this period, much of North America and Eurasia were submerged, with Eurasia separated from Gondwana by the Tethys Sea (Hay *et al.* 1993). North America was inundated by the Western Interior Seaway (Hay *et al.* 1993).

In the Paleocene, the continents composing Laurasia continued to move away from one another, but North America was still connected to Asia by the Bering Land Bridge and Europe by Greenland; this separated the future Pacific from the Bering Sea (Hopkins 1973). Likewise, the continents that composed Gondwana were shifting, with India moving toward Asia (Tapponier *et al.* 1986). Africa shifted toward Europe, continuing the narrowing of the Tethys. In the Eocene, North America became more separated from Eurasia as it broke apart from Greenland and Europe; this movement opened the North Sea, allowing it to deepen (O'Leary *et al.* 2004 ; Kender *et al.* 2012). Also during the Eocene, India collided with Asia forming the Himalayas (Tapponier *et al.* 1986).

During the Oligocene, the Atlantic continued to widen (Fairhead 1988). The continued growth of the Himalayas separated a section of the Tethys Sea, forming the Paratethys Sea (Rogl *et al.* 1999). The Paratethys Sea was fed with water from the North Sea and was connected to the Mediterranean (Rogl 1999). At this time, the African and Eurasian plates converged, isolating what remained of the Tethys Sea from other water bodies except for the Indian Ocean. The isolation of the Tethys from other water bodies was completed in the Miocene, sealing it in to today's Mediterranean Basin (Rogl *et al.* 1999). Today's Mediterranean Sea is the last remnant of the Tethys Sea (Rogl *et al.* 1999), and the Paratethys Sea formed the Black, Caspian, and Aral seas; these together with the Mediterranean Sea form the Ponto-Caspian Region (Hakanson *et al.* 2002). The isolation of the North Sea (from the Tethys) was followed by

Miocene shallowing of the North Sea and the formation of deltas in the area (Torsvik *et al.* 2002).

1.2.2 Geologic History: The Arctic Ocean and Bering Sea

The geologic history of the Arctic is poorly known and contentious (Clark 1974, 1975, Shepard *et al.* 2013, Moran *et al.* 2006). In the early Miocene, around 18 million years ago, a connection between the Arctic Ocean and the Atlantic was established in the form of the Fram Strait (Jacobson *et al.* 2007). Initially shallow, the strait allowed exchange of water between the two water bodies, transitioning the Arctic through an estuarine phase into the ocean of today (Thompson *et al.*2010; Engen *et al.* 2008). Fluctuations in sea level led to the opening and closing of the Bering Sea, alternately linking and isolating the Arctic from the Pacific (Elias *et al.* 1996). About 11,000 years ago, sea level rise caused the final inundation of the Bering land bridge, connecting the Bering Sea to the Pacific via the Bering Strait (Elias *et al.* 1996). Today, the Arctic and Atlantic oceans are linked by the Barents, Greenland, and Norwegian seas and the Denmark Strait. The Arctic Ocean remainsconnected to the Pacific via the Bering Sea.

1.2.3 Recent Geologic History: Baltic and North Seas

Pleistocene glaciations and interglacial periods caused several periods of drying and subsequent filling of the North Sea. During one of the interglacial periods, the Eemian (around 130,000 and 115,000 years ago), the North Sea expanded (Torvick *et al.* 2002). During this time, the first body of water known from the Baltic Basin, the Eemian Sea, formed (Torvick *et al.* 2002, Miettinen 2003). The Eemian Sea was larger than the Baltic Sea is today and had connections to the Atlantic Ocean and the North Sea (Miettinen 2003). The transformation from the Eemian Sea to today's Baltic Sea occurred through several stages, involving freezing, melting, different salinity regimes, and varying connections to the Atlantic Ocean and North Sea. The Weischlian Glaciation, the last glaciation of Earth, was characterized by the coverage of much of Eurasia by ice sheets (Siegert *et al.* 2001). This resulted in the drying of large portions of the North Sea to form plains (Torsvik *et al.* 2002, Konradi *et al.* 2005).

About 13,000-14,000 years ago, receding glaciers isolated sections of the Eemian Sea, which merged to form the Baltic Ice Lake (Morner 1995, Vassiljev and Saarse 2013). During the time of the Baltic Ice Lake, cooling and warming periods resulted in expansion and subsequent regression of ice sheets as well as uplift of land, changing the shape of the Baltic Ice Lake (Morner 1995, Vassiljev and Saarse 2013). The exact timing of the formation of the Baltic Ice Lake is controversial, but it is hypothesized to have been in existence between 13,800 and 10,300 years ago (Morner 1995, Vassiljev and Saarse 2013).

Toward the end of the Weisclian Glaciation, warming temperatures caused the final retreat of ice sheets and increased sea level. Water again began filling the North Sea, and once again a connection between the North Sea and the Baltic Ice Lake formed (Lepparanta and Myreberg 2009). North Sea waters inundated the lake, increasing salinity and forming the Yoldia Sea, which persisted until about 9,500 years ago (Andren *et al.* 2002, Lepparanta and Myreberg 2009).

Continuing uplift of land separated the Yoldia and North seas, resulting in the formation of the freshwater Ancylus Lake from the Yoldia Sea (Morner 1995). About 8,150 years ago, a Tsunami hit the North Sea, inundating the remaining dry areas (Weninger *et al.* 2008). Portions of Ancylus Lake were still ice-covered and uplift continued, gradually forming a new river (Dana River) between Ancylus Lake and the North Sea (Lepparanta and Myreberg 2009). The river widened into a strait, introducing saltwater into Ancylus Lake once again. This led to the formation of the Littorina Sea about 7,500 years ago (Sandgren *et al.* 2004). The exchange of water between the North and Littorina seas and sea level rise led to higher levels of salinity than present in the Baltic Sea (Sandgren *et al.* 2004). Influx of seawater into the Littorina Sea continued until about 4,000 years ago, when sea level rise ended and salinity began decreasing in what is now known as the Baltic Sea (Sandgren *et al.* 2004, Berglund *et al.* 2005). The Baltic Sea is brackish and continues to shrink due to uplift (Walday and Krogland 2011). It is connected to and exchanges water with the North Sea (Walday and Krogland 2011).

1.2.4 Recent Geologic History: East China, Yellow, and Japan Seas

The Sea of Japan first opened 32 million years ago, a process that continued until about 15 million years ago at which point it was an isolated body of water (Tada *et al.* 2013). About 2.6 million years ago, the oldest portion of the East China Sea began opening (Gallagher *et al.* 2015). Around this time, the Sea of Japan became connected to the Pacific Ocean by the opening of the Tsugaru and Tsushima straits (Uda 2016, Gallagher *et al.* 2015). Linkages between the Sea of Japan and the Pacific continued intermittently with sea level rise and fall (Liu and Milliman 2004). Today, the Sea of Japan maintains a connection to the Pacific Ocean via the Tsugaru Strait.

The East China Sea and Yellow Sea were still primarily terrestrial or subtidal areas 20,000 years ago (Lee 2012). Multiple phases of sea level rise inundated these areas and led to both the East China and Yellow Seas as we know them today. By 15,000 years ago, rising sea levels filled the East China Sea and reached the southern portion of the Yellow Sea (Liu and Milliman 2004, Yi *et al.* 2003). Subsequent periods of rapid sea level rise led to the gradual filling of the Yellow Sea, from the south to the north, over the next 5,000 years (Liu and Milliman 2004, Yi *et al.* 2003). The Tsushima Strait connects the East China Sea to the Sea of Japan today.

1.2.5 Recent Geologic History: Interior North America, The Mississippi and Missouri Rivers and their tributaries

The Mississippi River and its tributaries existed before the Pleistocene, but their courses were so different from those seen today that the history of these rivers is commonly considered to begin with the end of the last glacial stage (Leverett 1921). The Missouri River also formed before the last ice age, but its course was significantly diverted by glaciation (Trimble 1980). Much of North America was covered with ice sheets and glaciers during the Pleistocene (Leverett 1921, Trimble 1980). 13,000 years ago, glacial melt flooded the area now occupied by Lake Champlain and the Gulf of Saint Lawrence, forming the Champlain Sea (Russell and Cummings 2009, Mather 1917). As glacial retreat continued in the Holocene, continental rebound caused this arm of the Atlantic Ocean to shrink (Mather 1917). Further melting formed the Great Lakes Erie and Michigan, which drained into the Mississippi, about 10,000 years ago (Larson and Schaetzl 2001, UWSG 2013). By one thousand years later, Lake Superior, which also drained into the Mississippi, had formed. Sea level rise about 7,500 years ago flooded the Great Lakes, connecting them to form one large lake, Lake Nipissing, which drained into the Mississippi and the Atlantic (Larson and Schaetzl 2001, UWSG 2013). By 3,000 years ago, the level of the lake had fallen, resulting in the current configuration of the Great Lakes (Farrand 1988, UWSG 2013). The final retreat of glaciation also established the Mississippi and Missouri Rivers in the (approximate) courses they occupy today, draining into the Gulf of Mexico (Wickert *et al.* 2013, Trimble 1980, Leverett 1921).

2.0 Methods

In keeping with the cladistic nature of my dissertation and following the detailed example of Hoagstrom *et al.* (2014), I selected a comparative biogeographical method of analyses based on hierarchical relationships (Parenti and Ebach 2009). This method is rooted in the use of three-item analysis to infer relationships among geographic areas (Parenti and Ebach 2009). In three-item analysis, phylogenies of taxa that have overlapping distributions are converted to areagrams. An areagram is a cladogram on which species names have been replaced by the areas the taxon inhabits. It is the overlap of multiple species (with hypothesized evolutionary history) that provides the shared history informing relationships of the areas-termed areas of endemism. Analyses are then conducted to extract three-area relationships (i.e., hierarchical relationships reflecting the relationships among sets of three areas). Three-area relationships are combined to form optimal area cladograms, which are further broken down into an intersection tree, which displays the three area relationships recovered by all optimal area cladograms. These relationships may be explored by examining the geologic history of areas of endemism; geologic patterns considered alongside biological ones allow for reciprocal illumination- each method sheds light on the other. (Parenti and Ebach 2009).

2.1 Phylogenies for Comparison

The full mitogenome sturgeon (Figure 4) phylogeny was used for the exploration of area relationships based on sturgeon biogeography, as were phylogenies for lampreys (Figure 5; Gill *et al.* 2013), and pinnipeds (Figure 6; Arnason *et al.* 2006).. These taxa have overlapping ranges with sturgeons, and phylogenetic hypotheses were available for them.

The lamprey phylogeny (Gill *et al.* 2003) was based on 32 morphological and karyological characters. The original phylogeny included 18 lamprey species. Three of these species (*Mordacia mordax, M. lapicida,* and *Geotria australis*) do not have ranges that overlap with those of sturgeons or within my proposed areas of endemism and were pruned from the phylogeny, resulting in a 15 terminal taxa topology (Figure 5).

The pinniped phylogeny (Arnason *et al.* 2006) was based on protein-coding sequences of the mitogenome, and included 28 species. Sixteen of these species overlapped with sturgeons in the proposed areas of endemism, and the remaining species (*Otaria byronia, Arctocephalus pusilis, Phocarctos hookeri, Neophoca cinerea, A. forsteri, A. australis, Monachus schauinslandi,* *Ommatophoca rossi, Lobodon carcinophagus, Leptonychotes weddelli, Hydrurga leptonyx,* and *Pusa sibirica*) were pruned from the topology (Figure 6).

2.2 Species Ranges and Phylogenies

I conducted literature and fishery landing searches to map the ranges of extant sturgeon species in my full mitogenome Bayesian phylogeny of extant sturgeons (Figure 7). This sturgeon phylogeny (Figure 1; Chapter 1) was used in analyses described below. Maps were superimposed upon one another and were used, along with Bemis and Kynard's (1997) sturgeon biogeography study, to identify potential areas of endemism (Figure 4). Once these areas were identified, the ranges of lamprey and pinniped species were identified from the literature (Gill *et al.* 2003, Arnason *et al.* 2006), IUCN RedList species range maps and descriptions (IUCN 2014) and the global fish species database FishBase (Froese and Pauly 2016). Lamprey and pinniped taxa that do not inhabit the areas of endemism identified for sturgeons were pruned (using the command drop.tip in the ape package, version 3.4, Paradis *et al.* 2004, for R, R Core Team 2014) from the available phylogenies.2.3 Areas of Endemism and Three-Item Analysis

Areas of endemism and pruned phylogenies were converted into areagrams and coded for LISBETH version 1.3 (Bagils *et al.* 2012). LISBETH conducts three-item analysis using fractional weighting and paralogy free sub-tree analysis (Nelson and Ladiges 1996). Three-area statements were extracted from LISBETH and imported into PAUP*4b10 (Swofford 2003). PAUP was used to identify optimal taxon-area cladograms (by combining TACs). Optimal TACs were then imported into LISBETH, where an intersection tree including all three-area relationships was constructed. Results are reported with a "completeness index", which is the percentage of three-area statements in both the intersection trees and the paralogy free sub-trees. A high completeness index is indicative of agreement on area relationships among all taxa considered

(Hoagstrom *et al.* 2014). Results are discussed in comparison to geologic history of the areas of endemism considered.

3.0 Results

3.1 Areas of Endemism

Proposed areas of endemism followed "provinces" used by Bemis and Kynard (1997) to explore sturgeon biogeography. Modifications were made to Bemis and Kynard's (1997) provinces, and two additional areas of endemism were proposed. Proposed areas of endemism are given different names here to reflect that the regions have been modified from those of Bemis and Kynard (1997). Eleven putative areas of endemism were proposed (Figure 4) and are defined as follows.

1) Japan (JPN) Region
    a. Includes the Sea of Japan, Amur Basin, Sea of Okhotsk, and Tatar Strait,
    b. Corresponds to Bemis and Kynard's (1997) ASJ province, but is expanded to include the entire Sea of Okhotsk,
    c. Inhabited by *A. mikadoi, A. schrenckii,* and *H. dauricus.*
2) China (CHN) Region
    a. Includes the Yangtze Basin and the southern portion of the East China Sea,
    b. Corresponds to Bemis and Kynard's (1997) CH province, but does not include the part of the coast along the Yellow Sea,
    c. Inhabited by *A. dabryanus* and *A. sinensis.*
3) Yellow Sea (YLW) Region
    a. Includes the Yellow Sea itself, and the coast of China along the Yellow Sea,

   b. The Yellow Sea was not included in Bemis and Kynard's (1997) study, but part of the coast of China along the Yellow Sea was included in their CH province,
   c. Inhabited by *A. sinensis.*

4) Pacific Coast of North America (PNA) Region
   a. Includes the Pacific Coast and rivers draining into the Pacific,
   b. Corresponds to Bemis and Kynard's (1997) NEA province, but is slightly expanded (here) to the south,
   c. Inhabited by *A. medirostris* and *A. transmontanus.*
5) Bering Sea (BRG) Region
   a. Includes the Bering Strait,
   b. Was not included in Bemis and Kynard's (1997) study,
   c. Inhabited by *A. mikadoi.*
6) Baltic and North Seas (BNS) Region
   a. Includes the Gulf of Bothnia,
   b. Corresponds to Bemis and Kynard's (1997) NEA province, but is expanded to include the Gulf of Bothnia,
   c. Inhabited by *A. sturio.*
7) Ponto-Caspian (PTC) Region
   a. Includes the Black, Azov, Caspian, Mediterranean and Adriatic seas,
   b. Corresponds to Bemis and Kynard's (1997) PT province,
   c. Inhabited by *A. gueldenstaedtii, A. naccarii, A. ruthenus, A. stellatus, A. sturio, H. huso,* and *P. kaufmanni.*
8) Arctic (AES) Region

    a. Includes the Arctic Ocean and East Siberian Sea,
    b. Corresponds to Bemis and Kynard's (1997) SAO province, expanded to include Arctic North America,
    c. Inhabited by *A. baerii, A. mikadoi, A. ruthenus,* and *A. stellatus.*

9) Atlantic Coast of North America (ANA) Region
    a. Includes the Atlantic and rivers draining into it,
    b. Corresponds to Bemis and Kynard's (1997) NWA province, expanded slightly northward
    c. Inhabited by *A. brevisotrum* and *A. o. oxyrinchus.*

10) Gulf of Mexico Region (MGM)
    a. Includes the Mississippi and Mobile basins,
    b. Corresponds to Bemis and Kynard's (1997) MGM province,
    c. Inhabited by *A. fulvesencs, A. o. desotoi,* and *Scaphirhynchus.*

11) North American Lakes (NAL)
    a. Includes the Great Lakes, Hudson Bay, and Lake Champlain,
    b. Corresponds to Bemis and Kynard's (1997) GL province,
    c. Inhabited by *A. fulvescens* and *A. oxyrhynchus.*

Each sturgeon (Table 1), lamprey (Table 2), and pinniped (Table 3) species was assigned to all areas of endemism that they inhabit (see above). Areagrams were constructed using the phylogenies of each taxonomic group and the areas of endemism they inhabit; thees areagram is shown in Figure 8, the lamprey areagram in Figure 9, and the Pinniped areagram in Figure 10.

3.2 Intersection Tree

The Sturgeon-Lamprey-Pinniped intersection tree had a completeness index of 92.7%. This intersection tree (Figure 11) suggested two clades of related areas, a Pacific clade and a primarily Atlantic clade. The Pacific clade consisted of China and the Yellow Sea areas of endemism grouping as sister areas, with Pacific North America outside of them. Outside of this relationship was the Japan area of endemism, with the Arctic area just outside of that grouping of four areas. The Atlantic clade recovered the Ponto-Caspian and Baltic/North Sea areas of endemism as sister areas. The North American Lakes area of endemism was sister to the Gulf of Mexico area, with the Atlantic coast of North America area falling just outside. The Atlantic and Pacific clades grouped together, with the Bering Sea area of endemism outside of that sister-area relationship.

4.0 Discussion and Conclusions

4.1 Anadromy

Although I do not attempt to identify the ancestral habitat of sturgeons as marine or freshwater, or to identify the processes by which sturgeons came to inhabit their current ranges, it is interesting to consider the areas of endemism occupied by the strictly freshwater taxa (Figure 1). Exclusively freshwater sturgeons are present in the Japan, China, Ponto-Caspian, Arctic, Gulf of Mexico, and North American Lakes Regions. These regions are also occupied by anadromous species. Except for *Scaphirinchus,* none of the sturgeon clades in this analysis is exclusively freshwater. This may be due to taxon sampling. For example, only one species of *Pseudoscaphirhynchus* was included in this study. *Pseudoscaphirhynchus* is generally considered to be a clade also containing *P. kaufmanni, P. hermanni* (Birstein *et al.* 2002, Dillman *et al.* 2007, Hilton *et al.* 2011) and the presumed recently extinct *P. fedtschenkoi* (IUCN 2014). The three *Pseudoscaphirhynchus* taxa are freshwater species, and as such represent an entire freshwater clade (IUCN 2014). Parenti (2008) noted that, without making assumptions that the habitat of a taxon has changed from either marine to freshwater or vice versa, the ancestral habitat can be reconstructed as one that includes both fresh and saltwater areas. This analysis does not contradict this statement, and, in fact, supports it.

4.2 Area Relationships

Many of the area relationships hypothesized via three-item-analysis mirror those suggested by geologic history and by past biogeographic studies. Area relationships are divided into two primary area clades: 1) the Atlantic Clade containing three North American Regions (Lakes, Atlantic Coast, and Gulf of Mexico), the Baltic/North Sea Region, and the Ponto-Caspian Region; and 2) the Pacific Clade, composed of Pacific North America, China, Yellow Sea, Japan, and Arctic Regions. The Bering Sea Region was recovered separate from these two clades.

4.2.1 Atlantic Clade

4.2.1a Atlantic Clade: North American and Baltic/North Sea/Ponto-Caspian Regions

Connections between the two Atlantic Region cladesare exemplified in the relationships between two clades of sturgeons. In the first clade, *Acipenser sturio,* which ranges throughout the Ponto-Caspian and Baltic/North Sea regions, is sister to the two subspecies of *A. oxyrinchus,* which inhabit the North American Atlantic Regions (including Gulf of Mexico). The second clade contains six Ponto-Caspian species (*H. huso, A. ruthenus, A. naccarii, A. gueldenstaedtii, A. stellatus,* and *P. kaufmanni*), *A. brevirostrum* from the Atlantic Coast of North America, and *A. fulvescens,* which ranges in the Gulf of Mexico Region and North American Lakes Regions. Of interest, two of the Ponto-Caspian species, *A. stellatus* and *A. ruthenus,* are also found in the Arctic Region, as is *A. baerii* which is also part of this sturgeon clade.

Current geography also supports a relationship between the two Atlantic Region Area Clades. The North Sea connects the Atlantic Ocean to the Baltic Sea today; these waters have been intermittently connected since at least 115,000 years ago. The Atlantic is also currently connected to the Mediterranean (included in the Ponto-Caspian Region) via the Strait of Gibraltar. Although the other seas in the Ponto-Caspian region do not currently connect to the Atlantic Ocean, the origin of the Atlantic as part of the Tethys Sea in the Jurassic, and continued connection to the Tethys until the Cretaceous, supports these area relationships.

4.2.1b Atlantic Clade: North American Regions

The recovery of three area relationships among the Atlantic Coast of North America, the North American Lakes, and the Gulf of Mexico Regions, in which the latter two areas are most closely related to one another, is supported by three sturgeon taxa and geologic history. *Acipenser fulvescens* is found in both the Gulf of Mexico and North American Lakes Regions. Linkages between these two regions have occurred intermittently, with the Great Lakes draining into the Mississippi at some points over the last 10,000 years. None of the sturgeon taxa considered in this analysis inhabits all three North American regions, but *A. o. oxyrinchus* lives in the North American Lakes region and along the Atlantic Coast. Its subspecies *A. o. desotoi* is endemic to the Gulf of Mexico. These evolutionary relationships and the area relationships they suggest are strengthened by the geologic connections between the Atlantic Ocean and the other two North American Regions. The North American Lakes Region was historically connected to the Atlantic Ocean via the Champlain Sea, and has a current connection to the Atlantic through the Gulf of Saint Lawrence. The Gulf of Mexico is connected to the Atlantic Ocean by the Florida Straits, and the Mississippi River had historic connections to the Atlantic.

4.2.1c Atlantic Clade: Baltic/North Sea and Ponto-Caspian Regions

The Baltic/North Sea and Ponto-Caspian regions are hypothesized to have a sister-area relationship to one another. A single sturgeon species in this analyses, *Acipenser sturio,* inhabits the Baltic/North Sea Region; it is also found in the Ponto-Caspian Region. The Oligocene connection between the North Sea and Paratethys Sea supports this area relationship.

4.3 Pacific Clade

4.3.1 Pacific Clade: Asia and Pacific North America

A seven-species clade supports the area relationships among several Pacific regions: (((China Region + Yellow Sea Region) Pacific North America Region) Japan Region). The taxa reflecting this clade include two sister species, *A. dabryanus* and *A. sinensis,* that inhabit the East China Sea. *Acipenser sinensis* is also found in the Yellow Sea. Sister to *A. dabryanus* and *A. sinensis* is another pair of sister-species: *A. transmontanus*, which inhabits the Pacific North America region and *A. schrenckii,* which is from the Japan Region. Another pair of Pacific North America-Japan Region taxa, *Acipenser medirostris* and *A. mikadoi*, respectively, form a sister species relationship. Together with the Japan Region species *H. dauricus,* these taxa form a sister clade to the four taxa listed above.

The shared geologic history between the East China and Yellow Seas, which were filled sequentially by an increase in sea level, further supports the hypothesized biogeographic relationship between these two areas. Although it might seem that these areas should be most closely related to the Sea of Japan, as they are all marginal seas of the western Pacific, this analysis and geologic history do not support such a relationship.

In this analysis, the East China and Yellow seas are more closely related to Pacific North America than they are to the Sea of Japan. Although the Sea of Japan was connected to the Pacific long before the East China Sea was flooded, about 2.6 million years ago, these connections have been ephemeral. The Sea of Japan has also maintained intermittent connections with the East China Sea. This long history has provided for periods of exchange of water and taxa between the Sea of Japan and the Pacific, illustrated by the sister species relationships between *A. transmontanus* and *A. schrenckii*, and *A. medirostris* and *A. mikadoi*. The East China Sea and Yellow Sea, however, have had a constant connection to the Pacific Ocean, and therefore provided more consistent opportunity for taxonomic connections since they formed.

Current and past connections between the East China Sea and the Yellow Sea, along with the fact that the only sturgeon taxon that inhabits the Yellow Sea (*A. sinensis*) also inhabits the China Region, suggest that subsequent analyses should consolidate these into a single region or area of endemism.

4.3.2 Pacific Clade: Pacific and Arctic

Four turgeon species, *A. baerii, A. mikadoi, A. ruthenus,* and *A. stellatus* occur in the Arctic region. The relationships hypothesized in this analysis among the Arctic and other Pacific Clade Regions are not strongly supported by the phylogeny of sturgeons alone. A single taxon occurring in the Arctic Region (*A. mikadoi*) also occurs in other Pacific Clade Regions. This species is sister to a Pacific North American species, and part of a clade with other exclusively Pacific Region taxa. Two other taxa from in the Arctic (*A. ruthenus* and *A. stellatus*) occur in the Ponto-Caspian Region of the Atlantic Area clade. Another species (*A. baerii*) is sister to two Ponto-Caspian Region species (*A. naccarii* and *A. gueldenstaedtii*).

The examination of areagrams based on phylogenetic relationships of other taxa included in three-item analysis help clarify the history of the Arctic Region. Two lamprey taxa inhabit the Arctic region- *Lethenteron camtschaticum* and *Petromyzon marinus* (IUCN 2014). The first of these also inhabits the Pacific region, and is phylogenetically most closely related to two pairs of sister species; one species in each pair inhabits the Atlantic region, and one in each pair the Pacific (IUCN 2014, Gill *et al.* 2003). Along with the two species pairs, *L. camtschaticum* is part of a larger clade composed of five other Pacific and one Atlantic species (IUCN 2014, Gill *et al.* 2003). *Petromyzon marinus,* on the other hand, inhabits the Atlantic as well as the Arctic (IUCN 2014). This taxon is in a clade with three Atlantic species (Gill *et al.* 2003). Of the pinnipeds included in three-item analysis, *Odobenus rosmarus, Histriophoca fasciata, Pagophilus groenlandicus, Pusa hispida* and *Phoca vitulina* range in the Arctic (Arnason *et al.* 2006, IUCN 2014). *Pusa hispida* and *P. vitulina* are also be found in regions in the Atlantic and Pacific Regions of this analysis (Arnason *et al.* 2006, IUCN 2014). *Histriophoca fasciata* is found in the Pacific as well as the Arctic, while its sister species *P. groenlandicus* is found in the Atlantic (Arnason *et al.* 2006, IUCN 2014). *Phoca vitulina* is sister to a Pacific Clade species, and these sister species form a polytomy with *H. grypus,* an Atlantic species, and *P. hispida* (Arnason *et al.* 2006, IUCN 2014). These phylogenetic relationships suggest area relationships among the Arctic, Pacific, and Atlantic, but the evolutionary history of *O. rosmarus* supports inclusion in the Pacific Clade (IUCN 2014). *Odobenus rosmarus* can be found in the Bering Sea Region as well as the Arctic (IUCN 2014). Althought it is not found in the Atlantic or Pacific, it is part of a clade including four other pinnipeds that inhabit only Pacific Regions in our analysis (Arnason *et al.* 2006, IUCN 2014).

Although the Arctic became geographically connected to the Atlantic in the Cenozoic, long before its connection to the Pacific, other studies show a similarly close area relationship between the Arctic and Pacific as opposed to an Arctic-Atlantic area relationship. Early Cenozoic taxa in the Arctic and East Siberian Seas (my Arctic Region) also inhabited the Atlantic (Marincovich and Gladenkov 2001). During the mid-to-late Cenozoic, with the opening of the Bering Strait (Brigham-Grette 2001, Marincovich and Gladenkov 2001), the fauna of the Arctic became dominated by taxa common in the Pacific, indicating migration from the Pacific to the Arctic (Marincovich and Gladenkov 2001). This relationship also mirrors past studies that recover Croizat's "boreal" region, which encompasses my Arctic Region, with the Pacific Ocean rather than the Atlantic.

4.4 Bering Sea

Three-item analysis recovered the Atlantic and Pacific Area clades as "sister", with the Bering Sea Region outside of these clades, although this is not strongly supported by geologic history. Geographically, the Bering Sea provides a connection between the Pacific and Arctic oceans. The Bering and Arctic seas are also an aquatic "bridge" between the Pacific and Atlantic. Therefore, one might predict that taxa in the Bering Sea Region would be more closely related to the taxa in the Pacific Area Clade than to those of the Atlantic Clade. Past studies also support inclusion of the Bering Sea with the Pacific; Croizat's "boreal" region includes my Bering Sea Region, and has been repeatedly associated with the Pacific.

Among sturgeons, only *A. mikadoi* occurs in the Bering Sea; it also occurs in the Pacific and is closely related to other Pacific taxa (IUCN 2014). The two lamprey species occurring in the Bering Sea Region, *Lethenteron camtschaticum* and *Entosphenus tridentatus,* also occur in the Pacific Region (IUCN 2014). *Entosphenus tridentatus* groups in a clade with other Pacific species, while *L. camtschaticum* is in a clade with two Pacific and two Atlantic taxa (IUCN 2014, Gill *et al.* 2003). Of the pinniped taxa that occur in the Bering Sea, two occur in both the Atlantic and Pacific regions and four occur in Pacific Regions (including the Arctic) (IUCN 2014). The area relationships of the Bering Sea may be strongly influenced by widespread and highly migratory pinniped taxa that inhabit this region. These factors suggest re-examination of whether the Bering Sea Region should be a distinct area for future analyses. Further analyses including more taxa may help clarify the issue.

4.5 Conclusions

With the exception of the Bering Sea, the area relationships proposed here are supported by geologic history. Excluding the Bering Sea Region, broad area relationships, namely the inclusion of the Arctic Region within the Pacific Area Clade, reflect patterns that have been proposed in past studies (Marincovich and Gladenkov 2001). Future studies should consider different or additional taxa to clarify the issues with the hypothesized Bering Sea Region area relationships. They should also consider consolidation of the Yellow Sea Region into the China Region.

Mechanisms leading to current distributions of sturgeons were not discussed here for multiple reasons. First, the geologic history of some of the regions considered is contentious (e.g. the Arctic, see Clark 1974). Second, even with reliable dates for geologic events, a well-supported, time-calibrated sturgeon phylogeny would be needed to hypothesize potential specific mechanisms such as invasions or strandings. Finally, the methods of biogeography that I endorse (comparative biogeography) do not include selection of ancestral habitats as either strictly marine or freshwater. Because distributional mechanisms are not discussed, the evolution of migratory habits and how they relate to geologic history (previously discussed by Bemis and Kynard 1997) are not discussed here. Despite the lack of speculation on the ancestral habitat and on potential mechanisms for the evolution of anadromy, the occurrence of both anadromous and strictly freshwater taxa in nearly every region, as well as the occurrence of anadromy and freshwater habits in taxa at varying levels of the phylogeny, supports previous hypotheses that the ancestral habitat likely included both freshwater and marine environments.

Acknowledgements

Many thanks to my committee member Lynne Parenti, whose advice, support, and help with analyses has been invaluable from the development of this chapter onward. Thanks also to my advisor Eric Hilton for helpful comments and advice on the chapter, and to the rest of my committee, John Graves, Casey Dillman, and Emmett Duffy. Funding sources included National Science Foundation (DEB 0841691) and the GK-12 Program.

References

Abalev A.G., Taschi S.M., and Safarova S.A. 2002. New Data on the Cenozoic History of the Underwater Yamato Rise, Sea of Japan. Russian Journal of Marine Biology, 28(3): 208-212.

Alden W. C., 1958, Physiography and Glacial Geology of Eastern Montana and Adjacent Areas. U. S. Geological Survey Professional Paper 174.

Algeo T.J., Kuwahara K., Sano H., Bates S., Lyons S., Elswick E., Hinnov L., Ellwood B., Moser J., and Maynard J.B. 2011. Spatial variation in sediment fluxes, redox conditions, and productivity in the Permian–Triassic Panthalassic Ocean. Palaeogeography, Palaeoclimatology, Palaeoecology, 308(1-2): 65-83.

Andrén, T., Lindeberg, G. & Andrén, E., 2002: Evidence of the final drainage of the Baltic Ice Lake and the brackish phase of the Yoldia Sea in glacial varves from the Baltic Sea. Boreas 31, 226–238

Arnason U., Gullberg A., Janke A., Kullberg M., Lehman N., Petrov E.A., and Väinölä R. 2006. Pinniped phylogeny and a new hypothesis for their origin and dispersal. Molecular Phylogenetics and Evolution, 41(2): 345-354.

Artyukhin E. N. 1995. On biogeography and relationships within the genus Acipenser. Sturgeon Quarterly, 3(2): 6–8.Artyukhin, E. N. 2006. Morphological Phylogeny of the Order Acipenseriformes. Journal of Applied Ichthyology, 22: 66–69.

Artyukhin, E. N. 2006. Morphological Phylogeny of the Order Acipenseriformes. Journal of Applied Ichthyology, 22: 66–69.

Bagils R.Z., Ung V., Grand A., Vignes-Lebbe R., Cao N., and Ducasse J. 2012. LisBeth: New cladistics for phylogenetics and biogeography. Comptes Rendus Palevol, 11: 563-566.

Bartok P. 1993. Prebreakup geology of the Gulf of Mexico-Caribbean: It's relation to Triassic and Jurassic Rift Systems of the Region. Tectonics, 12(2): 441-459.

Bemis W.E. and Kynard B. 1997. Sturgeon rivers: an introduction to acipenseriform biogeography and life history. Environmental Biology of Fishes, 48(1): 167-183.

Berglund B.E., Sandgren P., Barnekow L., Hannon G., Jiang H., and Skog G. 2005. Early Holocene history of the Baltic Sea, as reflected in coastal sediments in Blekinge, southeastern Sweden. Quaternary International, 130(1): 111-139.

Birstein V.J. & DeSalle R. 1998. Molecular phylogeny of Acipenserinae. Molecular Phylogenetics and Evolution, 9(1): 141-155.

Birstein V. J., Doukakis P., DeSalle R., 2002. Molecular phylogeny of Acipenseridae: Nonmonophyly of Scaphirhynchinae. Copeia, 2: 287–301.

Brigham-Grette J. 2001. New perspectives on Beringian Quaternary paleogeography, stratigraphy, and glacial history. Quaternary Science Reviews, 20: 15-24.

Choudhury A. and Dick T.A. 1998. Special Paper: The Historical Biogeography of Sturgeons (Osteichthyes: Acipenseridae): A Synthesis of Phylogenetics, Palaeontology and Palaeogeography. Journal of Biogeography, 25(4): 623-640.

Clark D. L. 19.74. Late Mesozoic and early Cenozoic sediment cores from the Arctic Ocean. Geology, 2: 41-44.

Croizat L. 1958. Panbiogeography or an introductory synthesis of zoogeography, phytogeography, and geology; with notes on evolution, systematics, ecology, anthropology, etc. Vol. 1—The New World, Vol. 2a- The Old World, Vol. 2b- The Old World. L Croizat: Caracas.

Croizat, L., G. Nelson and D. E. Rosen. 1974. Centers of origin and related concepts. *Systematic Zoology* 23: 265-87.

Dillman C. B., Wood R. M., Kuhajda B. R., Ray J. M., Salnikov V. B., Mayden R. L. 2007. Molecular systematics of Scaphirhynchinae: an assessment of North American and Central Asian Freshwater Sturgeon Species. Journal of Applied Ichthyology, 23: 290–296.

Docker M.F., Youson J.H., Beamish R.J., and Devlin R.H. 1999. Phylogeny of the lamprey genus Lampetra inferred from mitochondrial cytochrome *b* and ND3 gene sequences. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2340-2349.

Elias S.A., Short S.K., Nelson C.H., and Birks H.H. 1996. Life and times of the Bering land bridge. Nature, 382: 60 – 63.

Engen O., Faleide1 J.I., Dyreng T.K. 2008. Opening of the Fram Strait gateway: A review of plate tectonic constraints. Tectonophysices, 450(1-4): 51-69.

Fairhead J.D., 1988. Mesozoic plate tectonic reconstructions of the central South Atlantic Ocean: The role of the West and Central African rift system. Tectonophysics, 155: 181- 191.

Farrand W.R. 1967. The Glacial Lakes around Michigan. Geological Survey Division- Michigan Department of Environmental Quality, Bulletin 4.

Findeis, E.K. 1997. Osteology and phylogenetic interrelationships of sturgeons (Acipenseridae).Environmental Biology of Fishes, 48(1-4): 73-126.

Froese, R. and D. Pauly. Editors. 2016. FishBase. World Wide Web electronic publication. www.fishbase.org.

Gallagher S.J., Kitamura A., Iryu Y., Itaki T., Koizumi I. and Hoiles P.W. 2015. The Pliocene to recent history of the Kuroshio and Tsushima Currents: a multi-proxy approach. Progress in Earth and Planetary Science, 2:17.

Gill H.S., Renaud C.B., Chapleau F., Mayden R.L., and Potter I.C. 2003. Phylogeny of Living Parasitic Lampreys (Petromyzontiformes) Based on Morphological Data. Copeia, 4: 687-703.

Golonka J. 2007. Late Triassic and Early Jurassic palaeogeography of the world. Palaeogeography, Palaeoclimatology, Palaeoecology, 244(1-4): 297-307.

Grande L., Hilton E.J. 2006. An exquisitely preserved skeleton representing a primitive sturgeonfrom the Upper Cretaceous Judith River Formation of Montana (Acipenseriformes: Acipenseridae). Journal of Paleontology, 80: 672-683.

Hakanson L., Andersson M., and Ryden L. 2002. The Baltic Sea Basin - Nature, History and Economy. In: Ryden L., Migula P. and Andersson M. Environmental Science. The Baltic University Press. pp 92-119.

Hay W.W., Eicher D. L., and Diner R. 1993. Physical oceanography and water masses in the Cretaceous Western Interior Seaway. Geological Association of Canada Special Papers, 39: 297-318.

Hilton E.J., Grande L., Bemis W.E. 2011. Skeletal anatomy of the shortnose sturgeon *Acipenserbrevirostrum* Lesueur, 1818, and the systematics of sturgeons (Acipenseriformes, Acipenseridae). Fieldiana Life and Earth Sciences, 3: 1-168.

Hoagstrom C.W., Ung V., and Taylor K. 2014. Miocene rivers and taxon cycles clarify the comparative biogeography of North American highland fishes. Journal of Biogeography, 41(4): 644-658.

Holzkamm, T.E. and L.G. Waisberg. 2004. Native American utilization of sturgeon. In G.T.O.LeBreton, F.W.H. Beamish and R.S. McKinley (eds.). Sturgeon and Paddlefish of NorthAmerica. Kluwer Academic Publishers, Dordreht. Fish and Fisheries Series 27: 22–39.

Hopkins D.M. 1973. Sea level history in Beringia during the past 250,000 years. Quarternary Research, 3(4): 520-540.

IUCN. 2014. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>.

Jakobsson M., Backman J., Rudels B., Nycander J., Frank M., Mayer L., Jokat W., Sangiorgj F., O'Regan M., Brinkhuis H., King J., and Moran K. 2007. The early Miocene onset of a ventilated circulation regime in the Arctic Ocean. Nature, 447: 986–990.

Kender S., Stephenson M.H., Riding J.B., Leng M.J., Knox R.W., Peck V.L., Kendrick C.P., Ellis M.A., Vane C.H., and Jamieson R. 2012. Marine and terrestrial environmental changes in NW Europe preceding carbon release at the Paleocene–Eocene transition. Earth and Planetary Science Letters, 353-354: 108-120.

Konradi P.B., Larsen B., and Sorensen A.B. 2005. Marine Eemian in the Danish eastern North Sea. Quaternary International, 133-134: 21-31.

Krieger J., Hett A. K., Fuerst P. A., Artyukhin E., Ludwig A. 2008. The molecular phylogeny of theorder Acipenseriformes revisited. Journal of Applied Ichthyology, 24 (supplement 1): 36–45.

Lane C.E., Mayes C., Druehl L.D., and Sauders G.W. 2006. A multi-gene molecular investigation of the kelp (Laminariales, Phaeophyceae) supports substantial taxonomic re-organization. Journal of Phycology, 42: 493-512.

Larson G. and Schaetzl R. 2001. Origin and Evolution of the Great Lakes. Journal of Great Lakes Research, 27(4): 518-546.

Lee E.Y. 2010. Subsidence history of the Gunsan Basin (Cretaceous-Cenozoic) in the Yellow Sea, offshore Korea. Austrian Journal of Earth Sciences, 103(1): 111-120.

Lepparanta M. and Myrberg K. 2009. The Baltic Sea: History and Geography. In: Lepparanta M. and Myrberg K., Geophysical Sciences: Physical Oceanography of the Baltic Sea. Springer-Verlag Berlin Heidelberg. pp 9-40.

Leverett F. 1921. Outline of Pleistocene History of Mississippi Valley. The Journal of Geology, 29(7): 615-626.

Liu J.P. and Milliman J.D., 2004. Reconsidering Melt-water Pulses 1A and 1B: Global Impacts of Rapid Sea-level Rise. Journal of Ocean University of China, 3(3): 183-190.

Ludwig A., Debus L., Lieckfeldt D., Wirgin I., Benecke N., Jenneckens I., Williot P., Waldman J.R., and Pitra C. 2002. Fish populations: When the American sea sturgeon swam east. Nature, 419: 447-448.

Marincovich, L. and Jr., Gladenkov, A.Yu., 2001. New evidence for the age of Bering Strait. Quaternary Science, 20:329-335.

Mather K.F. 1917. The Champlain Sea in the Lake Ontario Basin. The Journal of Geology, 25(6): 542-554.

McHone J.G., 2000. Non-plume magmatism and rifting during the opening of the central Atlantic Ocean. Tectonophysics 316, 287–296.

Miettinen A. 2003. Quaternary sea level highstands in the Eastern Baltic area. GI2S Coast Research Publication. Available at: <http://www.igcp437finalconference.unisalento.it/AbstractBook/Miettinen171_172.pdf>

Milinkovitch M.C., Orti G., and Meyer A. 1995. Novel Phylogeny of Whales Revisited but not Revised. Molecular Biology and Evolution, 12(3): 518-520.

Moran K., Backman J., Brinkhuis H., Clemens S.C., Cronin T., Dickens G.R., Eynaud F., Gattacceca J., Jakobsson M., Jordan R.W., Kaminski M., King J., Koc N., Krylov A., Martinez N., Matthiessen J., McInroy D., Moore T.C., Onodera J., O'Regan M., Pälike H., Rea B., Rio D., Sakamoto T., Smith D.C., Stein R., St John K., Suto I., Suzuki N., Takahashi K., Watanabe M., Yamamoto M., Farrell J., Frank M., Kubik P., Jokat W., and] Kristoffersen Y. 2006. The Cenozoic palaeoenvironment of the Arctic Ocean. Nature, 441: 601 – 605.

Morner N.A. 1995. Earth Rotation, ocean circulation and paleoclimate. GeoJournal, 37(4): 419-430.

Neall V.E. and Trewick S.A. 2008. The age and origin of the Pacific Islands: a geological overview. Philosophical Transactions of the Royal Society B, 363:3293-3308.

Nelson G. and Ladiges P.Y. 1992. Information content and fractional weight of three-item statements. Systematic Biology, 41: 490–494.

Nelson G. and Ladiges P.Y. 1996. Paralogy in cladistic biogeography and analysis of paralogy-free subtrees. American Museum novitates, 3167.

Nelson G. and Ladiges P.Y. 2001. Gondwana, vicariance biogeography and the New York School revisited. Australian Journal of Botany, 49(3): 389-409.

Nelson, G. and N. I. Platnick. 1981. Systematics and biogeography: Cladistics and vicariance. Columbia Univ. Press, New York.

O'Leary N., White N., Tull S., Bashilov V., Kuprin V., Natapov L., and MacDonald D. 2004. Evolution of the Timan–Pechora and South Barents Sea basins. Geological Magazine, 141(2): 141-160.

Parenti L. 2008. Life History Patterns and Biogeography: An Interpretation of Diadromy in Fishes. Annals of the Missouri Botanical Garden, 95(2): 232-247.

Parenti L. and Ebach M. 2009. Comparative Biogeography: Discovering and Classifying Biogeographical Patterns of a Dynamic Earth.

Peng Z., Ludwig A., Wang D., Diogo R., Wei Q., and He S. 2007. Age and biogeography of major clades in sturgeons and paddlefishes (Pisces: Acipenseriformes). Molecular Phylogenetics and Evolution, 42: 854–862. University of California Press.

Posadas P., Crisci J.V., and Katinas L. 2006. Historical biogeography: A review of its basic concepts and critical issues. Journal of Arid Environments, 66(3): 389-403.

Price S.A., Bininda-Edmonds O.R.P., and Gittleman J.L. 2005. A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). Biological Reviews, 80(3): 445-473.

Rogl F. 1999. Mediterranean and Paratethys Palaeogeography during the Oligocene and Miocene. In: Agusti J., Rook L., Andrews P. Hominoid Evolution and Climatic Change in Europe, Volume 1. The Evolution of Neogene Terrestrial Ecosystems in Europe. Cambridge University Press. pp 8-22.

Russell H.A.J. and Cummings D.I. 2009. Deglaciation of the Champlain Sea Basin, Eastern Ontario. 72nd Friends of the Pleistocene Reunion, Geological Survey of Canada.

Sandgren P., Subetto D.A., Berglund B.E., Davydova N.N., Savelieva L.A. 2004. Mid-Holocene Baltic Sea transgressions and their climatic implications, based on stratigraphic studies in coastal lakes of NW Russia. GFF Vol.126. pp.363-380

Schlee J.S. 1999. Our Changing Continent. United States Department of the Interior Geological Survey. Unnumbered Report. Available at: <http://pubs.usgs.gov/unnumbered/70039574/report.pdf>

Shepherd, A, E R Ivins, G A, V R Barletta, M J Bentley, S Bettadpur, K H Briggs. 2012. A Reconciled Estimate of Ice-Sheet Mass Balance. Science 338 (6111): 1183–89. doi:10.1126/science.1228102.

Siegert, M. J., J. C. Ellis-Evans, and M. Tranter et al. (2001), Physical, chemical and biological processes in Lake Vostok and other Antarctic subglacial lakes, Nature, 414(6864): 603 – 609

Sulak K.J. and Randall M. 2002. Understanding sturgeon life history: Enigmas, myths, and insights from scientific studies. Journal of Applied Ichthyology, 18(4-6): 519-528.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tada, R., Murray, R.W., and Alvarez Zarikian, C.A., 2013. Asian Monsoon: onset and evolution of millennial-scale variability of Asian Monsoon and its possible relation with Himalaya and Tibetan plateau uplift. IODP Sci. Prosp., 346. doi:10.2204/iodp.sp.346.2013

Tapponnier P., Peltzer G., and Armijo R. 1986. On the mechanics of the collision between India and Asia. Geological Society London, Special Publications, 19: 113-157.

Thiede, J. (1979) History of the North Atlantic Ocean: Evolution of an Asymmetric Zonal Paleo-Environment in a Latitudinal Ocean Basin, in Deep Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironment (eds M. Talwani, W. Hay and W. B. F. Ryan), American Geophysical Union, Washington, D. C.

Thompson B., Nilsson J., Nycander J., Jakobsson M., and Döös K. 2010. Ventilation of the Miocene Arctic Ocean: An idealized model study. Paleoceanography, 25(4).

Torsvik T.H., Carlos D., Mosar J., Cocks L.R.M. and Malme T., 2002. Global reconstructions and North Atlantic palaeogeography 400 Ma to Recent. In: Eide E.A., BATLAS – Mid Norway plate reconstructions atlas with global and Atlantic perspectives. Geological Survey of Norway. pp 18-39.

Trimble D.E. 1980. The Geologic Story of The Great Plains. Geological Survey Bulletin, 1493.

Uda M. 2016. Sea of Japan. *Encyclopædia Britannica: Encyclopædia Britannica Online.* Encyclopædia Britannica Inc., Web: 01 Jan. 2016.

University of Wisconsin Sea Grant Institute (UWSG) 2013. Great Lakes: The Formation of the Great Lakes: How They Were Made. Available at: <http://seagrant.wisc.edu/home/Default.aspx?tabid=510>

Vasil'eva E.D., Vasil'ev V.P., Shedko S.V., Novomodny G.V. 2009. The Revision of the Validity of Genus Huso (Acipenseridae) Based on Recent Morphological and Genetic Data with Particular Reference to the Kaluga H. dauricus*. Journal of Ichthyology, 49(10):861-867.

Vassiljev J. and Saarse L/ 2013. Timing of the Baltic Ice Lake in the eastern Baltic. Bulletin of the Geological Society of Finland, 85: 9-18.

Walday M. and Krogland T. 2011. Seas around Europe: The Baltic Sea. European Environment Agency Publication. Available at: <http://www.eea.europa.eu/publications/report_2002_0524_154909/regional-seas-around-europe/page141.html>

Weninger, B., Schulting, R., Bradtmoller, M., Clare, L., Collard, M., Edinborough, K., Hilpert, J., Jo¨ ris, O., Niekus, M., Rohling, E.J., Wagner, B., 2008. The catastrophic final flooding of Doggerland by the Storegga Slide tsunami. Documenta Prehistorica XXXV, 1–24.

Wickert A.D., Mitrovica J.X., Williams C., and Anderson R.S. 2013. Gradual demise of a thin southern Laurentide ice sheet recorded by Mississippi drainage. Nature, 502: 668-671.

Yang Z., Li Z., Qu L., Lu Z., Zhou H., Zhou T., Liu G., Liu B., and Wu. 1982. Acta Geologica Sinica. Translated by W. Downs. Available at: <http://paleoglot.org/files/Yang&_82.pdf>

Yi S., Yi S., Batten D.J., Yun H., and Park S. 2003. Cretaceous and Cenozoic non-marine deposits of the Northern South Yellow Sea Basin, offshore western Korea: palynostratigraphy and palaeoenvironments. Palaeogeography, Palaeoclimatology, Palaeoecology, 191(1): 15-44.

Ziegler P. A. 1975. The geological evolution of the North Sea area in the tectonic framework of North-Western Europe. Norges Geologiske Undersokelse. 316: 1-27.

Zonenshain L.P. and Pichon X. 1986. Deep basins of the Black Sea and Caspian Sea as remnants of Mesozoic back-arc basins. Tectonophysics, 123(1-4): 181-211.

Tables

Table 1: Sturgeon species ranges. Ranges of sturgeon species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

	JPN	CHN	YLW	PNA	AES	BNS	PTC	ANA	MGM	NAL	BRG
Acipenser fulvescens									√	√	
Acipenser baerii					√						
Acipenser brevirostrum								√			
Acipenser dabryanus		√									
Acipenser gueldenstaedtii							√				
Acipenser medirostris				√							
Acipenser mikadoi	√				√						√
Acipenser naccarii							√				
Acipenser oxyrinchus desotoi									√		
Acipenser oxyrinchus oxyrinchus								√		√	
Acipenser ruthenus					√		√				
Acipenser schrenckii	√										
Acipenser sinensis		√	√								
Acipenser stellatus					√		√				
Acipenser sturio						√	√				
Acipenser transmontanus				√							
Huso dauricus	√										
Huso huso							√				
Pseudoscaphirhynchus kaufmanni							√				
Scaphirhynchus albus									√		
Scaphirhynchus platorynchus									√		
Scaphirhynchus suttkusi									√		

Table 2: Lamprey species ranges. Ranges of lamprey species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

	JPN	CHN	YLW	PNA	AES	BNS	PTC	ANA	MGM	NAL	BRG
Caspiomyzon wagneri							√				
Entosphenus macrostomus				√							
Entosphenus minimus				√							
Entosphenus similis				√							
Entosphenus tridentatus				√							√
Eudontomyzon danfordi							√				
Eudontomyzon morii			√								
Ichthyomyzon unicuspis									√	√	
Ichthyomyzon bdellium									√		
Ichthyomyzon casteneus									√	√	
Lampetra ayresii				√							
Lampetra fluviatilis						√	√				
Lethenteron camtschaticum	√			√	√						√
Petromyzon marinus					√		√	√	√	√	
Tetrapleurodon spadiceus				√							

Table 3: Pinniped species ranges. Ranges of pinniped species in areas of endemism. JPN=Japan Region, CHN=China Region, YLW=Yellow Sea Region, PNA=Pacific Coast of North America Region, AES=Arctic Region, BNS=Baltic and North Sea Region, PTC=Ponto-Caspian Region, ANA=Atlantic Coast of North America Region, MGM=Gulf of Mexico Region, NAL=North American Lakes Region, BRG=Bering Sea Region

	JPN	CHN	YLW	PNA	BNS	PTC	ANA	MGM	NAL	AES	BRG
Monachus monachus						X					
Arctocephalus townsendi				X							
Callorhinus ursinus	X			X							X
Cystophora cristata					X		X		X		
Erignathus barbatus	X								X		X
Eumetopias jubatus	X			X							X
Halichoerus grypus					X		X		X		
Histriophoca fasciata	X									X	X
Mirounga angustirostris				X							
Odobenus rosmarus										X	X
Pagophilus groenlandicus							X		X	X	
Phoca largha	X	X									
Phoca vitulina	X			X			X		X	X	
Pusa caspica				X							
Pusa hispida	X	X			X		X		X	X	X
Zalophus californianus				X							

Figures

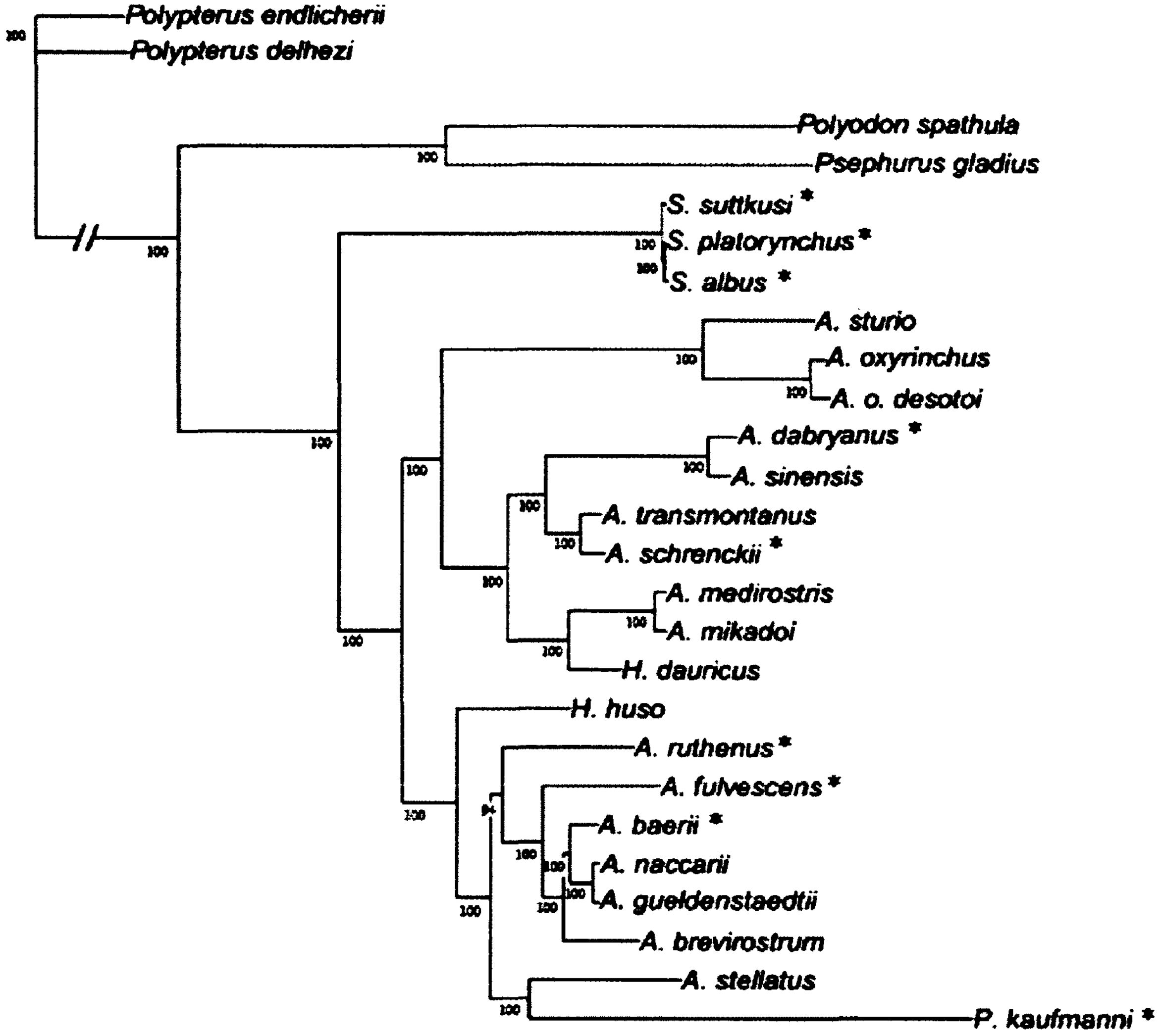


Figure 1: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships. Exclusively freshwater taxa are indicated by an asterisk (*).

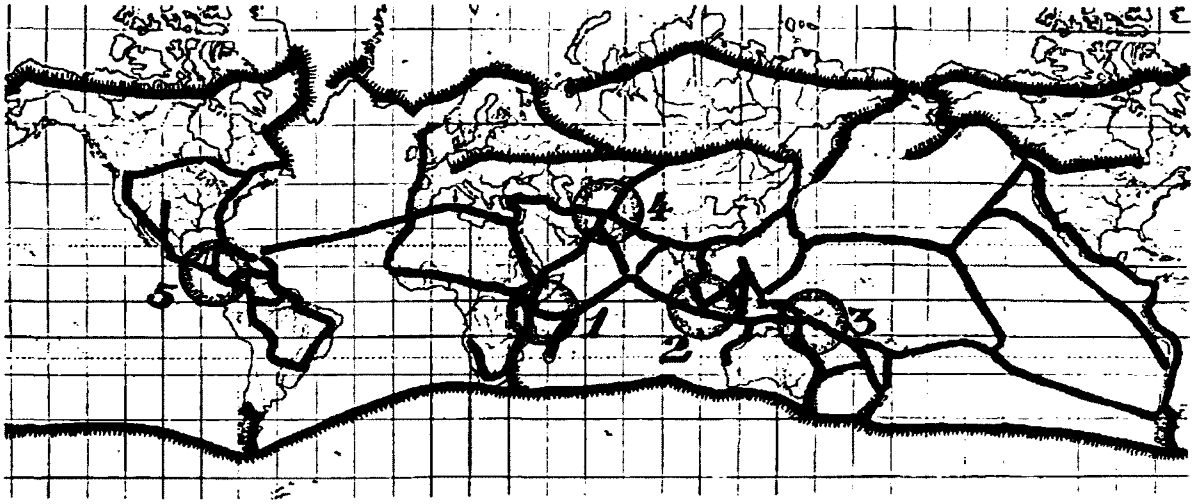


Figure 2: Croizat's global biogeography map. Austral and Boreal regions are depicted with hatched lines.

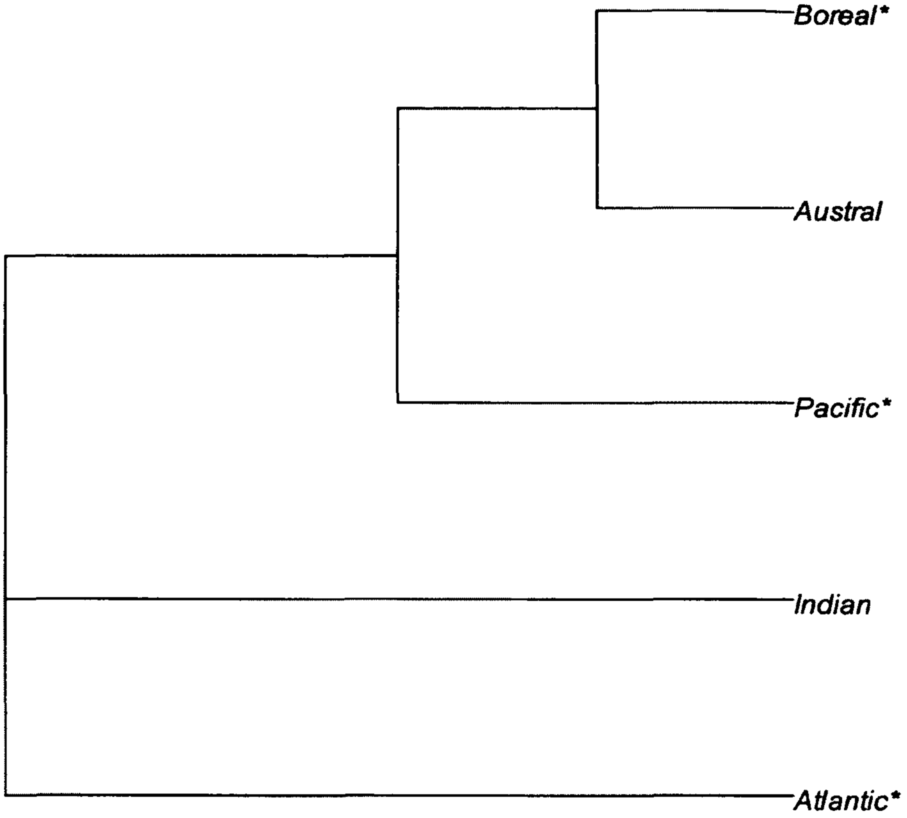


Figure 3: Areagram of Croizat's (1958) biogeographic regions following Nelson and Ladiges (2001) and Parenti (2008). Asterisks indicate regions inhabited by sturgeons.

Figure 4: Map of sturgeon Areas of Endemism; Regions are abbreviated as follows BNS: Baltic and North Sea, PTC: Ponto-Caspian, JPN: Japan, YLW: Yellow Sea, CHN: China, AES: Arctic, BRG: Bering Sea, PNA: Pacific North America, MGM: Gulf of Mexico, NAL: North American Lakes, and ANA: Atlantic North America.

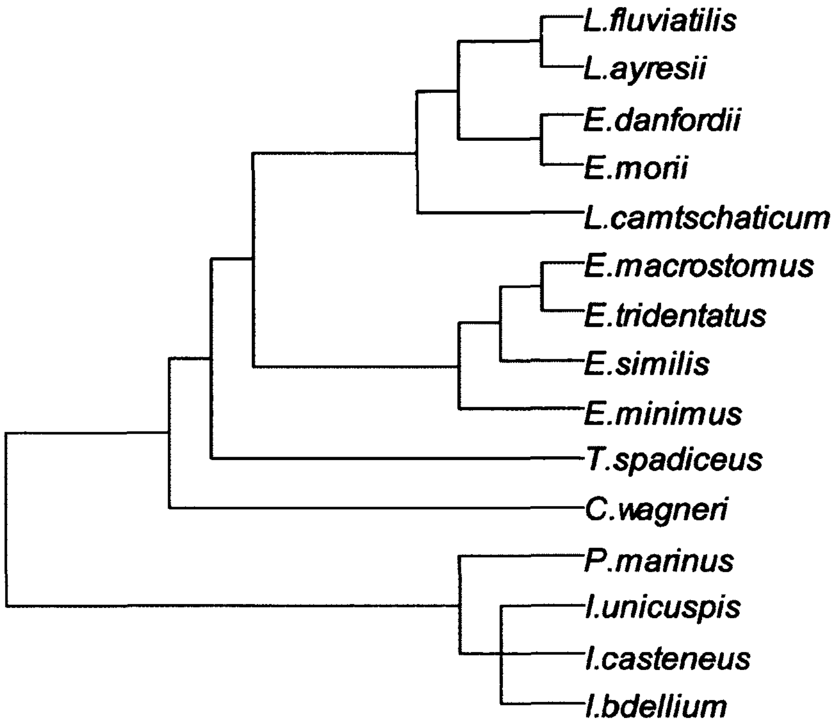


Figure 5: Morphological Lamprey Phylogeny modified from Gill et al. 2003; species that do not share area of endemism with sturgeons were pruned from the original topology.

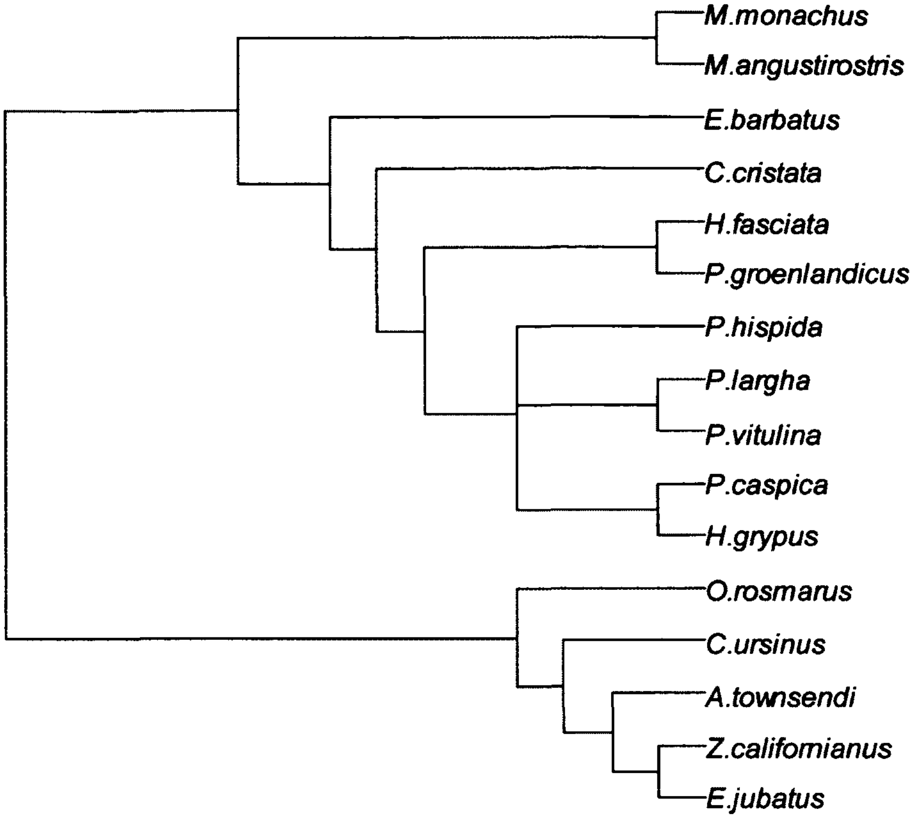
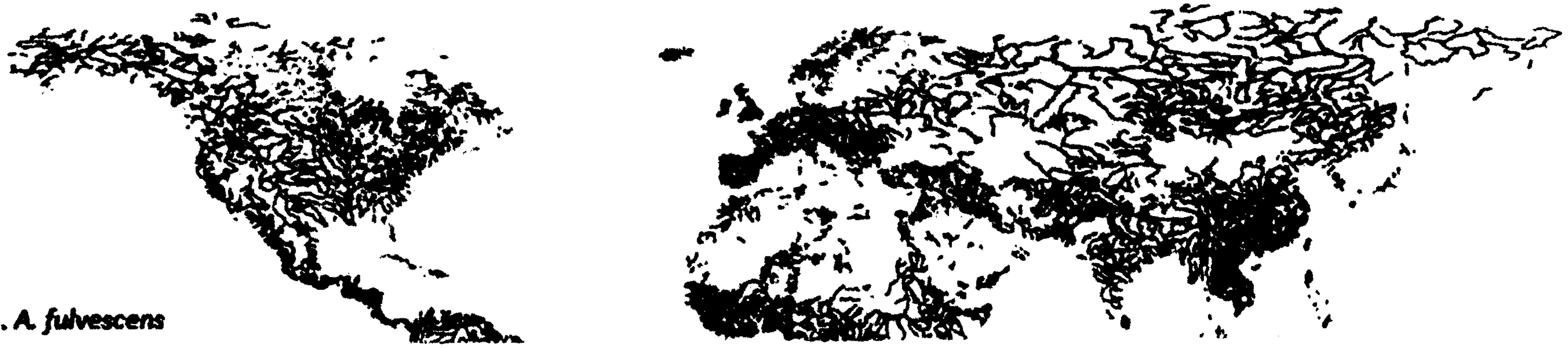
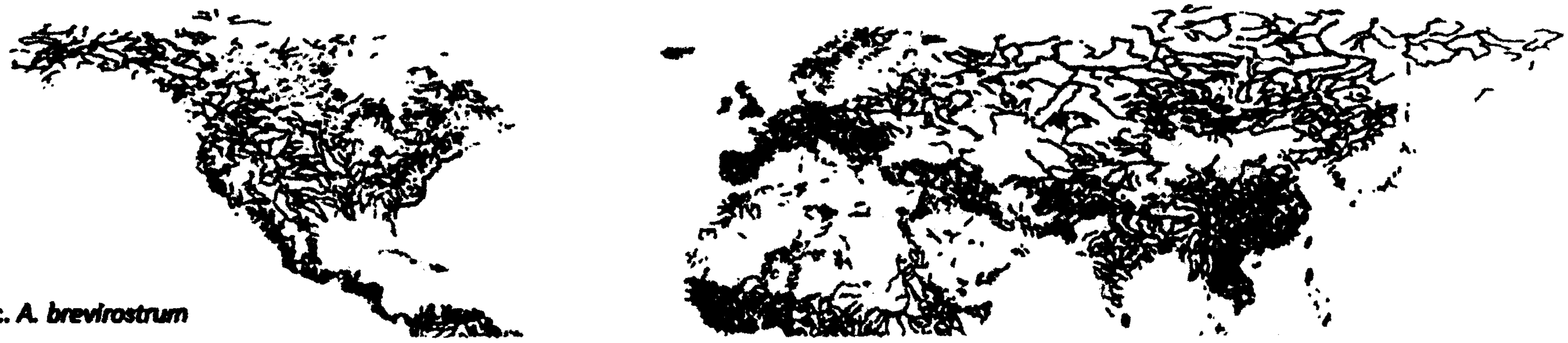


Figure 6: Molecular Pinniped Phylogeny modified from Arnason et al. 2006; species that do not share area of endemism with sturgeons were pruned from the original topology.

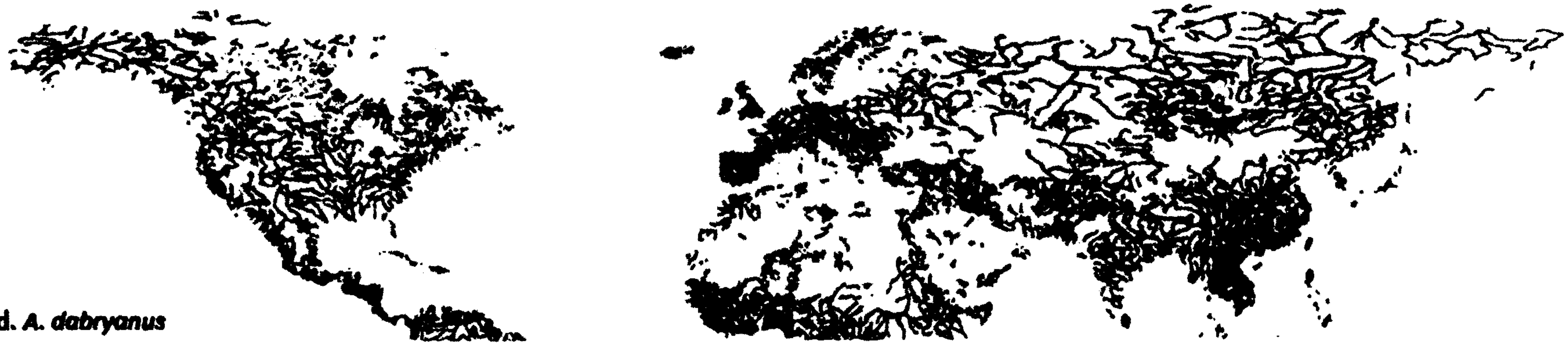


7a. *A. fulvescens*

7b. *A. baerii*



7c. *A. brevirostrum*



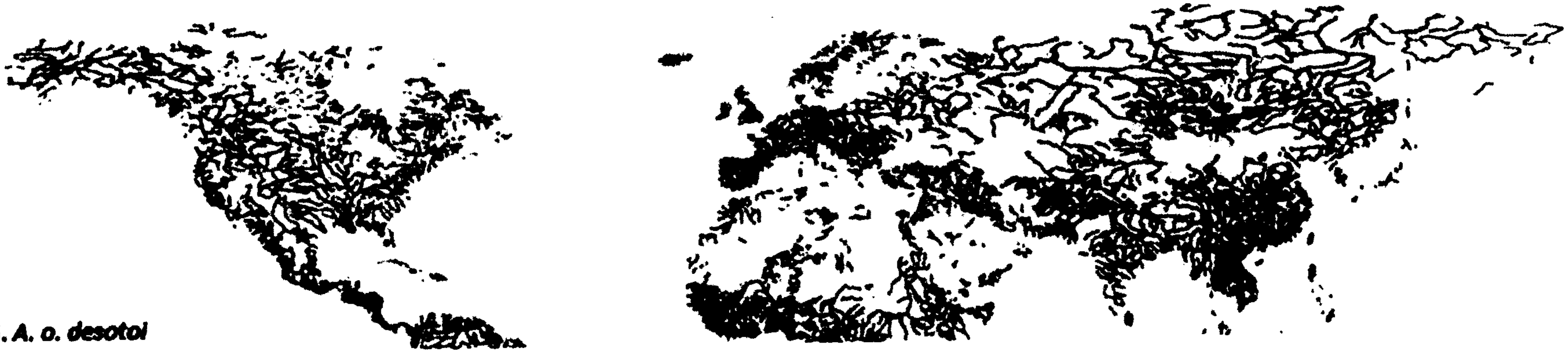
7d. *A. dabryanus*

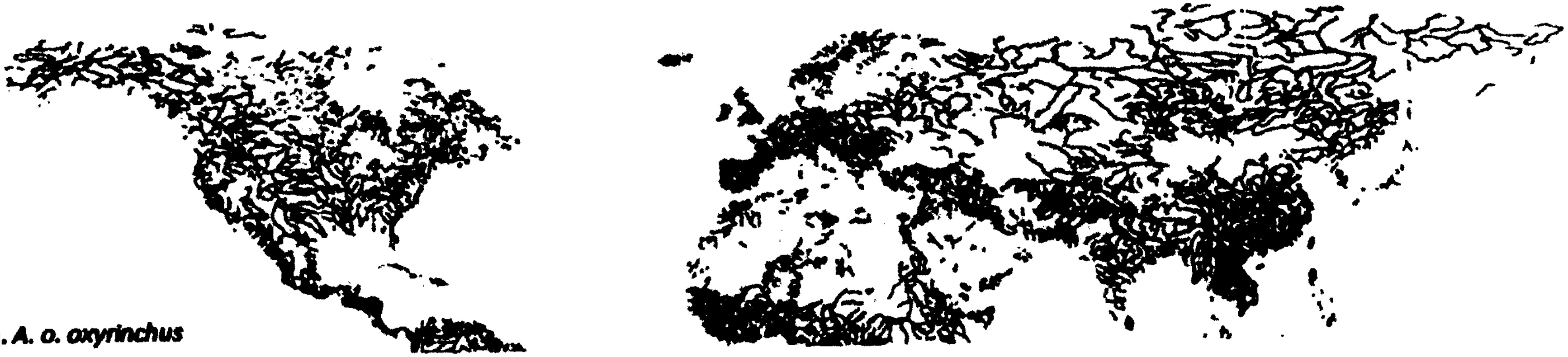
7e. A. gueldenstaedtii

7f. A. medirostris

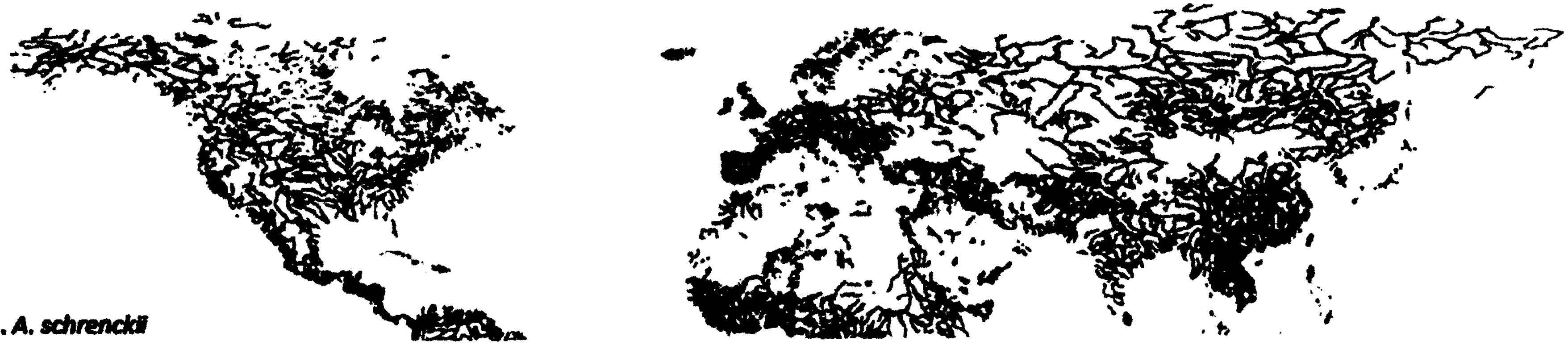
7g. A. mikadoi

7h. A. naccarii

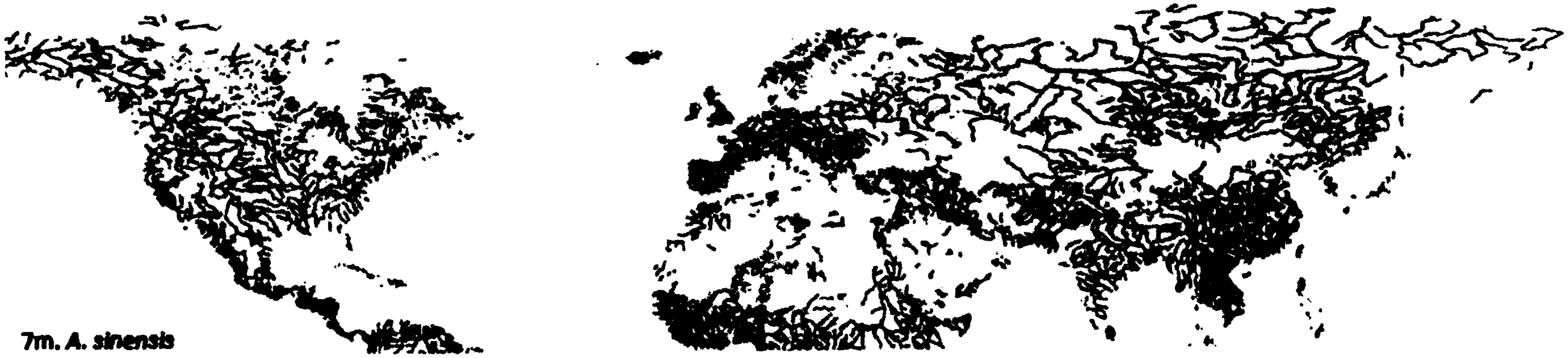

7i. *A. o. desotoi*


7j. *A. o. oxyrinchus*

7k. A. ruthenus

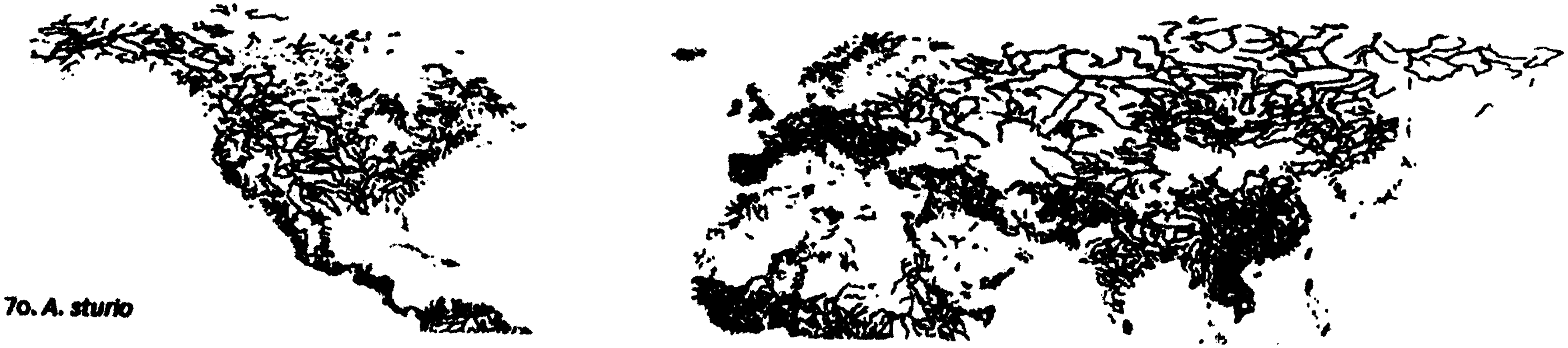


7l. A. schrenckii

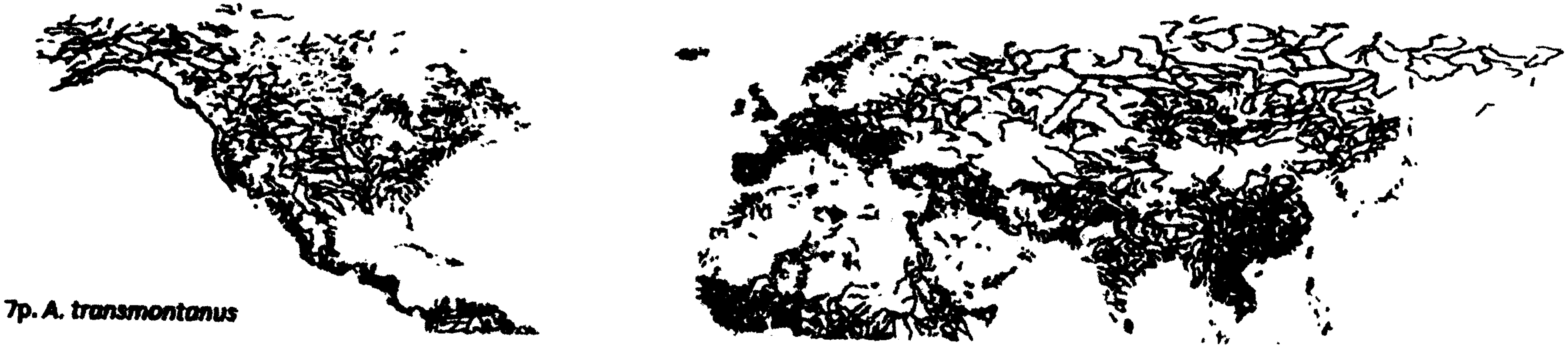


7m. A. sinensis

7n. A. stellatus



7o. *A. sturio*



7p. *A. transmontanus*

7q. *H. dauricus*

7r. *H. huso*

7s. *P. kaufmanni*

7t. *S. albus*

7u. S. platorhynchus

7v. S. suttkusi

Figure 7: Range maps of sturgeon species included in biogeographic analyses. Red indicates areas inhabited by each species.

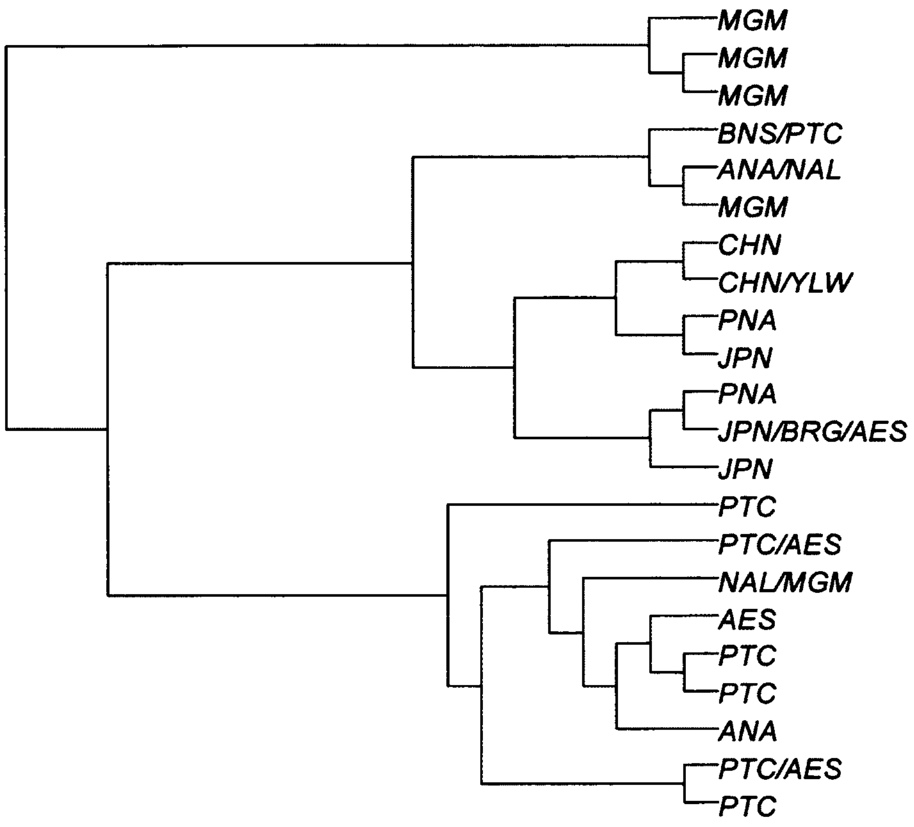


Figure 8: Sturgeon Areagram; species names are replaced on the full mitogenome phylogeny with areas of endemism

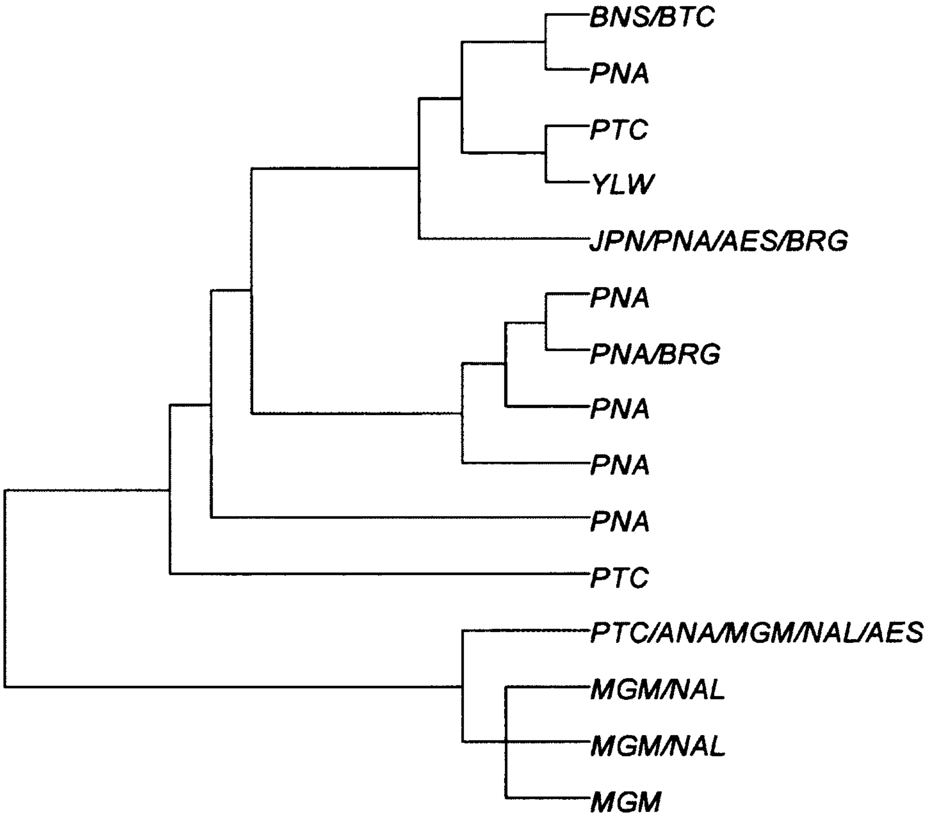


Figure 9: Lamprey Areagram. Species names are replaced on the pruned Gill *et al.* phylogeny with areas of endemism.

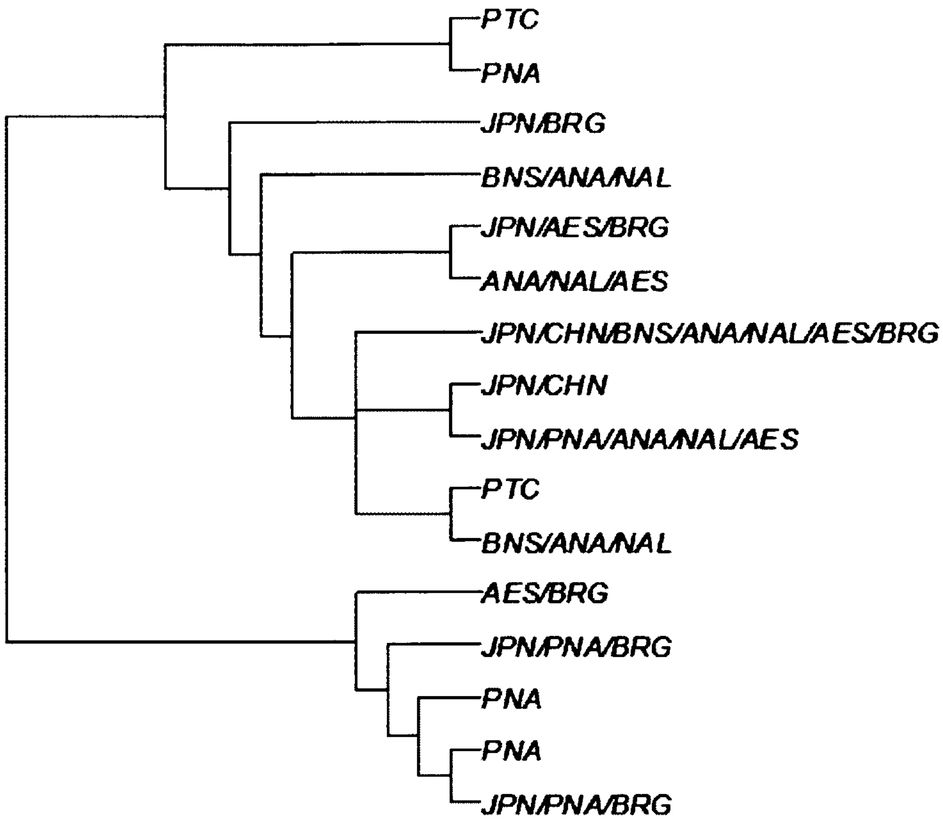


Figure 10: Pinniped Areagram; species names are replaced on the pruned Arnason et al. 2006 phylogeny with areas of endemism

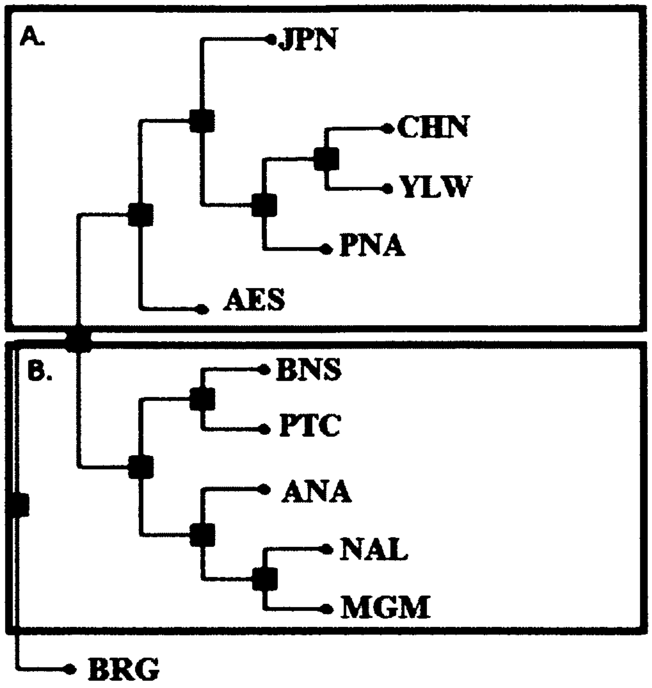


Figure 11: Sturgeon-Lamprey-Pinniped Intersection Tree showing area relationships for sturgeon, lamprey, and pinniped areas of endemism. Box A indicates the primarily Pacific clade, while box B indicates the Atlantic-Arctic clade. The completeness index 92.7%.

CHAPTER 3.

Utility of Ontogenetic Characters in Phylogenetic Studies

Abstract

Form, which is represented by morphology, and function, which is exemplified through behavior, are said to be inextricably linked. Sturgeons (Acipenseridae) exhibit dramatic morphological and behavioral shifts during ontogeny: developing then resorbing teeth and beginning their exogenous feeding behavior by swimming up into the water column before eventually becoming benthic feeders. Researchers often use behavioral and ontogenetic characters in the study of phylogeny, but the timing of morphological and behavioral shifts in sturgeons has not been investigated for phylogenetic signal. The primary goal of this chapter is to identify links between morphological and behavioral ontogeny, and to determine whether the timing of ontogenetic changes in their morphology and behavior carries phylogenetic signal. Correlation is found between one set of behavioral-morphological characters, the age at which migration of early life stage sturgeons is completed and the length at which tooth development is complete. One ontogenetic character, the age at which teeth are completely resorbed, shows phylogenetic signal.

1. Introduction

It is widely understood that form (morphology) and function (exemplified in behavior) are often inextricably linked (e.g. Russell 1982, Arnold 1983, Zelditch *et al.* 2004, Bertossa 2011). Comparative analyses among species, which have traditionally been used to test for correlations between morphological traits (Garland *et al.* 1992), may also be used to test for links between form (morphological traits) and function (behavioral traits) (Jordana *et al.* 1999). Traditional comparative methods, such as Pearson product moment correlations, assume that trait data under consideration are not statistically linked and that all taxa included in analysis are equally related to one another (Felsenstein 1985, Garland et al. 2005). However, the hierarchical nature of species relationships violates this assumption; traits are linked through the shared evolutionary history of their taxa (Garland et al. 2005). To remove the confounding factor of hierarchical relationships from comparative analyses and avoid making an assumption that is sure to be violated, phylogenetic relationships must be taken into account in comparative analyses. Phylogenetic comparative methods such as phylogenetically independent contrasts (PIC) use phylogenies to control for the confounding phylogenetic relationships. This allows for the straightforward testing for correlations among species traits. The 25 species of sturgeon, distributed amongst the genera *Acipenser, Scaphirhynchus, Pseudoscaphirhynchus,* and *Huso,* exhibit behavioral and morphological changes during development that may be linked, and may exhibit phylogenetic signal.

Frequently described as "ancient" fishes (Inoue *et al.* 2003, Birstein 1993, Ludwig *et al.* 2007, Katsu *et al.* 2008), sturgeons exhibit dramatic morphological and behavioral shifts between hatching and maturity (Kirschbaum and Williot 2011, Nelson *et al.* 2013, Dang and

Zhang 2014). Many sturgeon species are anadromous, moving between fresh and saltwater during their lives (Bemis and Kynard 1997, Peng 2007). They spawn and hatch as free embryos in riverine systems, where they develop into larvae and begin foraging (Kynard and Parker 2005, Kynard 1997, Kynard and Horgan 2002, Kynard *et al.* 2005, Parker 2007, Kynard and Parker 2010). The anadromous species move into the ocean by adulthood, but return to rivers to spawn (Nelson *et al.* 2013). Sturgeon larval development includes major changes to jaw morphology, the appearance and proliferation of electroreceptors on the rostrum, and the development of bony scutes in rows along the sides of the body (although scute development continues after the larval stage) (Parker 2007). Sturgeons hatch with forward facing, subterminal jaws, which shift to a ventral orientation during early stages of development (i.e., before the juvenile stage). Oral jaws (subsequently referred to as jaws) in juvenile and adult sturgeons are highly protrusible, enabling suction feeding. The jaws of sturgeons are toothless at hatching, but teeth develop on the dentary, dermopalatines, and palatopterygoid during the larval stage. The teeth are subsequently resorbed and the jaws of juvenile and adult sturgeons are edentulous. It is likely that teeth assist in prey capture while the jaws are still terminally positioned (B. Kuhajda, personal communication, 2011).

Sturgeons also display shifts in behavior during larval and early juvenile life stages. The links between behavioral and morphological development in sturgeons have been the focus of a variety of studies, and morphological and behavioral changes during early life history stages of sturgeons are known to occur concurrently (Dang and Zhang 2014, Parker 2007, Gisbert and Williot 1999). Dang and Zhang (2014) found that behavioral ontogeny, in general, correlated with morphological development in *A. fulvescens*. Parker (2007) showed that the peak of migration in *A. brevirostrum* is linked with the completion of development of the pectoral and pelvic fins and the presence of a pre-anal fin fold. Gisbert and Williot (1999) showed that the beginning of free feeding in *A. baerii* is accompanied by changes in digestive enzymes present in the gut. The ability to perform behaviors such as swimming downstream requires that certain morphological structures be developed (e.g., fins; Parker 2007). It follows, then, that the timing of tooth and oral jaw development should correlate with changes in feeding behavior.

Several studies have documented early life history changes in sturgeons (e.g. Buckley and Kynard 1981, Kynard 1997, Kynard and Horgan 2002, Kynard *et al.* 2005, Parker 2007, Richmond and Kynard 1995). The goals of these studies included informing conservation and restoration efforts and describing various aspects of behavior (feeding behavior, habitat use, upstream and downstream movement, activity patterns at different time of day and different light levels night, and position in the water column). They have also shown that different sturgeon species exhibit different behaviors during the larval stage. For example, the larvae of different species lie on the substrate, seek cover in rocky habitat, or swim up above the bottom during foraging migrations (Parker 2007, Gisbert and Williot 1997). Behaviors that are shared among taxa in the larval stage may diverge as the species develop; most species first feed on plankton and eventually progress to a diet of benthic organisms, but some species (e.g., *S. albus*) may become piscivorous as adults (Grohs *et al.* 2009, Parker 2007).

Developmental studies have noted a general trend for post-hatch embryonic (=free embryo) *A. oxyrinchus* and *A. brevirostrum* to seek and maintain cover until the larval stage is reached, upon which the fish emerge to migrate and forage (Buckley and Kynard 1981, Kynard 1997, Kynard and Horgan 2002, Richmond and Kynard 1995). *Acipenser transmontanus,* however, begins a short migration immediately upon hatching (Kynard and Parker 2005 and 2010). This species migrates during the day and night, but exhibits nocturnal foraging habits (Kynard and Parker 2005, 2010). *Acipenser medirostris* cannot swim well upon hatching, but can move short distances to reach cover (Kynard *et al.* 2005). Like *A. oxyrinchus* and *A. brevirostrum, A. medirostris* emerges from cover-providing habitat at the larval stage and begins migrating downstream, with brief foraging periods during migration (Kynard *et al.* 2005). Conversely, free embryos and larvae of *A. medirostris* exhibit primarily nocturnal migration and foraging habits (Kynard *et al.* 2005). *Acipenser oxyrinchus* is nocturnal during the early stages of its 12-day migration period, and decidedly diurnal during the later migration period (Kynard and Horgan 2002). *Acipenser brevirostrum* migrates diurnally (Kynard and Horgan 2002).

Patterns of development of sturgeons, including behavioral and morphological ontogeny, may reflect phylogenetic signal. For example, migration and habitat preference have been used as characters in phylogenetic studies of sturgeons (Mayden and Kuhajda 1996). Some researchers have suggested that similarities in behavior in sturgeon development is the result of convergent evolution rather than shared evolutionary history (Kynard and Horgan 2002). Morphological characters have been used in phylogenetic studies of sturgeons (e.g. Findeis 1997; Artyukhin 2006, Hilton et al. 2011), but the timing of development of these characters has not been explored in the context of the phylogenetic history of sturgeons. The goals of this study are twofold. First, I test whether the timing of specific behavioral and morphological shifts carries phylogenetic signal. Results are discussed in terms of which characters carry phylogenetic signal matching that of a full mitogenome Bayesian phylogeny (Figure 1). Second, I test for correlation between the timing of morphological and behavioral ontogenetic shifts related to feeding structures (teeth and oral jaw elements, specifically Meckel's cartilage and cartilage of the palatoquadrate) and feeding, respectively.

2. Methods

2.1 Taxon Sampling and Data Collection

2.1.1 Morphological Characters

Developmental series of known-age specimens (from one to 65 days post-hatch) of *Acipenser transmontanus* (VIMS 17716), *A. oxyrinchus* (VIMS uncataloged),, *A. brevirostrum* (VIMS 12091; also VIMS 12076-12082, VIMS 12086, VIMS12088, VIMS 12089, VIMS, 12101),, *A. medirostris* (VIMS 17715),, and *A. fulvescens* (VIMS 13577), were obtained from hatcheries and collections; all specimens were deposited in the VIMS Ichthyology Collection. Specimens were preserved in 4% paraformaldehyde and transferred to 70% ethanol. Specimens were cleared and stained for bone (stained red) and cartilage (stained blue) based on a protocol modified from Dingerkus and Uhler (1977) (examples of cleared and stained specimens are shown in Figure 2). All specimens were photographed and total length (TL) measured using a Zeiss Discovery V20 microscope. Specimens were examined to identify the earliest appearance feeding-related structures (teeth and oral jaw elements) in each species. Additionally, the development of teeth was tracked, as was the resorption of the yolk sac. The earliest age (days posthatch) and shortest length (total length) at which elements were first observed was recorded for each species. Age at occurrence for some elements is unknown. This resulted in 10 continuous morphological ontogenetic characters:

1. Age at which first appearance of mandibular arch (jaw) elements (specifically Meckel's cartilage and cartilage of the palatoquadrate) occurs;

2. Length at which first appearance of mandibular arch elements occurs;
3. Age at which teeth first appear;
4. Length at which teeth first appear;
5. Age at which all teeth are completely developed, and after which teeth begin to resorb, becoming thinner, shorter, and fewer (this milestone is identified by the maximum counts and lengths of teeth);
6. Length at which all teeth are completely developed;
7. Age when teeth are completely resorbed, defined as the presence of tooth sockets but absence of mineralized tissue protruding from sockets;
8. Length at which teeth are completely resorbed;
9. Age at which the yolk sac is completely digested; and
10. Length at which the yolk sac is completely digested.

2.1.2 Behavioral Characters

Extensive ontogenetic behavioral data for five species were available from the literature: *A. brevirostrum* (Richmond and Kynard 1995, Kynard 1997, Kynard and Horgan 2002), *A. medirostris* (Kynard *et al.* 2005), *A. oxyrinchus* (Kynard and Horgan 2002), *A. fulvescens* (Smith and King 2005) and *A. transmontanus* (Kynard and Parker 2005, 2010). These data include information about habitat use, preference for illumination, foraging behavior, migratory behavior, preference for dark or light substrate, swimming height at different life stages, nocturnal and diurnal behavior, and timing of the onset of arval migration. Data for some of these characters were not available for *A. fulvescens* (see Smith and King 2005). Habitat use was characterized by the percent of fish using rocky cover, the open bottom, the water column, and the water surface on each day for *A. transmontanus, A. brevirostrum,* and *A. medirostris;* this has not been investigated for *A. oxyrinchus* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Preference for light or dark habitats was measured by the mean percent of fish in illuminated sections of a tank as opposed to covered, dark sections each day for *A. transmontanus, A. brevirostrum, A. medirostris,* and *A. oxyrinchus* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Preference for different color substrates was determined by the mean percent of fish on white, rather than black, substrate in *A. transmontanus, A. oxyrinchus,* and *A. medirostris* (see Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). It was measured as the percent of individuals on black, gray, and white substrates for *A. brevirostrum* (Richmond and Kynard 1995). Swimming height was calculated as the mean swimming height of fish, in centimeters, each day for *A. transmontanus* and *A. medirostris* (Kynard *et al.* 2005, Kynard and Parker 2005). For *A. brevirostrum,* swimming height was measured as the mean swimming height of fish introduced at the water surface and after swimming to the bottom (Kynard and Horgan 2002). Information regarding *A. oxyrinchus* was not available for this trait. Migratory behavior was determined for *A. transmontanus, A. brevirostrum, A. medirostris,* and *A. oxyrinchus* by counting the number of times fish passed a set point moving upstream during daylight hours (Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005). Nocturnal migratory behavior was measured for all species in the same way as migratory behavior, but at night (Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005).

From these studies, I identified and analyzed five continuous behavioral ontogenetic characters that were measured in a standard format across studies, and comparable across four of the included species- *A. transmontanus, A. oxyrhynchus, A. brevirostrum,* and *A. medirostris.* These characters are:

1. The earliest age (in degree-days posthatch) at which fish begin to swim above the substrate;
2. The age at which downstream migration begins, regardless of whether migration is during the day or night (foraging begins at the same time);
3. The age at which migration ends;
4. The age at which fish first use open habitat, including the open bottom, the water column, and the water surface regardless of whether they previously hid in rocky substrate; and
5. The earliest age at which fish use open habitat after a period of time hiding in rocky substrate.

General habits of all five taxa, available from the literature, provided the basis for three categorical behavioral ontogenetic characters (Kempinger 1988, Richmond and Kynard 1995, Kynard and Horgan 2002, Kynard *et al.* 2005, Kynard and Parker 2005, Smith and King 2005).

1. Nocturnality: categorized as primarily nocturnal or diurnal during the larval stage;
2. Migration: timing of migration was determined to be either early (as free embryos) in development or late (as larvae) for all five species; and
3. Larval behavior related to phototaxis: photopositive or photonegative.

2.2 Metrics for Phylogenetic Signal in Individual Traits

To test for phylogenetic signal in individual ontogenetic characters, I calculated two autocorrelation indices (Moran's I and Abouheif's C-mean) following Munkemuller *et al.* (2012). A full mitogenome Bayesian phylogeny, originally including 22 sturgeon species (Figure 1), was used as the reference phylogeny for these tests. All species except *A. transmontanus, A. oxyrinchus, A. fulvescens, A. brevirostrum,* and *A. transmontanus* were pruned (using the command drop.tip in the ape package, version 3.4, Paradis *et al.* 2004, for R, R Core Team 2014) from this topology prior to comparison (Figure 3). For tests of traits for which data were missing for *A. fulvescens*, this species was also pruned from the phylogeny (Figure 4).

Moran's I was originally designed to measure spatial autocorrelation (Moran 1950), but was modified for use with phylogenies and trait data by Gittleman and Kot (1990). It measures phylogenetic signal based on the correlation between variation of the trait in question and the given phylogeny. To calculate Moran's I, I ran 999 observations under the Moran.I function in the ape package (version 3.4, Paradis *et al.* 2004) for R (R Core Team 2014). This package calculates the observed value of Moran's I and the value that would be expected (in the case of no correlation between the trait and the phylogeny in question). If the observed value is greater than the expected value, the trait in question positively correlated with the phylogeny, indicating phylogenetic signal. I used a 95% confidence interval to determine significance in the difference between the observed and expected values.

Abouheif's C_{mean} tests for independence between traits and a given phylogeny by squaring the differences between the trait values of sister species as defined by the phylogeny (Jombart and Dray 2008). I calculated this parameter with the abouheif.moran function in the adephylo package for R, using the method "oriAbouheif" and running 999 iterations (Jombart and Dray 2008). As with Moran's I, observed and expected (if no phylogenetic signal is present) values of Abouheif's C_{mean} are calculated, and an observed value greater than the expected value indicates phylogenetic signal. Again, I used a 95% confidence interval to determine significance.

2.3 Tree Comparison Metrics

A character matrix including *A. transmontanus, A. fulvescens, A. medirostris, A. oxyrinchus,* and *A. brevirostrum* was constructed. Characters included in the matrix were coded as categorical (nocturnality, migration, and phototaxic habit; Table 1) or continuous (ages at which teeth first appear, mandibular arch elements appear, teeth are fully developed, teeth are completely resorbed, and the yolk sac is completely absorbed; Table 2). Single-character trees were constructed using the "Heuristic Add and Rearrange Tree Search" function in Mesquite v. 3.04 (Maddison and Maddison 2015). Criteria used in tree searches were Parsimony Character Steps, using the SPR Tree Rearranger and storing 100 maximum trees. The trees were unrooted.

Rzhetsky-Nei (R-N) distances are a metric used to calculate the difference between two topologies using the number of bipartitions that differ between trees (Rzhetsky and Nei 1992). High R-N distances indicate divergent trees, low distances indicate similarity between trees, and an R-N distance of 0 indicates identical topologies (Duchene *et al.* 2011). Statistical significance tests are not available for R-N distances; this is a method for comparison between topologies only. Rzhetsky-Nei distances were calculated to examine similarity in phylogenetic signal between single-character trees and a reference phylogeny (a full mitogenome Bayesian sturgeon phylogeny pruned (using the command drop.tip in the ape package, version 3.4, Paradis *et al.* 2004, for R, R Core Team 2014) to include only the species under consideration) using the dist.topo function in the ape package in R (Paradis *et al.* 2004, R Core Team 2014).

2.4 Phylogenetic Comparative Tests

To test whether morphological characters correlate with behavioral characters (i.e., are linked), I conducted Phylogenetically Independent Contrast (PIC) and Pearson's Product Moment of Correlation analyses. PIC allows for tests of correlation with the confounding factor of shared evolutionary history removed, but when few taxa are under consideration, it is recommended that correlation tests in which no branch transformations occur (e.g., Pearson's Product Moment of Correlations) are run alongside PIC (Swenson 2014, Paradis 2006). Therefore, Pearson's Correlations were also calculated, and correlations were considered significant if they were found to be so by both PIC and Pearson's Correlations.

Phylogenetically Independent Contrasts were run for each continuous ontogenetic character (10 morphological and five behavioral characters) using the pic function in the ape package (version 3.4, Paradis *et al.* 2004) for R (R Core Team 2014) and the pruned full mitogenome phylogeny. PICs were plotted between every possible morphological-behavioral character pair. Because plots revealed that observed correlations were, for the most part, linear, a simple linear regression with the regression line forced through the origin was performed for each pair (Swenson 2014, Paradis 2006). Correlation coefficients were calculated on the PIC-transformed data for morphological and behavioral character pairs using the cor.test function in R, and a 95% confidence interval was used to determine significance. Pearson's product moment correlation coefficients were calculated for all character pairs using the cor.test function in R and a 95% confidence interval to determine significance.

3. Results

3.1 Traits

3.1.1 Morphological Ontogenetic Traits: Age

Morphological milestones (Table 2) in *A. oxyrinchus* and *A. transmontanus* occurred in the same order, as did traits in *A. medirostris* and *A. fulvescens. Acipenser brevirostrum* reached the measured milestones in a unique order.

In *A. oxyrinchus* and *A. transmontanus*, the first appearance of mandibular arch elements occurs at two and three days posthatch, respectively, followed by the appearance of teeth at four days posthatch in both species. The yolk sac is completely resorbed at six days posthatch in *A. oxyrinchus* and 9 days posthatch in *A. transmontanus.* Teeth are completely developed at 19 and 20 days posthatch in *A. oxyrinchus* and *A. transmontanus*, respectively. After the development of teeth is complete, they are lost via resorption; teeth gradually become thinner and shorter until the only evidence they ever existed is the presence of shallow tooth sockets in the underlying bones. This occurs in *A. oxyrinchus* at 23 days posthatch, and slightly earlier, at 21 days posthatch, in *A. transmontanus.* Eventually the tooth sockets become smoothed out, leaving no evidence of teeth; the available developmental series ended prior to some species reaching this stage, so it was not included in the analyses.

In *A. medirostris* and *A. fulvescens,* teeth are completely developed before the yolk sac is completely resorbed. In *A. fulvescens,* first appearance of mandibular arch elements occurs at day seven and teeth first appear at day eight, whereas both of these are reached at seven days posthatch in *A. medirostris.* Completion of tooth development (13 days) and yolk sac resorption (15 days) occur earlier in *A. fulvescens* than in *A. medirostris* (14 and 18 days posthatch, respectively). Although tooth resorption begins later in *A. medirostris* than in *A. fulvescens*, teeth are completely resorbed in *A. medirostris* earlier than in *A. fulvescens* (21 and 44 days post hatching, respectively).

Mandibular arch elements first appear in *A. brevirostrum* at six days posthatch. Eight days after hatching, the teeth first appear. The complete development of teeth and completion of yolk sac resorption occur at 13 days posthatch. Teeth are completely resorbed, with tooth sockets present, 52 days after hatching.

3.1.2 Morphological Ontogenetic Traits: Length

Lengths at which milestones are reached are provided in Table 2. Commonalities and differences in the order in which these milestones are reached are discussed here.

In *A. oxyrinchus, A. transmontanus, A. brevirostrum,* and *A. fulvescens,* the mandibular arch elements appear first, followed by first appearance of teeth. In *A. medirostris,* the first appearance of teeth and the oral jaws occur at the same size and mark the earliest milestone reached. *Acipenser oxyrinchus* and *A. transmontanus* both have a completely resorbed yolk sac prior to complete tooth development. In contrast, the completion of tooth development occurs prior to complete resorption of the yolk sac in *A. medirostris, A. fulvescens,* and *A. brevirostrum.* The complete resorption of the teeth is the final milestone reached in all five taxa.

3.1.3 Behavioral Ontogenetic Traits: Age

Four of the five sturgeon species under consideration, *A. oxyrinchus, A. medirostris, A, fulvescens,* and *A. brevirostrum*, migrate as larvae, whereas *A. transmontanus* undergoes a short migration during the free embryo phase (Table 2). *Acipenser transmontanus, A. medirostris,* and

A. fulvescens are primarily nocturnal during their larval migration phases (Smith and King 2005, Kynard *et al.* 2005, Kynard and Parker 1995 and 2010). *Acipenser oxyrinchus* and *A. brevirostrum* are diurnal during migratory life stages (Kynard and Horgan 2002).

Phototaxic behavior was characterized for free embryos and larval *A. oxyrinchus, A. brevirostrum, A. transmontanus,* and *A. medirostris* based on tests for preference for illuminated or dark habitat (Kynard 1997, Kynard *et al.* 2005, Kynard and Horgan 2002, Kynard and Parker 2005 and 2010); phototaxis data were not available for *A. fulvescens* (Table 1). All four species are photonegative as free embryos. Both *A. oxyrinchus* and *A. brevirostrum* become photopositive when they reach the larval stage. *Acipenser transmontanus* and *A. medirostris* continue photonegative behavior into the larval stage of life.

3.2 Pruned (reference) topologies

When pruned (using the command drop.tip in the ape package, version 3.4, Paradis *et al.* 2004, for R, R Core Team 2014) to include only the five taxa in these analyses, the full mitogenome Bayesian topology included two monophyletic groups: one in which *A. oxyrinchus* is sister to (*A. transmontanus, A. medirostris*), and one in which *A. fulvescens* and *A. brevirostrum* are sister species (Figure 3). When *A. fulvescens* was pruned out of the topology for analyses including only four taxa, *A. brevirostrum* fell outside of the *A. oxyrinchus,* (*A. transmontanus, A. medirostris*) clade (Figure 4). These topologies were used as the reference phylogenies in tree comparison and PIC.

3.3 Phylogenetic Signal in Individual Traits

Moran's I and Abouheif's C_{mean} both returned statistically significant results for one morphological character (Table *3*). The age at which teeth are completely resorbed but tooth sockets are still present carries phylogenetic signal matching that of the reference phylogeny. Neither length (Table 4) nor behavioral characters exhibited phylogenetic signal.

3.4 Tree Comparisons

3.4.1 Morphological Character Tree Comparisons

The topology recovered by the tree based on age at which teeth are completely resorbed but tooth sockets are present exactly matched the reference phylogeny (Rzhetsky-Nei distance = 0; Figure 3).

The trees based on the number of days at which teeth first appear, at which teeth are completely developed, and at which the yolk sac is completely resorbed also recovered *A. brevirostrum* and *A. fulvescens* as sister taxa, although they recovered *A. medirostris* as sister to these species, and (*A. transmontanus, A. oxyrinchus)* as sister to this clade (Rzhetsky-Nei distance = 2; Figure 5). The topology based on first appearance of mandibular arch elements also recovered *A. transmontanus* and *A. oxyrinchus* as sister taxa, but with *A. medirostris* sister to *A. fulvescens,* and *A. brevirostrum* outside of that group (Rzhetsky-Nei distance = 2; Figure 6).

3.4.2 Behavioral Character Tree Comparisons

The tree based on the timing of initial migration in sturgeon larvae recovered a polytomy consisting of *A. brevirostrum, A. fulvescens, A. medirostris,* and *A. oxyrinchus,* to which *A. transmontanus* is sister (Rzhetsky-Nei distance = 2; Figure 7). Two different topologies, both with a Rzhetsky-Nei distance = 3, were recovered by nocturnal behavior (Figure 8) and phototaxis (Figure 9).

Nocturnal behavior recovered a polytomy composed of *A. fulvescens, A. medirostris,* and *A. transmontanus*, with *A. brevirostrum* and *A. oxyrinchus* as sister taxa (Figure 8). Data on the phototaxic habits of *A. fulvescens* were not available. This trait recovers a topology with two sister groups: *A. medirostris* and *A. transmontanus;* and *A. brevirostrum* and *A. oxyrinchus* (Figure 9). No behavioral traits recovered the same topology as the full mitogenome.

3.5 Correlations between Characters

Correlation was shown in one pair of behavioral and morphological characters, the age at which migration ends and the length at which teeth are completely developed. Both PIC and Pearson's product moment correlation tests identified a negative correlation between these characters (Pearson: c=-0.97, p=0.03; PIC: adjusted R-squared=0.92, p=0.03; Figures 7, 5, and 10), showing that correlation is present independent of phylogeny. This indicates that species whose teeth are fully developed at smaller sizes end their larval migration at older ages than those whose teeth finish developing later. No other morphological-behavioral trait pair showed evidence of correlation in both transformed (PIC) and non-transformed (Pearson) analyses.

4. Discussion and Conclusion

4.1 Phylogenetic Signal

The phylogenetic signal shown, by autocorrelation indices, in the age at which teeth are completely resorbed but tooth sockets are still present, is refelected by the fact that this trait recovered the same topology as the full mitogenome. The failure to detect phylogenetic signal in other traits does not necessarily mean that phylogenetic signal is not present. Autocorrelation tests do not pick up on all aspects of phylogenetic signal, so signal could be present but undetected by the methods used. Additionally, few species were included in these analyses, and tests for phylogenetic signal perform better when more species are considered (Blomberg *et al.* 2003, Munkemuller *et al.* 2012). Previous evaluations of the effect of sample size (number of nodes in a phylogeny) found that autocorrelation tests perform better at larger sample sizes (Munkemuller *et al.* 2012); these tests have not been evaluated for use with fewer than eight species (Pavoine *et al.* 2008). Still, my results suggest further consideration of the patterns detected and factors other than phylogenetic signal that may drive these patterns.

Considering behavioral characters in the context of the full mitogenome phylogeny revealed that both the full mitogenome and phototaxic behavior link *Acipenser medirostris* and *A. transmontanus* as sister species. This might suggest some phylogenetic signal in phototaxis even though signal was not detected by autocorrelation metrics. However, factors other than common ancestry may contribute to this similarity between these species. Although these species are sympatric, the specimens examined in behavioral analyses were not from the same river of origin so I cannot conclude that habitat-related pressures drive the similarity. Convergent evolution, as suggested by previous studies (Kynard and Horgan 2002), may play a role in shared behavioral traits. Further analyses with data on more species could help shed light on this similarity.

Behavioral trait similarities between other species are influenced by non-phylogenetic factors; *Acipenser oxyrinchus* and *A. brevirostrum* are not closely related (based on the full mitogenome phylogeny), but they share several larval behavioral traits. Similarity in their diurnal/nocturnality behavior and phototaxic habits, coupled with the fact that these characters did not appear to exhibit phylogenetic signal, suggests that factors other than phylogeny influence these behaviors. Kynard and Horgan (2002) concluded that similarities in the larval behavior of these species was likely due to common adaptations. Kynard *et al.* (2005) described predation, foraging, competition, and visual acuity as factors that contribute to nocturnal behavior. While these species are sympatric, individuals used in behavioral studies did not come from the same river of origin. Pressures influencing these behaviors cannot, therefore, be directly linked to shared habitat but may reflect similarities between habitats.

4.2 Correlations between Characters

Because teeth may play a role in the foraging that occurs during migration, the negative correlation between the age at which migration of early life stage sturgeons is completed and the length at which tooth development is complete is somewhat counter-intuitive (Kynard 1997, Kynrd and Horgan 2002, Kynard and Park 2005 and 2010, Kynard *et al.* 2005). Species such as *A. oxyrinchus*, in which tooth development is complete at relatively small sizes, also complete their migration later than those (e.g. *A. medirostris*) that develop teeth at larger sizes. Although the end of migration does not correlate with the age at which tooth development is complete, and length data are unavailable for the individuals used in behavioral studies, teeth are present when foraging begins (Richmond and Kynard 1995). The developmental stage of the teeth during this migratory and foraging behavior is unknown.

Although the correlation between the age at which migration ends and completion of tooth development supports my hypothesis that the feeding-related morphological characters I measured would be correlated to the behavioral traits examined, the lack of evidence for correlation between most morphological and behavioral characters must be considered.

Factors other than an actual lack of correlation might have influenced my analyses. Behavioral characters obtained from the literature (Kynard *et al.* 1997, 2002, 2005, 2010) were measured as age in degree days at which milestones occurred. Age in degree-days was calculated by Kynard *et al.* (1997, 2002, 2005, 2010) as the sum of 0.5 x temperature ° C for all days post-hatch. Because data on the temperature at which all specimens in my developmental series were not available, morphological traits were measured in days post-hatch rather than degree-days post-hatch. This difference in age measurement may have contributed to the lack of correlation between age-based morphological and behavioral characters.

Similar to autocorrelation metrics, PICs are sensitive to sample size. While Ackerly (2000) found that PIC is relatively reliable for the exploration of links between continuous traits even without complete taxon sampling, the use of non-phylogenetically-transformed data necessarily accompanied PIC due to the small sample size included in this study. The inclusion of additional taxa in future analyses would make for more reliable analyses.

The timing of behavioral and morphological shifts may be influenced by different factors; predation risk may influence behavior while temperature influences morphological development. Some of the behavioral characters I evaluated, such as use of open habitat, are likely linked to predation risk. For example, the manner in which larval sturgeons migrate has been hypothesized to be a habit used to find cover and avoid predation (Richmond and Kynard 1995). The use of open habitat is linked to predation and food availability- sturgeons begin using open habitat and foraging when predation risk is not prohibitively high (Parker 2007).

In addition to helping shed light on the lack of correlation between the timing of morphological and behavioral shifts, the influence of different factors on growth may explain discrepancies in the age and length at which morphological milestones are reached. While food availability affects both behavior and growth (Gisbert and Williot 1997), the exact nutritional components of the food available can strongly impact growth and skeletal development of larval fishes (Gawlicka *et al.* 2002, Cahu *et al.* 2003). Temperature is widely known to impact growth of both fresh and saltwater fishes, and colder temperatures have been linked to slow growth in *A. brevirostrum* and *A. oxyrinchus* (see Houde 1989, Claramunt and Wahl 2000, Hardy and Litvak 2004). Light and day length have also been shown to impact larval fish growth, with longer days increasing growth (Boeuf and Bail 1999).

4.3 Concluding Remarks

Overall, I did not find the timing of the behavioral ontogenetic milestones considered here to carry phylogenetic signal. The timing of one morphological ontogenetic milestone did carry phylogenetic signal; the use of this type of data needs to be considered carefully in phylogenetic analyses. This study included few behavioral and morphological milestones in the context of a small (five taxon) phylogeny. Other studies have found these types of characters useful in the recovery (e.g. Tucker *et al.* 1993), resolution (e.g. Scholtz *et al.* 2009), and polarization (Bryant 2001) of phylogenies. The timing of the aspects of morphological and behavioral development in larval sturgeons that were considered here are likely impacted by factors other than common ancestry, such that any phylogenetic signal is swamped by selection pressures. Food availability and quality, temperature, and other habitat factors are known to influence early life history (Houde 1989, Claramunt and Wahl 2000, Gawlicka *et al.* 2002, Cahu *et al.* 2003, Hardy and Litvak 2004). Therefore, signal may be present in these character, but that signal may be swamped out by other factors that play a role in the timing of development.

The correlation between one morphological and one behavioral trait support the fact that form and function are inextricably linked, but lack of correlation among other traits suggests that behavior and morphological development are impacted by differing factors. While form and function are definitely linked (larval sturgeons feed in the water column until their oral jaws are ventrally oriented to allow for benthic feeding), the characters considered here appear to be influenced more strongly by outside factors than by one another.

Finally, my findings may be biased by the small number of taxa considered in analysis. The best metrics available for the number of taxa included were selected for phylogenetic signal testing, and the restricted taxon sampling was accounted for in correlation tests by including non-transformed data along with PIC. However, the inclusion of more taxa and consideration of more traits is recommended in future studies of this kind.

Acknowledgements

Thanks to the Hilton lab, specifically Todd Clardy and Alison Deary for help with brainstorming and implementing methods. Thanks to Eric Hilton for support throughout this chapter, and to Bernie Kuhajda for insightful correspondence. Peter Struffenegger and Sterling Caviar and Dave Sien and Manning/Cedarville Hatcheries, Maryland Department of Natural Resources provided specimens for study. Funding sources included National Science Foundation (DEB 0841691), the GK-12 Program, Virginia Institute of Marine Science Dean and Director's Office and Fisheries Department, and John E. Olney Sr. Ichthyology Award.

References

Ackerly D.D. 2000. Taxon Sampling, Correlated Evolution, and Independent Contrasts. Evolution, 54(5): 1480-1492.

Arnold S.J. 1983. Morphology, Performance and Fitness. American Zoologist, 23(2): 347-361.

Asgari, R., Rafiee, G., Eagdei, S., Shahrooz, R., Pourbagher, H., Agh, N., Gisbert, E., 2013. Ontogeny of the digestive system in hatchery produced Beluga (*Huso huso* Linnaeus, 1758); a comparative study between Beluga and genus Acipenser. Aquaculture Nutrition, 20(6): 595-608.

Bemis W.E. and Kynard B. 1997. Sturgeon rivers: an introduction to acipenseriform biogeography and life history. Environmental Biology of Fishes, 48:167-183.

Bertossa R.C. 2011. Morphology and behaviour: functional links in development and evolution. Philosophical Transactions of the Royal Society B, 355(1574): 2056-2068.

Birstein V. 1993. Sturgeons and Paddlefishes: Threatened Fishes in Need of Conservation. Conservation Biology, 7(4): 773-787.

Blomberg S.P., Garland, T., Ives, A.R. 2003. Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits are More Labile. Evolution, 57(4): 717-745.

Boeuf G. and Bail P.L. 1999. Does light have an influence on fish growth? Aquaculture, 177(1-4): 129-152.

Brannon E., Brewer S., Setter A., Miller M., Utter F., Hershberger. 1985. Columbia River White Sturgeon (*Acipenser transmontanus)* Early Life History and Genetics Study. Report, National Marine Fisheries Service, Northwest and Alaska Fisheries Center.

Bryant H. N. 2001. Character polarity and the rooting of cladograms. In: G. P. Wagner (Ed.). The character concept in evolutionary biology. Academic Press. Pp. 321-340.

Buckley J. and Kynard B. 1981. Spawning and rearing of shortnose sturgeon from the Connecticut River. Prog. Fish-Cult. 43: 74-76.

Cahu C., Infante J.Z., and Takeuchi T. 2003. Nutritional components affecting skeletal development in fish larvae. Aquaculture, 227(1-4): 254-258.

Claramunt R.M. and Wahl D.H. 2000. The Effects of Abiotic and Biotic Factors in Determining Larval Fish Growth Rates: A Comparison Across Species and Reservoirs. Transactions of the American Fisheries Society, 129(3): 835-851.

Dang T and Zhang C. 2014. Postembryonic Ontogeny of Lake Sturgeon *Acipenser fulvescens* with Quantitative Data of Forebrain Growth. Annals of Aquatic Research, 1(1): 1004.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Findeis, E.K. 1997. Osteology and phylogenetic interrelationships of sturgeons (Acipenseridae). Environmental Biology of Fishes, 48(1-4): 73-126.

Fink W.L. and Zelditch M.L. 1995. Phylogenetic Analysis of Ontogenetic Shape Transformations: A Reassessment of the Piranha Genus Pygocentrus (Teleostei). Systematic Biology, 44(3): 343-360.

Froese, R. and D. Pauly. Editors. 2015. FishBase. World Wide Web electronic publication. www.fishbase.org, (10/2015)

Garland T.J., Bennett A.F., and Rezende E.L. 2005. Phylogenetic approaches in comparative physiology. Journal of Experimental Biology, 208: 3015-3035.

Garland T.J., Harvey P.H., Ives, A.R. 1992. Procedures for the Analysis of Comparative Data Using Phylogenetically Independent Contrasts. Systematic Biology, 41 (1): 18-32.

Gawlicka A., Herold M.A.m Barrows F.T.m Noue J.D.L., and Hung S.S.O. 2002. Effects of dietary lipids on growth, fatty acid composition, intestinal absorption and hepatic storage in white sturgeon (Acipenser transmontanus R.) larvae. Journal of Applied Ichthyology, 18(4-6): 637-681.

Gisbert E. and Williot P. 1997. Larval behaviour and effect of the timing of initial feeding on growth and survival of Siberian sturgeon (*Acipenser baerii*) larvae under small scale hatchery production. Aquaculture, 156: 63-76.

Gittleman, J.L. and Kot, M. (1990) Adaptation: statistics and a null model of estimating phylogenetic effects. Systematic Zoology, 39, 227–241.

Grohs K.L., Klumb R.A., Chops S.R., Wanner G.A. 2009. Ontogenetic patterns in prey use by pallid sturgeon in the Missouri River, South Dakota and Nebraska. Journal of Applied Ichthyology, 25(2): 48-53.

Hardy R.S. and Litvak M.K. 2004. Effects of Temperature on the Early Development, Growth, and Survival of Shortnose Sturgeon, Acipenser brevirostrum, and Atlantic Sturgeon,Acipenser oxyrhynchus, Yolk-Sac Larvae. Environmental Biology of Fishes, 70(2): 145-154.

Houde E.D. 1989. Comparative Growth, Mortality, and Energetics of Marine Fish Larvae: Temperature and Implied Latitudinal Effects.Fishery Bulletin, 87(3): 471-495.

IUCN 2015. The IUCN Red List of Threatened Species. Version 2015-4. <http://www.iucnredlist.org>. Downloaded on 19 November 2015.

Jombart T. and Dray S. (2008) adephylo: exploratory analyses for the phylogenetic comparative method.

Jordana J., Manteca X., and Ribo O. 1999. Comparative analysis of morphological and behavioral characters in the domestic dog and their importance in the reconstruction of phylogenetic relationships in canids. Genetics and Molecular Biology, 122(1).

Katsu Y., Kohno S., Hyodo S., Ijiri S., Adachi S., Hara A., Guillette L.J., and Iguchi T. 2008. Molecular Cloning, Characterization, and Evolutionary Analysis of Estrogen Receptors from Phylogenetically Ancient Fish. Endocrinology, 149(12): 6300-6310.

Kempinger, J.J. 1988. Spawning and early life history of lake sturgeon in the Lake Winnebago system. Wisconsin. Pp. 110-122 in 11th annual larval fish conference. American Fisheries Society, Symposium 5. Bethesda. Maryland. R.D. Hoyt, editor.

Kirschbaum F. and Williot P. 2011. pp 65-80 in Ontogeny of the European Sturgeon, *Acipenser sturio*. Biology and Conservation of the European Sturgeon *Acipenser sturio* L 1758. Springer Berlin Heidelberg, editor.

Kynard, B. 1997. Life history, latitudinal patterns, and status of shortnose sturgeon. Environmental Biology of Fishes 48:319-334.

Kynard B. and Horgan M. 2002. Ontogenetic behavior and migration of Atlantic sturgeon, *Acipenser oxyrinchus oxyrinchus*, and shortnose sturgeon, *A. brevirostrum*, with notes on social behavior. Environmental Biology of Fishes 63: 137–150.

Kynard, B. and Parker, E. 2005. Ontogenetic behavior and dispersal of Sacramento River white sturgeon, *Acipenser transmontanus*, with a note on body color. Environmental Biology of Fishes. 74: 19-30.

Kynard, B., Parker, E. 2010. Ontogenetic Behavior of Kootenai River white sturgeon, *Acipenser transmontanus*, with a note on body color: A laboratory study Environmental Biology of Fishes. 88: 65-77.

Kynard B., Parker E., Parker T. 2005. Behavior of early life intervals of Klamath River green sturgeon, *Acipenser medirostris*, with a note on body color. Environmental Biology of Fishes, 72: 85–97.

Ludwig A. 2006. A sturgeon view on conservation genetics. European Journal of Wildlife Research. 52(1): 3-8.

Maddison, W. P. and D.R. Maddison. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04 <http://mesquiteproject.org>

Mayden R. L., Kuhajda B. R. 1996. Systematics, taxonomy, and conservation status of the endangered Alabama sturgeon, *Scaphirhynchus suttkusi* Williams and Clemmer (Actinopterygii, Acipenseridae). Copeia, 1996(2): 241–273.

Moran P.A.P. 1950. Notes on continuous stochastic phenomena. Biometrika, 37: 17-23.

Münkemüller T., Lavergne S., Bzeznik B., Dray S., Jombart T., Schiffers K., Thuiller W. 2012. How to measure and test phylogenetic signal. Methods in Ecology and Evolution, 3(4): 743–756.

Near T.J., Eytan R.I., Dornburg A., Kuhn K.L., Moore J.A., Davis M.P., Wainwright P.C., Friedman M., and Smith W.L. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proceedings of the National Academy of Sciences of the United States of America, 109:13698-13703.

Nelson T.C., Doukakis P., Lindley S.T., Schreier A.D., Hightower J.E., Hildebrand L.R., Whitlock R.E. and Webb M.A.H. 2013. Research Tools to Investigate Movements, Migrations, and Life History of Sturgeons (Acipenseridae), with an Emphasis on Marine-Oriented Populations. PLOS One, DOI: 10.1371/journal.pone.0071552.

Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in Rlanguage. Bioinformatics 20: 289-290.

Paradis E. 2006. Analysis of Phylogenetics and Evolution with R. Springer. New York.

Parker E.L. 2007. Ontogeny and Life History of Shortnose Sturgeon (*Acipenser brevirostrum* Leseur 1818): Effects of Latitudinal Variation and Water Temperature. Dissertation; Graduate School of the University of Massachusetts Amherst, Amherst, Massachusetts.

Pavoine S., Ollier S., Pontier D., Chessel D. 2008. Testing for phylogenetic signal in phenotypic traits: New matrices of phylogenetic proximities. Theoretical Population Biology, 73: 79-91.

Peng Z., Ludwig A., Wang D., Diogo R. , Wei Q., He S. 2007. Age and biogeography of major clades in sturgeons and paddlefshes (Pisces: Acipenseriformes). Molecular Phylogenetics and Evolution, 42: 854-862.

Revell L. J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, 3:217-223. doi:10.1111/j.2041-210X.2011.00169.x

Richmond A.M. and Kynard B. 1995. Ontogenetic Behavior of Shortnose Sturgeon, *Acipenser brevirostrum*. Copeia, 1995(1): 172-182

Russell E.S. 1982. Form and Function: a Contribution to the History of Animal Morphology. Chicago: University of Chicago Press.

Scholtz G., Abzhanov A., Alwes F., Biffis C. and Pint, J. 2009. "Development, genes, and decapod evolution" In: Decapod Crustacean Phylogenetics (eds. Martin, J.W., Crandall, K.A. and Felder D.L.) Taylor & Francis, CRC Press, Boca Raton pp. 31- 46

Smith K.M. and King D.K. 2005. Dynamics and extent of larval lake sturgeon *Acipenser fulvescens* drift in the Upper Black River, Michigan. Journal of Applied Ichthyology, 21(3): 161–168.

Swenson N. 2014. Functional and Phylogenetic Ecology in R. Springer. New York.

Tucker S.C., Douglas A.W>, and Han-Xing L. 1993. Utility of Ontogenetic and Conventional Characters in Determining Phylogenetic Relationships of Saururaceae and Piperaceae (Piperales). American Society of Plant Taxonomists, 18(4): 614-641.

Zelditch M.L., Swiderski D.L., Sheets H.D., Fink W.L.. 2004. Geometric Morphometrics for Biologists: A Primer Elsevier Academic Press, New York and London, 437p.

Tables

Table 1: Character matrix of three behavioral traits. For migration, character state 1 indicates that first migration occurs during the free embryo stage, 2 indicates first migration at larval stage. For nocturnality, 1 indicates nocturnal habits during the larval stage, 2 indicates a diurnal habit. For Phototaxis, 0 indicates no data, 1 indicates photonegative behavior, and 2 photopositive behavior.

TRAITS	**CHARACTER STATE**				
	A.t.	***A.b.***	***A.m.***	***A.o.***	***A.f.***
Migration	1	2	2	2	2
Nocturnality	1	2	1	2	1
Phototaxis	1	2	1	2	0

Table 2: Earliest occurrences of ontogenetic changes in four species. KEY: A.t.=*A. transmontanus*; A.o.=*A. oxyrinchus*; A.b.=*A. brevirostrum*; A.m.=*A. medirostris*; A.f.=*A. fulvescens*

TRAITS	Days posthatch					Size (mm TL)				
Timing of morphological shifts	***A.t.***	***A.o.***	***A.b.***	***A.m.***	***A.f.***	***A.t.***	***A.o.***	***A.b.***	***A.m.***	***A.f.***
1st appearance of mandibular arch elements	3	2	6	7	7	9.9	8.2	13.1	18.5	16
1st appearance of teeth	4	4	6	7	8	12.7	9.8	16	18.5	17.6
Yolk sac completely resorbed	9	6	13	18	15	18.5	15.3	19.2	21.8	21.8
Teeth completely developed	20	19	13	14	13	26.6	16	19.2	21.2	19.1
Teeth completely resorbed	21	23	52	21	44	36.2	26.6	58.7	40.3	33.4
Timing of behavioral shifts	***A.t.***	***A.o.***	***A.b.***	***A.m.***						
Fishes swim above benthos	1	8	0	15						
Downstream migration begins	0	8	18	10						
Downstream migration ends	7	27	20	21						
Fishes us open habitat	0	8	9	0						
Downstream migration ends	7	27	20	21						
Fishes us open habitat	0	8	9	0						

Table 3: Calculations of phylogenetic signal in morphological characters based on age at which milestones first occur. *s indicate significance.

	Moran's I			Abouheif's C_{mean}		
TRAITS	**I observed**	**I expected**	**p-value**	**C_{mean} observed**	**C_{mean} expected**	**p-value**
1st appearance of mandibular arch elements	0.02	-0.22	0.31	-0.01	-0.01	0.51
1st appearance of teeth	0.73	0.05	0.30	1.38	0.23	0.21
Yolk sac completely digested	-0.43	-0.38	0.31	-0.67	-0.13	0.80
Teeth completely developed	0.52	-0.03	0.27	1.05	0.15	0.21
Teeth completely resorbed	2.43	0.43	.00*	2.81	0.56	0.04*

Table 4. Calculations of phylogenetic signal in morphological characters based on total length (mm) at which milestones first occur.

	Moran's I			Abouheif's Cmean		
TRAITS	I observed	I expected	p value	Cmean observed	Cmean expected	p value
1st appearance of mandibular arch elements	-0.42	-0.48	0.53	-0.17	-0.48	0.76
1st appearance of teeth	-0.29	-0.13	0.35	-0.04	-0.11	0.36
Yolk sac completely resorbed	-0.33	-0.27	0.47	-0.04	-0.12	0.37
Teeth completely developed	-0.17	0.36	0.33	0.04	0.23	0.36
Teeth completely resorbed	-0.34	-0.53	0.77	-0.01	-0.41	0.66

Figures

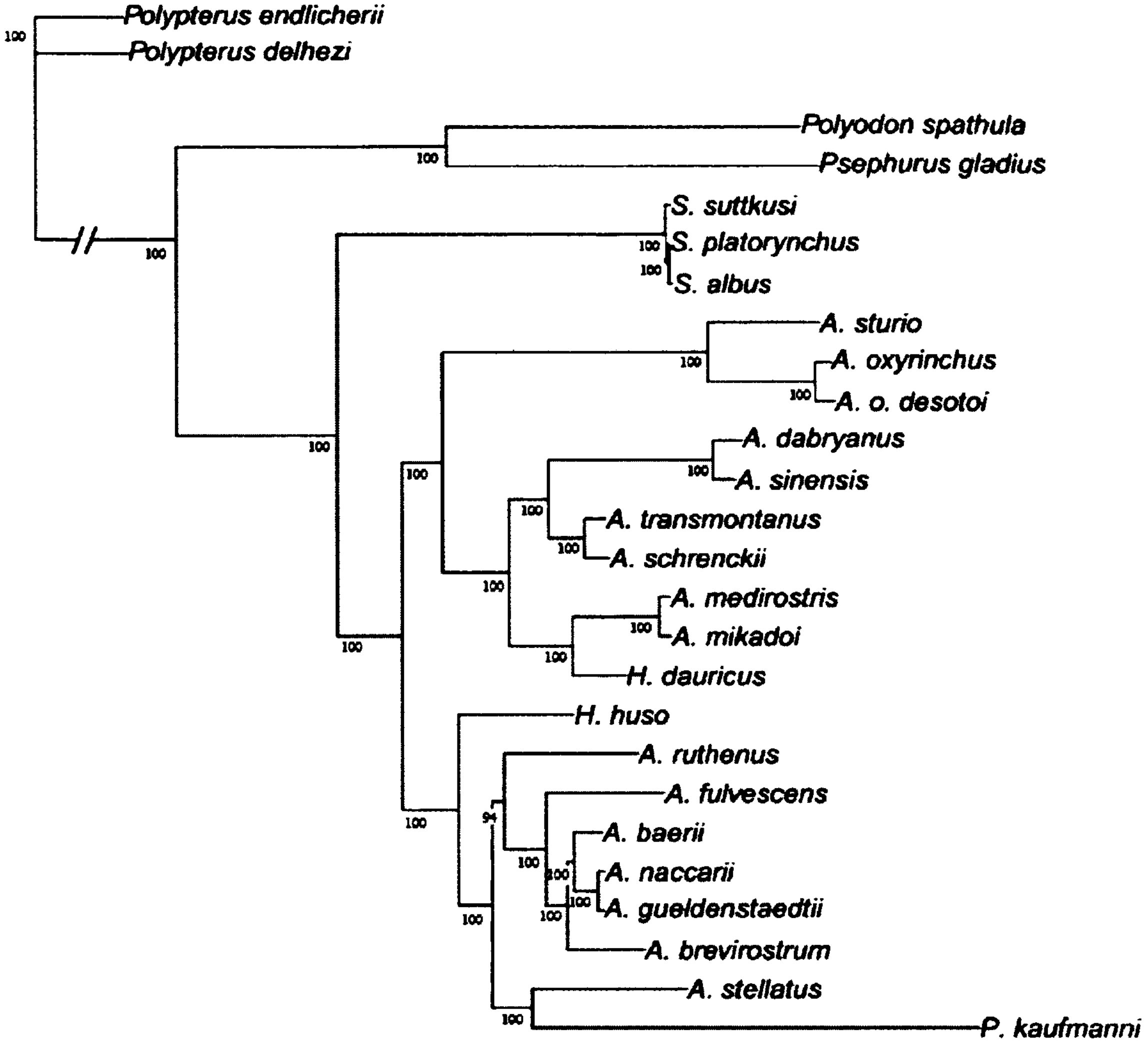


Figure 1: Bayesian full mitogenome phylogeny. Posterior probability values are to the left of and below each node. Branch lengths indicate estimated amount of evolutionary (or genetic) change. Long branches (pertaining to outgroups) have been broken to allow for better visualization of ingroup relationships.

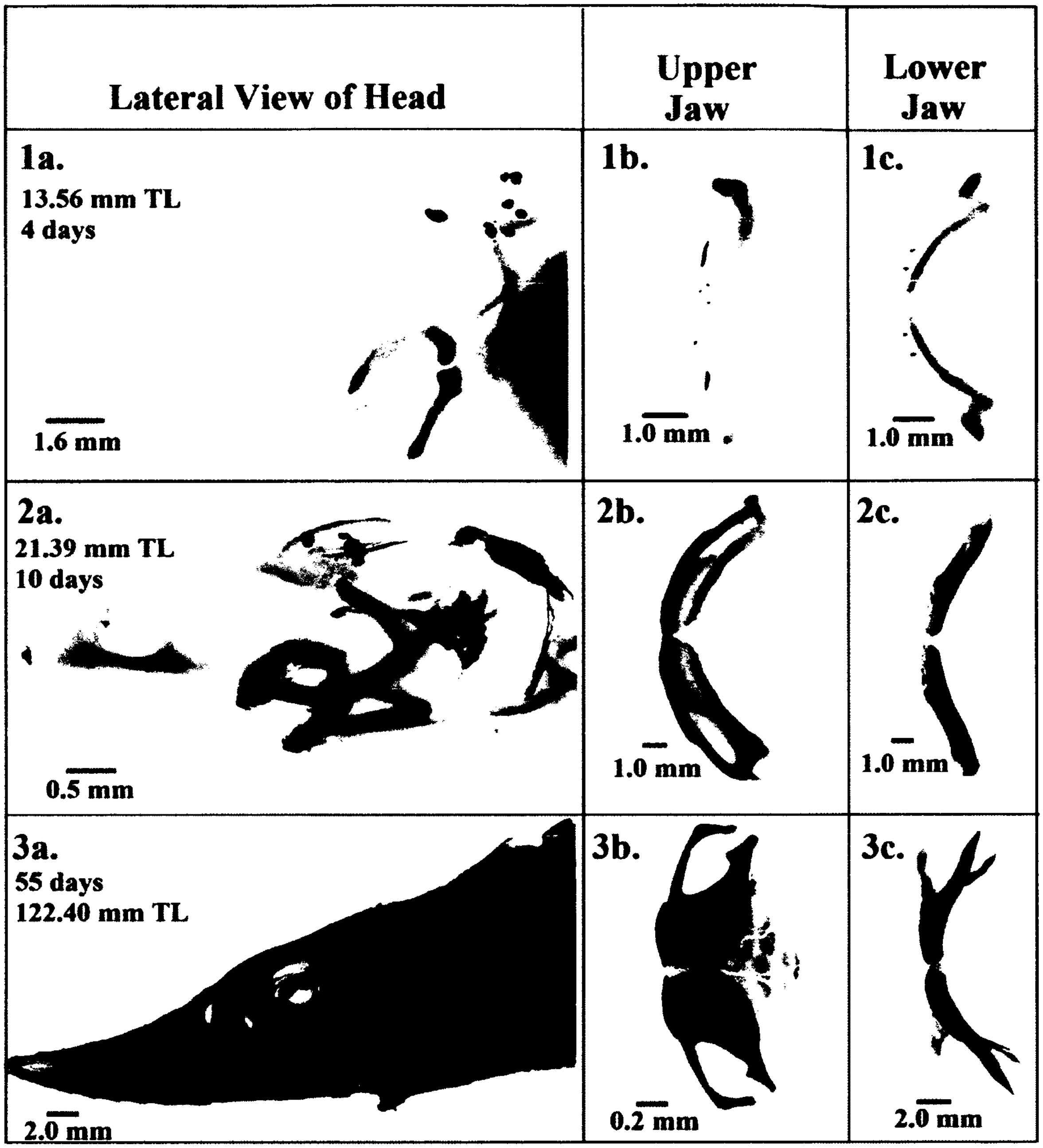


Figure 2: Example of cleared and stained specimens (*A. transmontanus*) illustrating various stages of tooth and mandibular arch elements. Boxes 1a,1b, and 1c illustrate appearance of teeth and mandibular arch elements. Boxes 2a, b, and c show developing teeth. Boxes 3a, b, and c show a specimen with teeth completely resorbed.

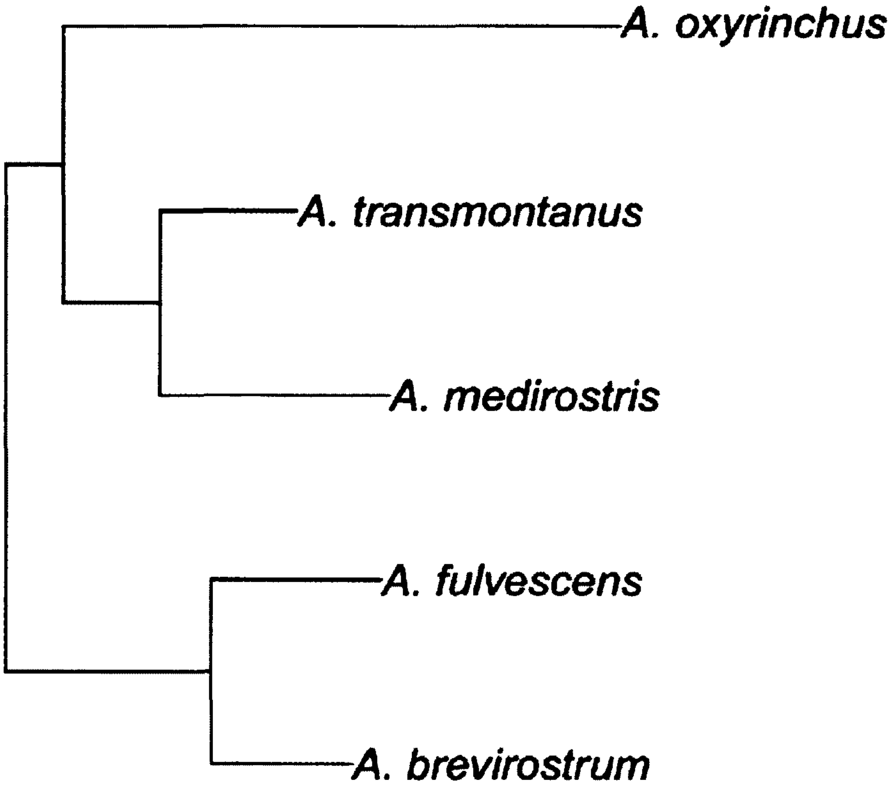


Figure 3: Full mitogenome Bayesian phylogeny, pruned so that only Acipenser oxyrinchus, A. brevirostrum, A. transmontanus, A. medirostris, and A. fulvescens appear. This topology is also shown by single-character tree based on the age at which complete resorption of teeth except for sockets occur.

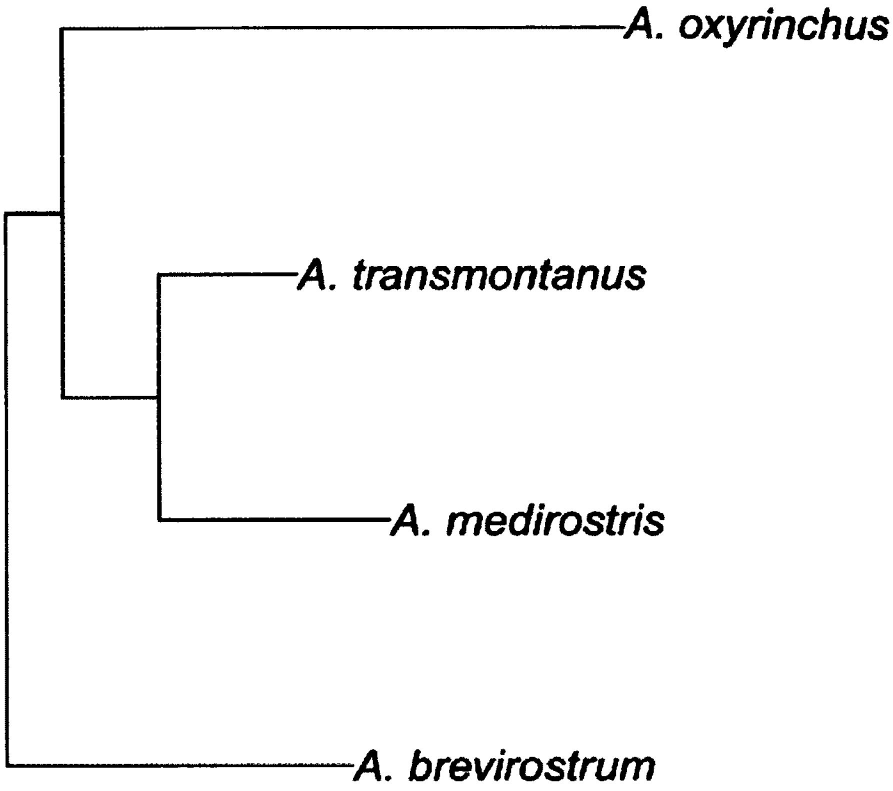


Figure 4: Full mitogenome Bayesian phylogeny, pruned so that only Acipenser oxyrinchus, A. brevirostrum, A. transmontanus, and A. medirostris appear.

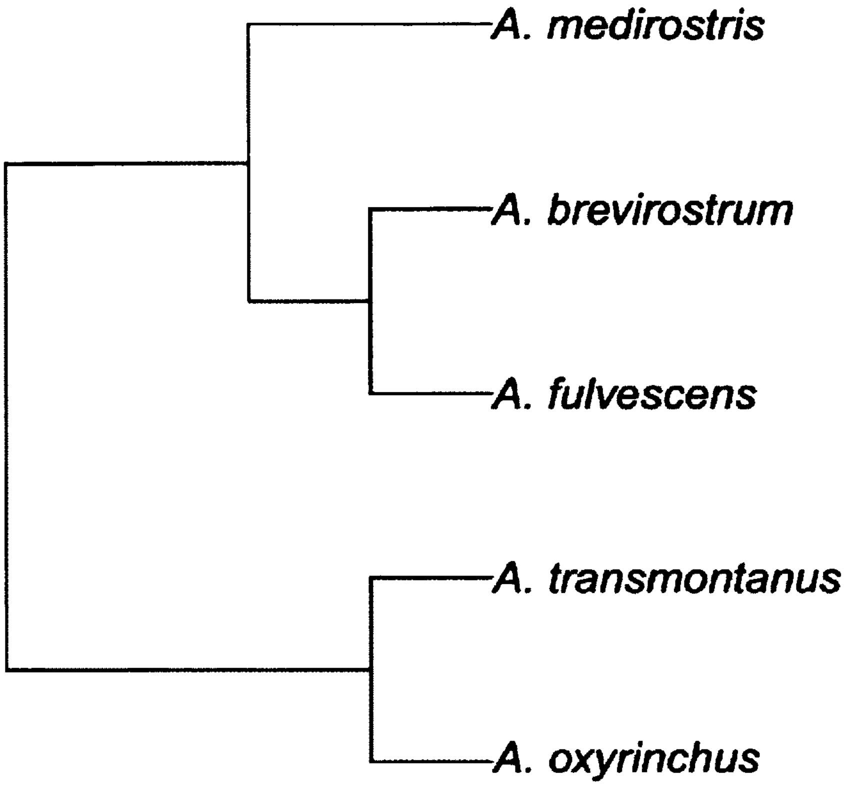


Figure 5: Topology shown by single-character (yolk sac and teeth) trees based on age at which first appearance of teeth, complete development of teeth, and complete resorption of yolk sac occur.

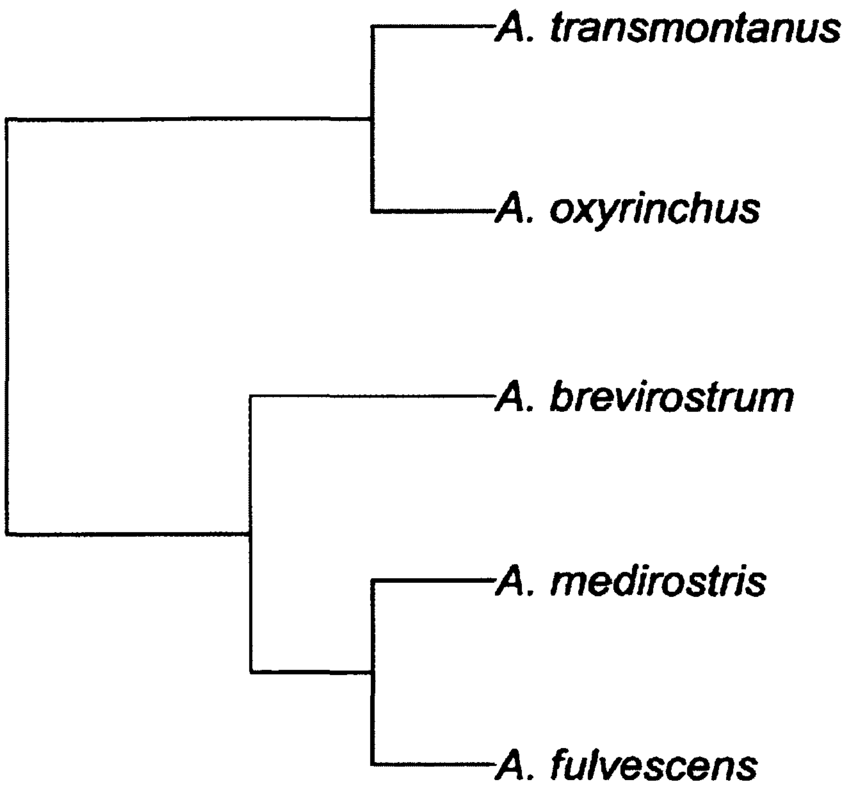


Figure 6: Topology shown by single-character (jaws elements) tree based on age at which first appearance of mandibular arch elements occurs.

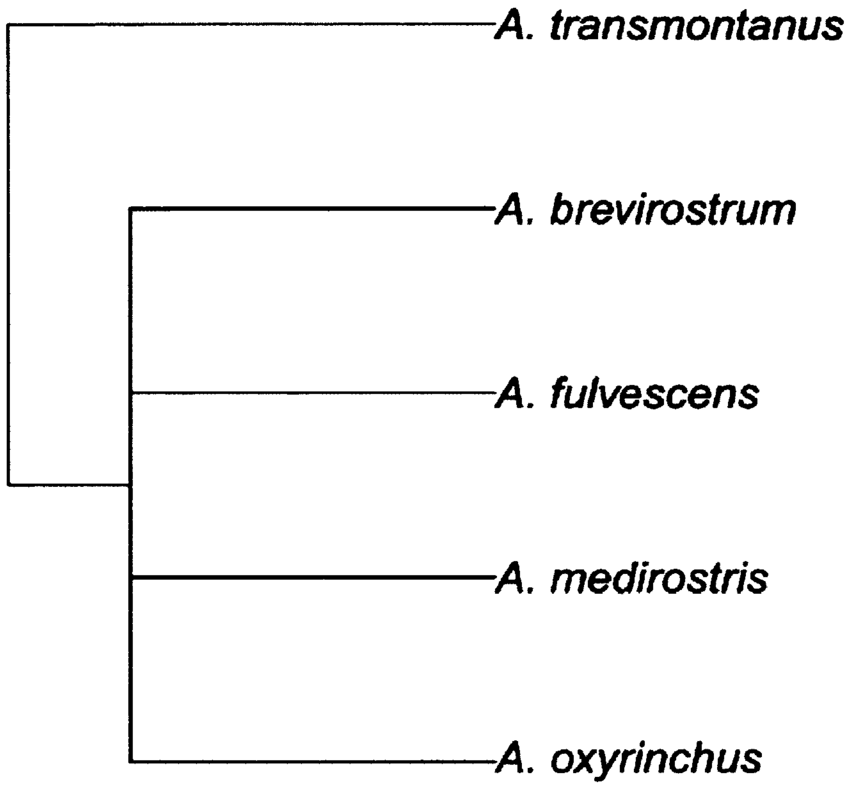


Figure 7: Topology shown by single-character (migration) tree based on age at which migration begins.

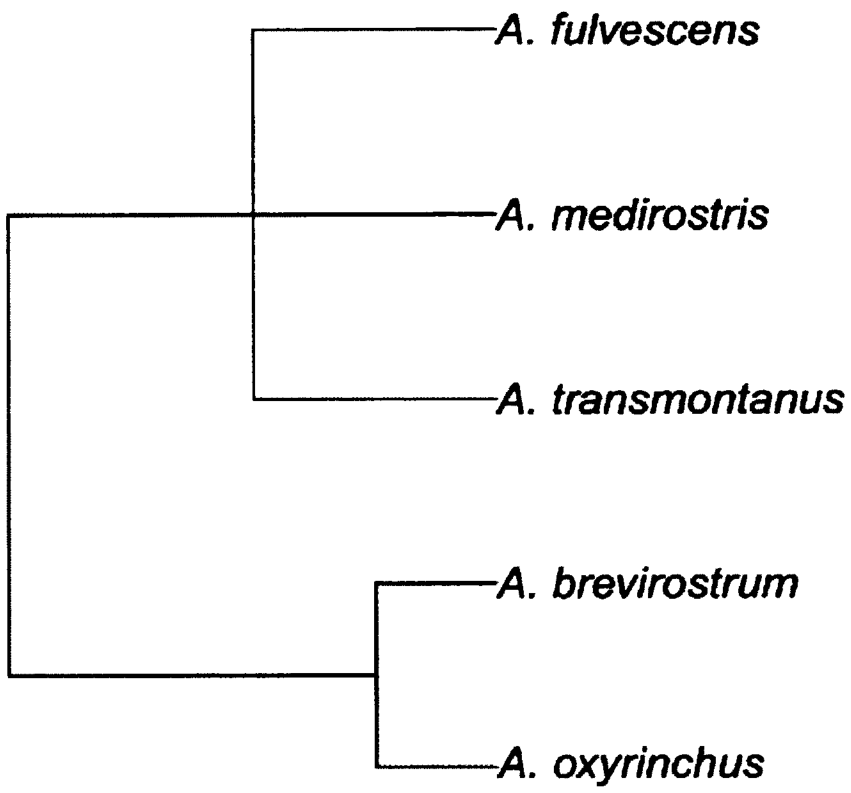


Figure 8: Topology shown by single-character (nocturnality) trees based on larval nocturnal behavior.

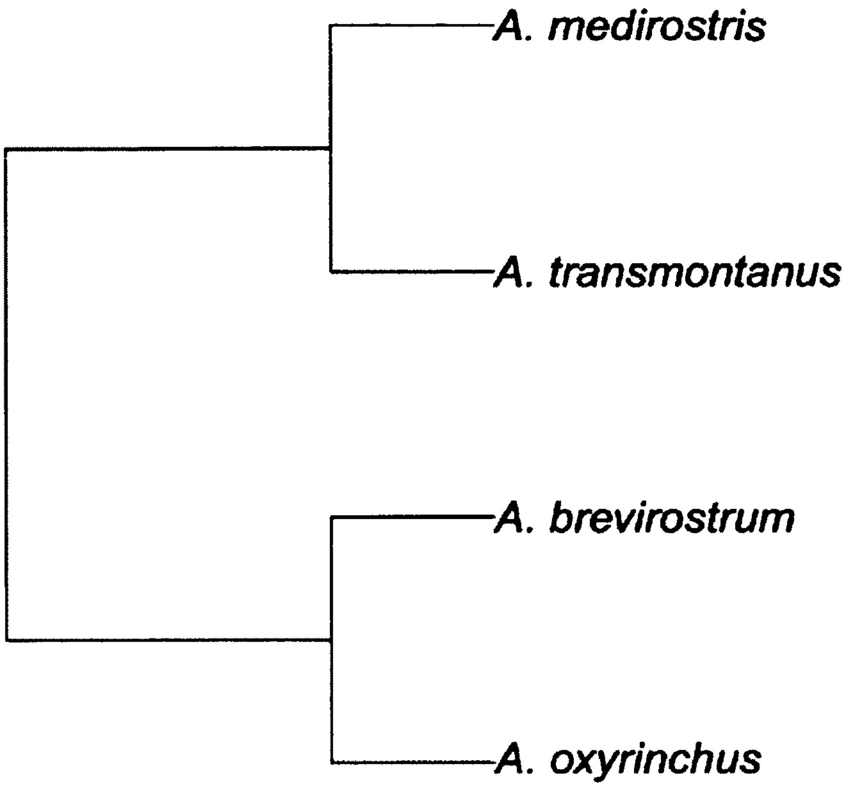
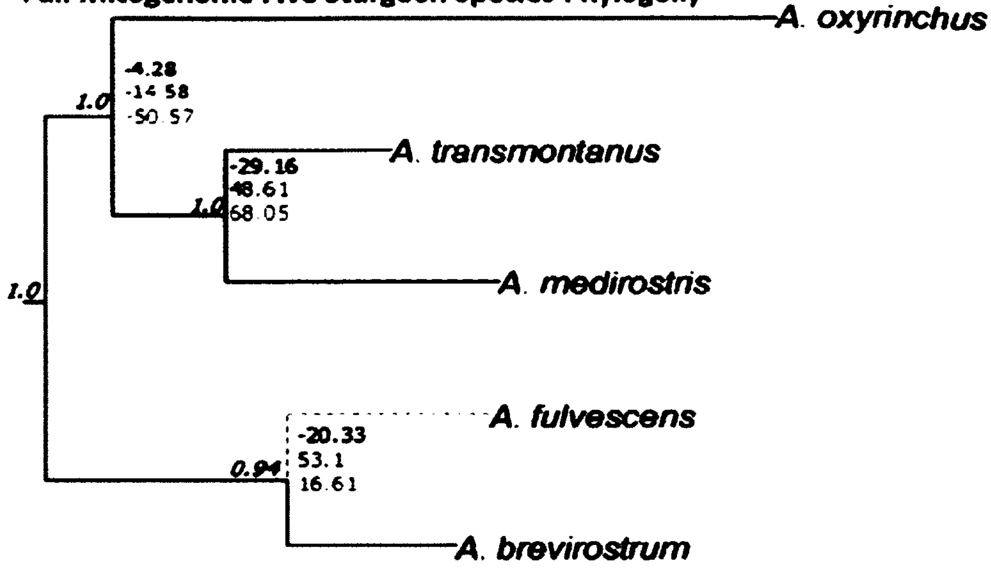


Figure 9: Topology shown by single-character (phototaxis) trees based on larval phototaxic behavior.

Full Mitogenome Five Sturgeon Species Phylogeny



Key:
Dashed lines indicate species included only in strictly morphological analyses.
Posterior probability values: Italics to the left of each node
PIC values for traits:
Black- length at which teeth are completely developed
Red- age at which migration begins
Green- age at which migration ends

Figure 10: Phylogenetic Independent Contrast values used to calculate correlations between traits.

CONCLUSIONS

Evolutionary relationships of sturgeons have been contentious, as phylogenetic studies based on different datasets have resulted in varied hypotheses of sturgeon evolution. Although general trends hold from study to study, specific relationships, namely the position of the genus *Scaphirhynchus* and the composition of subfamilies, have remained questionable. I used the most comprehensive dataset available thus far, over 14,000 base pairs of the sturgeon mitogenome for 22 species and subspecies, to produce a new sturgeon phylogenetic hypothesis. This hypothesis indicates, with strong support, a basal position for *Scaphirhynchus* and proposes the subfamily Husinae to be monotypic, including just *H. huso.*

My new phylogenetic hypothesis provided a framework for the exploration of phylogenetic signal among individual mitochondrial genes and gene families. Recent studies of the evolutionary relationships of various taxa (sturgeons, fungi, dolphins) have investigated signal carried by genes and gene families in an attempt to identify which genes are most useful to phylogenetic analyses. I sought to determine which genes most strongly influence phylogenetic signal in the full mitochondrial genome of sturgeons, and how this has impacted previous phylogenetic studies. I found that the protein coding genes and the rRNA 16s have signal similar to that of the full mitogenome, while D-loop, 12s rRNA, and the tRNAs have signal less similar to that of the full mitogenome. Previous phylogenetic analyses of sturgeons have commonly relied heavily upon the protein coding genes, the rRNAs, and D-loop. The fact that some of these regions (D-loop and 12s) have topologies that conflict with that of the full mitogenome while others have signal matching it 1) illustrates that different signal is carried by different genes and 2) indicates why studies using different genes have proposed many different phylogenetic hypotheses for sturgeons.

Sturgeon evolutionary relationships proposed in the full mitogenome phylogeny provided insight into the relationships among regions inhabited by these fishes. A clear division between Pacific and Atlantic regions was supported, and patterns of relationships within each of those groupings were further supported by the exploration of geologic history. These area relationships were further supported by those proposed by past biogeographic studies of other taxa. These results may be added to by future analyses incorporating additional sturgeon species and other taxa that inhabit similar ranges.

Finally, morphological and behavioral ontogenetic characters present in sturgeons were explored. A link between the age at which migration ends and the length at which teeth are completely developed supported the often-made statement that "form and function are inextricably linked". Although links between other morphological-behavioral characters were not found, they may be present but undetectable. Additionally, ontogenetic characters were evaluated for phylogenetic signal, with signal matching the full mitogenome identified in one character- the age at which the resorption of teeth is complete. Many factors contribute to the timing of development in fishes- from temperature and food availability to predation risk. Ontogenetic characters for which signal was not found may be more influenced by these or other environmental factors, they may be present but not detected, or signal differing from that of the full mitogenome might be present.

This new, well-supported sturgeon phylogenetic hypothesis provided a framework for many analyses, from exploration of phylogenetic signal to the investigation of relationships among areas sturgeons inhabit. These analyses provide new information not only about the interrelationships of sturgeons, but about the type of data that may be useful in phylogenetic analyses. My findings may be useful in future studies of sturgeons, including those on biogeography and early life history, and the data used may be incorporated into future phylogenetic analyses. The utility of this study does not end here. As imperiled, ancient fishes, sturgeons are frequently the subject of conservation efforts. The information revealed by my research may be useful to these efforts. For example, sturgeons are more and more frequently raised in aquaculture facilities, and hybridization between captive-raised escaped fishes has been identified as a threat to wild populations (Ludwig *et al.* 2009). Understanding species relationships and habitat needs of sturgeon species may help to identify where hybridization between species is likely and to prevent or mitigate this threat (e.g. identify if escaped fish would be able to 1: survive in habitat available and 2: hybridize with wild sturgeon(s) present). This study may further aid in conservation efforts by providing information that will help prioritize species or geographic areas for conservation (e.g. 1: the evolutionarily "distinctiveness" of species and 2: regions with the highest overall diversity of sturgeons). Ultimately, the data, analyses, and interpretations presented here have added to our knowledge of sturgeons, and it is my hope that these data and my findings will be useful in future research and in the conservation of these remarkable fishes.

References

Ludwig A., Lippold S., Debus L., and Reinartz R. 2009. First evidence of hybridization between endangered sterlets (*Acipenser ruthenus*) and exotic Siberian sturgeons (*Acipenser baerii)* in the Danube River. Biological invasions 11(3): 753-760.

Appendix 1. Alligned data matrix used in phylogenetic analyses presented in Chapter 1

```
P. delhezi ATGACCCTCA TTACACTAAT C------ATC AACCCACTTA TGTATATTAT
P. endlich ATGACCTTTA TTACACTAAT T------ATT AACCCACTTA TATATATTAT
A. fulvesc ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
P. kaufman ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
S. suttkus ATGACCTCTC ACCTGCTAAC CTACCTAATC AACCCACTAG CATACATCGT
S. albus   ATGACCTCTC ACCTGCTAAC CTACCTAATC AACCCACTAG CATACATCGT
S. platorh ATGACCTCTC ACCTGCTAAC CTACCTAATC AACCCACTAG CATACATCGT
A. sturio  ATGACCCCTC ACCTGCTAAC TTACCTAATT AACCCACTAG CATACATCAT
A. o. deso ATGACCTCTC ACCTGCTAAC TTACCTAATT AACCCACTAG CATACATCAT
A. o. oxyr ATGACCTCTC ACCTGCTAAC TTACCTAATT AACCCACTAG CATACATCAT
A. dabryan ATGACCTCTC ACCTATTAAC TTACCTAATC AATCCACTAG CATACATCAT
A. sinensi ATGACCTCTC ACCTATTAAC TTACCTAATC AATCCACTAG CATACATCAT
A. schrenc ATGACCTCTC ACCTACTAAC TTACCTAATC AATCCATTAG CATACATCGT
A. transmo ATGACCTCTC ACCTACTAAC TTACCTAATC AATCCATTAG CATACATCGT
H. dauricu ATGACCTCTC ACCTACTAAC CTACCTAATT AATCCACTAG CATACATCGT
A. mediros ATGACCTCTC ACCTACTAAC CTACCTAATC AACCCACTAG CATACATCGT
A. mikadoi ATGACCTCTC ACCTACTAAC CTACCTAATC AACCCACTAG CATACATCGT
A. stellat ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
H. huso    ATGACCTCTC ACCTACTGAC CTACCTAATT AACCCACTAG CATACATCGT
A. ruthenu ATGACCTCTC ACCTACTAAC CTACCTAATC AACCCACTAG CATACATCGT
A. breviro ATGACCTCTC ACCTACTAAC CTACCTAATT AATCCACTAG CATACATCGT
A. guelden ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
A. naccari ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
A. baerii  ATGACCTCTC ACCTACTAAC CTACCTAATT AACCCACTAG CATACATCGT
P. spathul ATGACCTCTC ACCTGCTAAC CTACCTAATC AATCCATTAG CATACATCAT
P.gladius  ATGACCTCCC ACCTACTGAC CCACCTAATC AACCCACTAG CATACATTGT

TCCTATTCTC TTAGCGATAG CCTTCCTCAC CCTAGTAGAA CGAAAAATTC TAGGTTATAT
TCCGATTCTT CTAGCAATAG CTTTCCTTAC TCTAGTTGAA CGAAAAATTC TCGGCTATAT
TCCANNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
TCCAATCCTA TTAGCAGTAG CATTTTTAAC CCTCATCGAC CGAAAAGCAC GACGCAAGAA
TCCAGTCCTA CTAGCAGTAG CATTCTTAAC CCTCATCGAA CGAAAAGTAT TAGGCTACAT
TCCAGTCCTA CTAGCAGTAG CATTCTTAAC CCTCATCGAA CGAAAAGTGT TAGGCTACAT
TCCAGTCCTA CTAGCAGTAG CATTCTTAAC CCTCATCGAA CGAAAAGTAT TAGGCTACAT
TCCGATTCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTAC TAGGCTATAT
TCCGATTCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTAC TAGGCTATAT
TCCGATTCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTAC TAGGCTATAT
TCCAATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTAT TAGGCTACAT
TCCAATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTAC TAGGCTACAT
TCCAATTCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAAGTGC TAGGCTACAT
TCCAATTCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAAGTGC TAGGCTACAT
TCCAATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAAGTAC TAGGCTATAT
TCCAATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAGGTAC TAGGCTACAT
TCCAATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAGGTAC TAGGCTACAT
TCCAATCCTA CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAAGTGC TAGGCTACAT
TCCAATCCTA CTAGCAGTAG CATTTTTAAC CCTCATCGAG CGAAAAGTGC TGGGCTATAT
TCCGATCCTG CTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTGC TAGGCTATAT
TCCAATCCTA TTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTGC TAGGTTATAT
TCCAATCCTA TTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTGC TAGGTTATAT
TCCAATCCTA TTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTGC TAGGTTATAT
TCCAATCCTA TTAGCAGTAG CATTTTTAAC CCTCATCGAA CGAAAAGTGC TAGGTTATAT
CCCCATCCTA CTAGCAGTAG CATTCCTAAC CCTTATTGAA CGAAAAGTAC TAGGCTACAT
TCCCATCCTG CTAGCAGTAG CATTCTTAAC TCTCATCGAA CGAAAAGTTC TAGGCTACAT

ACAGCTACGA AAAGGCCCAA ATATTGTTGG CCCCTACGGA CTTCTTCAAC CAATTGCAGA
ACAACTACGA AAAGGACCTA ATATTGTTGG CCCCTACGGA CTTCTACAAC CAATCGCAGA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
ACAAAAACGA AAAAGCCCAA ATATCGTTGG GCCTTAGGTC CTGCAGCAAC CGATTGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ATATCGTTGG GCCTTACGGC CTTCTGCAAC CGATTGCTGA
ACAATTACGA AAAGGCCCAA ATATCGTTGG GCCTTACGGC CTACTGCAAC CGATTGCTGA
ACAATTACGA AAAGGCCCAA ATATCGTTGG GCCTTACGGC CTACTGCAAC CGATTGCTGA
ACAACTACGA AAAGGCCCAA ACATTGTTGG ACCTTATGGC CTCCTACAGC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATTGTTGG ACCTTATGGC CTTCTACAGC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCTTATGGT CTTCTACAGC CCATCGCTGA
```

```
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCTTATGGT CTTCTACAGC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCTTATGGT CTTCTACAGC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTCGG ACCCTATGGT CTTCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTCGG ACCCTATGGT CTTCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTCGG ACCCTACGGC CTTCTACAGC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTTCTACAAC CCATCGCTGA
GCAACTACGA AAAGGCCCAA ACATCGTTGG ACCTTACGGC CTGCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTATGGC CTCCTACAAC CCATCGCTGA
GCAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
GCAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCAA ACATCGTTGG ACCCTACGGC CTCCTACAAC CCATCGCTGA
ACAACTACGA AAAGGCCCCA ACATCGTTGG ACCCTACGGC TTACTACAAC CCATTGCTGA
ACAACTACGA AAAGGCCCCA ATATCGTAGG ACCCTACGGA CTACTACAAC CCATTGCCGA

CGGGGTAAAA CTATTCATTA AAGAGCCAGT AAAACCATCA ACCGCCTCCC CCACACTTTT
TGGAGTAAAA CTATTTATTA AAGAACCAGT AAAACCATCA ACCTCCTCCC CAACCCTTTT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CGGCATCAAA CTATTCATCA AAGAACCTGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATTAAA CTATTCATTA AAGAGCCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATTAAA CTATTCATTA AAGAGCCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATTAAA CTATTCATTA AAGAGCCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
CGGCATCAAA CTATTCATCA AAGAACCTGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
CGGCATCAAA CTATTCATCA AAGAACCTGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
CGGCATCAAA CTATTCATCA AAGAACCTGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
CGGCATCAAA CTATTCATCA AAGAACCCGT GCGCCCCACC ACAGCCTCTC CATTTTTATT
CGGCATCAAA CTATTCATCA AAGAACCCGT GCGCCCCACC ACAGCCTCTC CATTTTTATT
CGGTATCAAG CTATTTATCA AAGAGCCCGT GCGTCCCACC ACAGCCTCCC CATTTTTATT
CGGTATCAAG CTATTTATCA AAGAGCCCGT GCGTCCCACC ACAGCCTCTC CATTTTTATT
CGGTATCAAG CTATTTATCA AAGAACCCGT GCGCCCTACC ACAGCCTCTC CATTTTTATT
CGGTATCAAA CTATTTATTA AAGAACCCGT GCGCCCTACC ACAGCCTCTC CATTTTTATT
CGGTATCAAA CTATTTATTA AAGAACCCGT GCGCCCTACC ACAGCCTCTC CATTTTTATT
CGGTATCAAA CTATTTATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGCATCAAA CTATTTATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTCTATT
TGGCATCAAA CTATTTATTA AAGAACCCGT GCGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATCAAA CTATTTATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATCAAA CTATTCATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATCAAA CTATTCATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
TGGTATCAAA CTATTTATTA AAGAACCCGT ACGCCCCACC ACAGCCTCCC CATTTTTATT
CGGTATCAAG CTATTCATTA AAGAACCTGT CCCGCCTACC ACAGCCTCCC CACTTCTATT
CGGCATTAAA CTATTTATTA AAGAGCCTGT CCGCCCCACC ACAGCCTCCC CACTACTATT

TCTACTAACC CCAGCACTAG CACTCACTTT AGCACTGATC CTGTGAATCC CACTTCCAAT
CTTACTTACC CCAACACTTG CCCTCACACT AGCCCTTATT TTATGAATTC CTCTCCCCAT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
TTTAGCAACC CCCATAATAG CACTAACCTT AGCCCTTACC CGATGAATAC CGCTACCAAT
CTTAGCAACC CCAATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCAAT
CTTAGCAACC CCAATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCAAT
CTTAGCAACC CCAATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCAAT
CTTAGCAACC CCCATTATAG CACTAACCTT GGCCCTCACC CTGTGAATAC CCCTACCAAT
CCTAGCAACC CCCATTATGG CGCTAACCTT GGCCCTCACC CTATGGATAC CACTACCAAT
CCTAGCAACC CCCATTATGG CGCTAACCTT GGCCCTCACC CTATGGATAC CACTACCAAT
TTTGGCAGCC CCCATCATAG CACTAACCCT GGCCCTCACC CTATGAATAC CCCTGCCAAT
TTTGGCAGCC CCCATCATAG CACTAACCCT GGCCCTCACC CTATGAATAC CCCTGCCAAT
TTTAGCAGCC CCCATCATAG CACTAACCTT AGCCCTCACC CTATGAATAC CCCTACCAAT
TTTAGCAGCC CCCATCATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCAAT
TTTAGCAGCC CCCATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CACTACCCAT
CTTAGCGGCC CCCATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCCAT
CTTAGCGGCC CCCATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CCCTACCCAT
TTTAGCAACC CCCATCATAG CACTAACCCT GGCCCTCACC CTGTGAATAC CACTACCAAT
TTTAGCAGCC CCCATTATAG CACTAACCTT GGCCCTCACC CTATGAATAC CGCTACCAAT
CTTAGCAACC CCCATCATAG CACTAACCTT GGCCCTCACC CTGTGAATAC CACTACCAAT
TTTAGCAACC CCCATTATAG CACTAACCTT AGCCCTTACC CTATGAATAC CGCTACCAAT
TTTGGCAACC CCCATTATAG CACTAACCCT AGCCCTTACC CTATGAATAC CACTACCAAT
TTTGGCAACC CCCATTATAG CACTAACCCT AGCCCTTACC CTATGAATAC CGCTACCAAT
TTTGGCAGCC CCCATTATAG CACTAACCTT AGCCCTTACC CTATGAATAC CGCTACCAAT
CTTAGCAACT CCAATTCTAG CACTAACCTT AGCCCTCACC CTCTGAATAC CCCTACCCAT
CTTAACAGCC CCAATTATAG CACTAACCCT GGCCCTCACC CTCTGAATAC CCCTACCAAT
```

```
ACCACTAGCA CTCACAGATT TAAACCTAAC TATCCTCTTC ATCCTGGCAG TATCCAGCCT
ACCACTTACA CTAACAGACT TAAACCTTAC CATCCTCTTC ATCTTAGCAG TCTCAAGTCT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
ACCAGACCCA GACGCAGACC TCAACCTAGG TATCCTATTC ATCCTAGCCC TATCAAGTCT
GCCCTACCCA ATCGCAGACC TCAACCTGGG TATCCTATTC ATCCTAGCCC TCTCCAGTCT
GCCCTACCCA ATCGCAGACC TCAACCTGGG TATCCTATTC ATCCTAGCCC TCTCCAGTCT
GCCCTACCCA ATCGCAGACC TCAACCTGGG TATCCTATTC ATCCTAGCCC TCTCCAGTCT
ACCCTATCCA GTCGCAGACC TAAACTTGGG CATCCTATTT ATCCTAGCCC TATCCAGTCT
GCCCTACCCA GTCGCAGACC TAAACCTGGG CATCCTATTT ATCCTAGCCC TGTCCAGTCT
GCCCTACCCA ATCGCAGACC TAAACCTGGG CATCCTATTT ATCCTAGCCC TGTCCAGTCT
GCCCTACCCA ATCGCAGACC TCAACCTAGG CATCCTATTT ATTCTAGCCC TATCCAGTCT
GCCCTACCCA ATCGCAGACC TCAACCTAGG CATCTTATTT ATTCTAGCCC TATCCAGTCT
ACCCTACCCA GTTGCAGACC TCAACCTAGG CATCCTATTT ATTCTAGCCC TATCCAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATTCTAGCCC TATCCAGTCT
GCCCTACCCA GTCGCAGACC TCAACCTAGG TATCCTATTT ATCCTAGCCC TATCCAGCCT
ACCCTACCCA ATCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGCCT
ACCCTACCCA ATCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGCCT
GCCCTATCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCAAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCT TATCCAGCCT
GCCCTACCCG GTCGCAGACC TCAACCTAGG TATCCTATTT ATTCTAGCCC TATCCAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGTCT
ACCCTACCCA GTCGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TATCCAGTCT
GCCCTACTCA ATTGCAGACC TCAACCTAGG CATCCTATTT ATCCTAGCCC TCTCCAGTCT
GCCCTACCCA ATTGCAGATC TTAACCTAGG CATCCTATTT ATCCTAGCCC TTTCCAGCCT

TTCAGTATAC TCCATCCTCG GCTCAGGCTG AGCCTCAAAT TCTAAATACG CTTTAATTGG
CTCTGTCTAC TCTATTTTGG GTTCAGGCTG AGCCTCAAAT TCAAAATATG CCCTAATTGG
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
GGCCGTATAC TCAATCCTAG GCTCCGGCTG AGCCTCTAAC TCAAAATACG CCCTAATCGG
GGCCGTATAC TCAATCCTAG GCTCCGGCTG AGCCTCTAAC TCAAAATACG CCCTAATCGG
GGCCGTATAC TCAATCCTAG GCTCCGGCTG AGCCTCTAAC TCAAAATACG CCCTAATCGG
AGCTGTATAC TCAATCTTAG GTTCCGGCTG AGCCTCCAAT TCAAAATACG CCCTCATCGG
AGCTGTATAC TCAATCTTAG GTTCCGGCTG AGCCTCCAAT TCAAAATACG CTCTTATCGG
AGCTGTATAC TCAATCTTAG GTTCCGGCTG AGCCTCCAAT TCAAAATACG CTCTTATCGG
AGCCGTCTAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATATG CTCTCATTGG
AGCCGTGTAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATATG CTCTCATTGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATATG CTCTCATCGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATATG CTCTCATCGG
AGCCGTGTAT TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CTCTCATCGG
AGCCGTATAT TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CTCTCATTGG
AGCCGTATAT TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CTCTCATTGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
AGCCGTATAC TCGATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATTGG
AACCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAT TCAAAATACG CCCTAATCGG
AGCCGTATAC TCAATCTTAG GCTCCGGTTG AGCCTCCAAC TCAAAATACG CCCTAATCGG
AGCAGTCTAC TCAATCCTAG GATCCGGCTG AGCCTCCACT TCAAAATACG CCCTCATCGG
AGCAGTATAC TCAATCTTAG GCTCTGGCTG AGCCTCCAAT TCAAAATACG CCCTCATCGG

TGCTCTCCGA GCAGTAGCAC AAACAATCTC ATACGAAGTC ACCCTAGGCC TCATCATCCT
TGCACTCCGA GCAGTAGCCC AAACAATTTC TTACGAAGTT ACATTAGGCC TGATCATCCT
GGGCTTTCGA GCGGTAGCAC AAACAATCTC CTACGAAGTA AGCCTTGGCT TGATCCTCTT
GGCACTCCGA GCGGTAGCAC AAACAATCTC CTACGAAGTA AGCCTCGGCT TAATCCTCTT
GGCACTTCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCCTAGGCT TAATCCTCTT
GGCACTTCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCCTAGGCT TAATCCTCTT
GGCACTTCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCCTAGGCT TAATCCTCTT
GGCACTCCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCCTCGGCT TAATTCTCTT
AGCACTCCGA GCAGTAGCAC AAACAATTTC CTACGAAGTA AGCCTTGGCT TAATCCTCTT
AGCACTCCGA GCAGTAGCAC AAACAATTTC CTACGAAGTA AGCCTTGGCT TAATCCTCTT
AGCACTCCGA GCGGTAGCAC AAACAATCTC CTACGAAGTA AGCCTCGGCT TGATCCTCTT
AGCACTCCGA GCGGTGGCAC AAACAATCTC CTACGAAGTA AGCCTCGGCT TGATCCTCTT
```

```
AGCACTTCGA GCGGTGGCAC AAACAATCTC CTATGAAGTA AGCCTCGGCT TGATCCTCTT
AGCACTTCGA GCGGTGGCAC AAACAATCTC CTATGAAGTA AGCCTCGGCT TGATCCTCTT
GGCACTTCGA GCAGTGGCAC AAACAATCTC CTATGAAGTA AGCCTGGGCT TAATTCTCTT
GGCACTCCGA GCGGTGGCAC AAACAATCTC CTATGAAGTA AGCCTGGGCT TAATTCTCTT
GGCACTCCGA GCGGTGGCAC AAACAATCTC CTATGAAGTA AGCCTGGGCT TAATTCTCTT
GGCACTCCGA GCGGTAGCAC AAACAATTTC CTACGAAGTA AGCCTCGGCT TAATCCTCTT
GGCACTTCGA GCAGTAGCAC AAACAATCTC CTATGAAGTA AGCCTTGGCT TAATCCTCTT
GGCACTTCGA GCGGTAGCAC AAACAATCTC CTATGAAGTA AGCCTTGGCT TGATCCTCTT
GGCACTTCGA GCGGTAGCGC AAACAATCTC CTACGAAGTA AGCCTTGGCT TGATCCTCTT
AGCACTCCGA GCGGTAGCAC AAACAATCTC CTACGAAGTA AGCCTTGGCT TGATCCTCTT
AGCACTCCGA GCGGTAGCAC AAACAATCTC CTACGAAGTA AGCCTTGGCT TGATCCTCTT
GGCACTTCGA GCGGTAGCGC AAACAATCTC CTACGAAGTA AGCCTTGGCT TGATCCTCTT
GGCACTTCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCCTAGGCC TAATTCTCCT
AGCACTTCGA GCAGTAGCAC AAACAATCTC CTACGAAGTA AGCTTAGGCC TAATTCTCCT

CTCACTCATT ATATTTACAG GGGCCTTTAC ACTAATTACA TTCAACACCG CACAAGAAGC
TTCCCTAATT ATATTTACCG GAGCATTTAC ACTAACTATA TTCAACACCG CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC TCTACACACC TTCAATACTA CACAAGAGGC
ATGTATAATT GTCTTCACTG GCAATTTCAC CCTCCACACC TTCAACATCA AACAAGAGGC
GTCCATGATT ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAATGTTA CACAAGAGGC
GTCCATGATT ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAATGTTA CACAAGAGGC
GTCCATGATT ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAATGTTA CACAAGAGGC
ATGTATAATC ATCTTCACTG GCAATTTCAC CCTCCATACC TTTAATGTTA CACAAGAAGC
ATGTATAATT ATCTTCACCG GCAACTTCAC CCTCCATACC TTTAATGTTA CACAAGAGGC
ATGTATAATT ATCTTCACCG GCAACTTCAC CCTCCATACC TTTAATGTCA CACAAGAGGC
ATGTATAATC ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAACATTA CACAAGAGGC
ATGTATAATC ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAACATTA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAATATTA CACAAGAAGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTCCACACC TTCAATATTA CACAAGAGGC
ATGTATGATT ATCTTCACTG GCAATTTTAC CCTCCACACC TTCAATACCA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTTAC CCTCCACACC TTCAATACCA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTTAC CCTCCACACC TTCAATACCA CACAAGAGGC
ATGTATAATT GTCTTCACTG GCAATTTCAC CCTCCACACC TTCAATGTTA CACAAGAGGC
ATCTATAATT ATCTTCACCG GCAATTTCAC CCTCCACACC TTCAATGTCA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAACTTTAC CCTCCACACC TTCAATATTA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTACACACC TTCAATATCA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTACACACC TTCAATATTA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTACACACC TTCAATATTA CACAAGAGGC
ATGTATAATT ATCTTCACTG GCAATTTCAC CCTACACACC TTCAATATTA CACAAGAGGC
ATGCATGATT GTCTTCACCG GCAATTTTAC CCTTCACACC TTTAACACCG CACAAGAAGC
ATGCATAATC ATCTTCACTG GCAACTTTAC CCTTCACACC TTCAACATCA CACAAGAAGC

CGTATGATTA ATTTTACCAG CCTGACCATT AGCCGCAATA TGATTTATTT CTACCCTAGC
AGTATGACTT ATTTTACCAG CCTGACCACT AGCCACAATA TGATTTATCT CCACCTTAGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTA CTAGCCCCAG GCTGACCACT AGCAGCAATA AGAAACATCT CTACCCTCAC
AATCTGACTA CTAGCCCCGG GCTGACCCCT CGCAGCAATA TGATATATCT CCACCCTCGC
AATCTGACTA CTAGCCCCGG GCTGACCCCT CGCAGCAATA TGATATATCT CCACCCTCGC
AATCTGACTA CTAGCCCCGG GCTGACCCCT CGCAGCAATA TGATATATCT CCACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CCACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CCACCCTCGC
AATTTGGCTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGGTATATCT CCACCCTCGC
AATTTGGCTG CTAGCCCCGG GCTGACCCCT CGCAGCAATA TGGTATATCT CCACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCACT CGCAGCAATA TGATATATCT CCACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATATATCT CCACCCTCGC
AATCTGACTA CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATATATCT CCACCCTAGC
AATCTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CCACCCTAGC
AATCTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CCACCCTAGC
AATTTGACTA CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CCACCCTAGC
GATTTGACTA CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCCCCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATTTGACTG CTAGCACCAG GCTGACCCCT CGCAGCAATA TGATACATCT CTACCCTCGC
AATCTGACTA CTCGCCCCAG GCTGACCCCT CGCAGCAATA TGGTATATTT CCACCCTTGC
```

```
AATCTGACTA CTCGCCCCAG GCTGACCCCT CGCAGCAATA TGATTCATCT CTACCCTAGC

TGAGACAAAC CGTGCCCCGT TTGACCTTAC AGAAGGAGAA TCCGAACTTG TCTCCGGGTT
AGAAACAAAC CGAGCCCCAT TTGACCTCAC AGAAGGAGAA TCTGAACTAG TTTCGGGCTT
CGAAACAAAC CGAGCCCCTT TCGATCTCAC AGAGGGGGAA TCAGAACTGG TCTCCGGCTT
AGAAACAAAC CGAGCACCAT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGATT
CGAAACAAAC CGAGCCCCAT TCGACCTCAC AGAAGGAGAA TCAGAACTAG TTTCCGGCTT
CGAAACAAAC CGAGCCCCAT TCGACCTCAC AGAAGGAGAA TCAGAACTAG TTTCCGGCTT
CGAAACAAAC CGAGCCCCAT TCGACCTCAC AGAAGGAGAA TCAGAACTAG TTTCCGGCTT
CGAAACAAAC CGAGCCCCAT TCGACCTAAC AGAAGGAGAA TCAGAGCTAG TCTCTGGCTT
CGAAACAAAC CGAGCCCCAT TCGATCTAAC GGAAGGGGAA TCAGAGCTCG TCTCCGGCTT
CGAAACAAAC CGAGCCCCAT TCGATCTAAC GGAAGGGGAA TCAGAGCTCG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TTTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAGGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAGGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCGT TTGACCTCAC AGAAGGAGAA TCAGAACTAG TCTCTGGCTT
CGAAACAAAC CGAGCCCCAT TTGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGATT
CGAAACAAAC CGAGCTCCAT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAGGGGGAA TCAGAACTAG TCTCCGGCTT
CGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAGGGGGAA TCAGAACTAG TCTCCGGCTT
AGAAACAAAC CGAGCCCCTT TCGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCCGGCTT
AGAAACAAAC CGAGCCCCAT TTGACCTCAC AGAAGGGGAA TCAGAACTAG TCTCTGGCTT
CGAAACAAAC CGAGCCCCAT TTGACCTCAC AGAGGGAGAA TCAGAACTAG TCTCCGGCTT

TAACGTCGAG TATGCAGGAG GCCCCTTCGC CCTATTTTTC CTTGCCGAGT ACACAAATAT
TAATGTTGAA TATGCAGGAG GACCCTTCGC CCTTTTCTTC CTTGCAGAAT ACACAAATAT
CAACGTAGAA TATGCAGGAG GACCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGGG GGCCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGGG GGCCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGGG GGCCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGGG GGCCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAATGTAGAG TACGCAGGAG GGCCATTCGC CCTATTCTTC CTAGCTGAGT ACGCCAATAT
CAATGTAGAA TACGCAGGAG GGCCGTTCGC CCTATTTTTC CTGGCTGAGT ACGCCAATAT
CAATGTAGAA TACGCAGGAG GGCCGTTCGC CCTATTTTTC CTGGCTGAGT ACGCCAATAT
CAACGTAGAA TATGCAGGAG GGCCATTCGC CCTATTTTTT CTAGCTGAAT ACGCCAATAT
CAACGTAGAA TATGCAGGAG GGCCATTCGC CCTATTTTTT CTAGCTGAAT ACGCCAATAT
CAACGTAGAA TACGCAGGGG GACCATTCGC CCTATTTTTT CTAGCTGAAT ACGCCAACAT
CAACGTAGAG TACGCAGGAG GACCATTCGC CCTATTTTTT CTAGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGAG GACCATTCGC CCTATTTTTT CTAGCTGAAT ACGCCAACAT
CAACGTAGAA TATGCAGGAG GGCCATTCGC CCTATTTTTT CTAGCCGAAT ATGCCAACAT
CAACGTAGAA TATGCAGGAG GGCCATTCGC CCTATTTTTT CTAGCCGAAT ATGCCAACAT
CAACGTAGAA TATGCAGGGG GACCCTTCGC CCTGTTTTTC TTGGCTGAAT ACGCCAACAT
CAACGTAGAA TACGCAGGGG GGCCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TATGCGGGAG GACCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TATGCGGGAG GACCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
TAACGTAGAA TATGCAGGAG GACCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
TAACGTAGAA TATGCAGGAG GACCCTTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
TAACGTAGAA TATGCAGGAG GACCATTCGC CCTATTTTTC TTAGCTGAAT ACGCCAACAT
CAACGTAGAA TATGCAGGAG GACCATTCGC CCTATTCTTC CTAGCCGAAT ACGCCAACAT
CAATGTAGAA TATGCCGGAG GGCCATTCGC CCTATTCTTC CTAGCCGAAT ACGCCAACAT

CTTATTTATA AACGCCTTAT CCGCAATTCT ATTCCTCGGA TGAGCCCTTA CATCTTATTT
CTTACTAATA AATGCTCTTT CTGCTATTCT ATTCCTCGGA TGAGCCCATA CATCCTATCA
CCTCCTAATA AACACCCTCT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTCTTAGGA TGAACCCATA CGTACTTGCC
CCTGCTAATA AATACCCTTT CCACAACCCT GTTTCTAGGA TGAACCCATA CGTACTTGCC
CCTGCTAATA AATACCCTTT CCACAATCCT GTTTCTAGGA TGAACCCATA CGTACTTGCC
CCTGCTAATA AATACCCTTT CCACAATCCT GTTTCTAGGA TGAACCCATA CGTACTTGCC
TCTCCTAATA AACACCCTCT CCACAATCCT ATTCTTAGGA TGAACCCCTA CGTACTTGCC
TCTCCTAATA AACACCCTCT CCACAATCCT ATTCTTAGGA TGAACCCCTA CGTACTTGCC
TCTCCTAATA AACACCCTCT CCACAATCCT ATTCTTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACCATCCT ATTCCTGGGA TGAACCCCTA CGTACTTGCC
```

CCTCCTAATA AACACCCTCT CCACCATCCT ATTCCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACAATCCT ATTTCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACAATCCT ATTCCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACAATCCT ATTCCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACAATCCT ATTCCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTAATA AACACCCTCT CCACAATCCT ATTCCTTGGA TGAACCCCTA CGTACTTGCC
CCTCTTGATA AACACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTTCTGGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AATACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
CCTCCTGATA AACACCCTTT CCACAATCCT ATTCCTAGGA TGAACCCCTA CGTACTTGCC
TCTTCTAATA AACACCCTCT CCACAATCCT ATTCCTCGGA TGAACCCCTA CGTATTTGCC
CCTCCTAATA AATACCCTCT CCGCAATCCT ATTCCTCGGA TGAACCCCTA CGTACTTGCC

ATTATACTTA CAAGCCTAGG ACTAGGCACT ACTCTAACTT TTGCCAGCTC CAACTGACTT
ACAATACTTA TTAGCCTAGG ATTAGGAACA ACTCTAACCT TTGCTAGCTC CAACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCTTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
GTCCTGCTTT CAAGCCTGGG AATTGGAACC ACCCTAACAT TTGTAAGCTC CCACTGACTC
GTCCTGCTTT CAAGCCTGGG AATTGGGACC ACCCTAACAT TTGTAAGCTC CCACTGACTC
GTCCTGCTTT CAAGCCTGGG AATTGGAACC ACCCTAACAT TTGTAAGCTC CCACTGACTC
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTCT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTCT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
GTCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTGACAC TTGCAAGCTC CCACTGACTT
GTCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTGACAC TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACC ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCACTGACTT
ATCCTGCTTT CAAGTCTAGG AATCGGAACC ACCCTAACAT TTGCAAGCTC CCACTGACTC
ATCCTGCTCT CAAGCCTAGG AATTGGAACT ACCCTAACAT TTGCAAGCTC CCATTGACTC

CTCGCATGAA TAGGCCTAGA AATTAATACT CTTGCCATTA TTCCCCTCAT AGCCCACCAC
CTTGCATGAA TAGGTTTAGA AATTAATACT CTAGCTATTA TTCCCCTTAT AGCCCGCCAC
TTAGCATGAA TGGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCTTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA AATCAGTACA ATAGCCATCA TCCCCCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA GATCAGTACA CTAGCCATCA TCCCTCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA GATCAGTACA CTAGCCATCA TCCCTCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA GATCAGTACA CTAGCCATCA TCCCTCTAAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCTTAAT AGCACAACAG
TTGGCATGAA TGGGCCTAGA AATCAGCACA CTGGCCATCA TTCCCTTAAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA AATCAGCACA CTGGCCATCA TCCCCTTAAT AGCACAACAA
TTAGCATGAA TGGGACTAGA GATCAGTACA CTAGCCATCA TCCCCCTTAT AGCACAACAA
TTAGCATGAA TGGGACTAGA GATCAGTACA CTAGCCATCA TCCCCCTTAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA GATTAGTACA TTAGCCATCA TCCCCCTCAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA GATTAGTACA TTAGCCATCA TCCCCCTCAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA GATTAGTACA CTAGCCATCA TCCCCCTTAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA GATTAGTACA CTAGCCATCA TCCCCCTTAT AGCACAACAA
TTGGCATGAA TGGGCCTAGA GATTAGTACA CTAGCCATCA TCCCCCTTAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA GATTAGTACA CTAGCCATCA TCCCCCTAAT AGCACAACAA
TTGGCATGAA TAGGCCTAGA AATCAATACA CTAGCCATCA TCCCCCTGAT AGCACAACAA
TTAGCATGAA TGGGCCTAGA AATCAGTACG CTAGCCATCA TCCCCTTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCCTAAT AGCACAACAA
TTAGCATGAA TAGGCCTAGA AATCAGTACA CTAGCCATCA TCCCCCTAAT AGCACAACAA

```
CTAGCATGAA TAGGCCTAGA AATCAACACA CTAGCCATTA TCCCACTAAT AGCACAACAA
CTAGCATGAA TAGGCCTAGA AATCAGTACT TTAGCCATCA TCCCCCTAAT AGCGCAACAA

CATCACCCAC GAGCCGTAGA AGCCGCAACA AAATACTTCA TTACACAAGC TGCAGCCGCA
CACCACCCAC GAGCTGTTGA AGCCGCAACA AAATACTTTA TTACTCAAGC TGCAGCTGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTCC TCACCCAAGC AACAGCCGCA
CATCACCCTC GAGCAGTCGA AGCCACAACT AAATATTTTC TCACCCAAGC CACAGCCGCA
CATCACCCCC GAGCAGTAGA AGCCACAACC AAATATTTTC TTACCCAAGC CACAGCCGCA
CATCACCCCC GAGCAGTAGA AGCCACAACC AAATATTTTC TTACCCAAGC CACAGCCGCA
CATCACCCCC GAGCAGTAGA AGCCACAACC AAATATTTTC TTACCCAAGC CACAGCCGCA
CACCACCCTC GAGCAGTAGA AGCCACAACT AAATATTTTC TCACCCAAGC AACAGCCGCA
CACCACCCTC GAGCAGTCGA AGCCACAACT AAATATTTTC TCACCCAAGC GACGGCCGCA
CACCACCCTC GAGCAGTCGA AGCCACAACT AAATATTTTC TCACCCAAGC AACGGCCGCA
CACCACCCCC GAGCAGTCGA AGCCACAACT AAATACTTCC TCACCCAAGC AACGGCCGCA
CACCACCCCC GAGCAGTCGA AGCCACAACT AAATACTTCC TCACCCAAGC AACGGCCGCA
CACCACCCCC GAGCAGTCGA AGCCACGACT AAATACTTCC TCACCCAAGC AACAGCCGCA
CACCACCCCC GAGCAGTCGA AGCCACAACT AAATACTTCC TCACCCAAGC AACAGCCGCA
CACCACCCCC GAGCAGTCGA AGCCACAACT AAATACTTCC TCACCCAAGC AACAGCTGCA
CACCACCCCC GAGCAGTTGA AGCCACAACC AAATATTTCC TCACCCAAGC AACAGCTGCA
CACCACCCCC GAGCAGTTGA AGCCACAACT AAATACTTCC TCACCCAAGC AACAGCTGCA
CATCACCCTC GAGCAGTCGA AGCCACAACT AAATATTTTC TCACCCAAGC CACGGCCGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTTC TCACCCAGGC GACAGCCGCA
CACCACCCTC GAGCAGTCGA AGCCACAACT AAATATTTTC TTACCCAAGC AACAGCCGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTTC TCACCCAAGC AACAGCCGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTTC TCACCCAAGC AACAGCCGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTTC TCACCCAAGC AACAGCCGCA
CATCACCCTC GAGCAGTCGA GGCCACAACT AAATACTTTC TCACCCAAGC AACAGCCGCA
CACCACCCCC GAGCAGTTGA GGCCACAACC AAGTACTTTC TAACCCAAGC AACGGCCGCA
CACCACCCCC GAGCAATTGA AGCCACAACC AAATACTTTC TCACCCAAGC AACGGCCGCA

GCTTTACTTC TGTTTGCCGG CCTATTTAAT GCTTGACAGT CCGGACAATG ACTTATCCAA
GCACTACTTC TATTCACTGG CTTATTTAAT GCCTGACAAT CAGGACAATG ACTAATTCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAGCAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GACACAAAAA CAGCAGAATG AAACATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAATAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAATAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAATAA CAGGAGAATG AAATATCCAA
GCTATAATCT TATTTGCCAG TACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATCT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATCT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG GGGTATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG GAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG AAACATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAATAA CAGGAGAATG AAATATCCAA
GCTATGATTT TATTTGCCAG CACCACCAAT GCTTGAATAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAGTAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAACAA CGGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG CACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAC GCTTGAACAA CAGGAGAATG GAATATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAATAA CAGGAGAATG AAACATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCTATAATTT TATTTGCCAG TACCACCAAT GCTTGAACAA CAGGAGAATG AAATATCCAA
GCCATAATTT TATTTGCTAG CACTACAAAT GCCTGATTGA CAGGAGAATG AAACATTCAA
GCTATAATCT TATTTGCCAG TACTACAAAT GCCTGAGTGA CGGGAGAATG AAACATTCAA

GACAT-TACA ATACC--AGT CTCCACCCTT ATGACTCTAG CAATTGCCAT TAAACTCGGG
GATATATCAA TACC---AGT ATCTGCACTT ATAACACTAG CAATTGCCAT CAAACTAGGC
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAAATTGGA
GAAATATCCA AACCCACACA CTTAGCCCTA AGTACCATCC CCCGAACCCT AAAAATCGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAGATTGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCCTAG CCCTGGCCCT AAAGATTGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAGATTGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAAATCGGG
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAAATTGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAAATTGGA
```

```
GAAATATCCA ACCCCACAGC CTTAGCCCTA ATCACCATGG CCCTGGCCCT AAAAATAGGA
GAAATATCTA ACCCCACAGC CTTAGCCCTA ATCACCATGG CCCTGGCCCT AAAAATAGGA
GAGATGTCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAGATCGGA
GAAATGTCCA ACCCCACAGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAGATCGGA
GAGATATCCA ACCCCACAGC CTTAACCCTA ATCACCATGG CCCTAGCCCT AAAAATCGGA
GAGATATCCA ACCCCATAGC CTTAGTCCTA ATTACCATGG CCCTGGCCCT AAAAATCGGA
GAGATATCCA ACCCCATAGC CTTAGTCCTA ATTACCATGG CCCTGGCCCT AAAAATCGGA
GAAATATCTA ACCCCACAGC CTTAGTCCTA ATTACCATAG CCTTAGCCCT AAAAATTGGA
GAAATATCCA ACCCCACGGC CTTAGTCCTA ATCACCATAG CCCTGGCCCT AAAAATTGGA
GAGATATCCA ATCCCACAGC CTTAGTCCTA ATTACCATAG CCCTGGCCCT AAAAATTGGA
GAAATGTCCA ACCCCACAGC CTTAGTCCTA ATTACCATGG CCCTGGCCCT AAAAATTGGA
GAAATGTCCA ACCCCACAGC CTTAGTCCTA ATTACCATAG CCCTGGCCCT AAAAATTGGA
GAAATGTCCA ACCCCACAGC CTTAGTCCTA ATTACCATAG CCCTGGCCCT AAAAATTGGA
GAAATATCCA ACCCCACAGC CTTAGTCCTA ATTACCATAG CCCTGGCCCT AAAAATTGGA
GAAATATCCA ACCCCACAGC CCTAACCTTA ATCACCATAG CCCTGGCCCT TAAAATCGGA
GAAATATCTA ACCCCACAGC CCTAGCCTTA ATCACCATAG CCCTAGCCCT TAAAATCGGG

ATTGCCCCAA TACACTTCTG ACTGCCCGAA GTCATACAAG GAATCACATT AAATACAGGA
GTAGCCCCTA TACACTTCTG ATTACCAGAA GTAATACAAG GAATTACACT AAATACAGGA
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGT
CTAGCCCCCG TCCACTACTT ACTCCCAGAA GTACTACAAG GCCTCGACCT CACCACGGGT
CTCGCCCCCG TCCACTACTG ACTCCCAGAG GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTCCCAGAG GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTCCCAGAG GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTTCCAGAA GTCCTACAAG GCCTCGATCT CACCACAGGA
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTTCTACAAG GCCTCAATCT CACCACAGGA
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTTCTACAAG GCCTCAATCT CACCACAGGA
CTCGCCCCCG TCCACTACTG ACTTCCAGAA GTACTACAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTTCCAGAA GTACTACAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTTCCGGAA GTACTGCAAG GCCTCGACCT CACCACAGGC
CTAGCCCCCG TCCACTACTG ACTTCCAGAG GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTTCCAGAG GTACTGCAAG GCCTCGACCT CACCACAGGC
CTCGCCCCCG TCCACTACTG ACTCCCAGAG GTACTGCAAG GCCTCGACCT CACCACGGGT
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTACTACAAG GCCTCGACCT CACCACAGGT
CTCGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GTCTCGACCT CACCACAGGT
CTAGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGT
CTAGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGT
CTAGCCCCCG TCCACTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGT
CTAGCCCCCG TCCATTACTG ACTCCCAGAA GTACTGCAAG GCCTCGACCT CACCACAGGT
CTTGCCCCCA TCCACTACTG ATTGCCAGAA GTATTACAAG GCCTTGACCT CACCACAGGA
CTCGCCCCTG TCCACTACTG ACTGCCAGAA GTACTACAAG GCCTCGACCT CACCACAGGA

CTAATCTTAG CCACCTGACA AAAACTCGCA CCATTAGCAC TCCTATACCA AATT--TCCA
CTAATCTTAG CCACCTGACA AAAACTTGCA CCTCTAGCAC TCCTCTACCA AATT--TCTA
CTCATCCTCT CTACCTGACA AAAACTGGCC CCATTTGCCC TAATTTATCA AATCAGTCCA
CTTATCCTCT CGACCTGACA AAAACTGGCC CCATTCGCCC TAATTTATCA AATTAGCCCA
CTCATCCTCT CAACCTGACA AAAACTGGCC CCATTTGCCC TAATTTACCA AATCAGCCAA
CTCATCCTCT CAACCTGACA AAAACTGGCC CCATTTGCCC TAATTTACCA AATCAGCCAA
CTCATCCTCT CAACCTGACA AAAACTGGCC CCATTTGCCC TAATTTACCA AATCAGCCAA
CTTATCCTCT CAACCTGACA AAAACTAGCC CCCTTTGCCT TAATTTATCA AATTAGCCCA
CTTATCCTCT CGACCTGACA AAAACTAGCC CCCTTTGCCT TAATTTATCA GATTAGCCCA
CTTATCCTCT CGACCTGACA AAAACTAGCC CCCTTTGCCT TAATTTATCA GATTAGCCCA
CTCATCCTCT CAACCTGACA GAAACTAGCC CCGTTTGCCC TAATTTATCA AATCAGCCCA
CTCATCCTCT CAACCTGACA GAAACTGGCC CCGTTTGCCC TAATTTATCA AATCAGTCCA
CTCATCCTTT CAACCTGACA AAAACTAGCC CCATTCGCCC TAATTTATCA AATCAGTCCA
CTCATCCTTT CAACCTGACA AAAACTAGCC CCCTTCGCCC TAATTTATCA AATCAGTCCA
CTCATCCTCT CAACCTGACA AAAACTAGCC CCATTTGCCC TAATCTATCA AATTAGCCCA
CTCATCCTCT CAACCTGACA AAAACTAGCC CCATTTGCCC TAATTTATCA AATTAGCCCA
CTCATCCTCT CAACCTGACA AAAACTAGCC CCATTTGCCC TAATTTATCA AATTAGCCCA
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTCGCCC TTATTTATCA AATTAGCCCA
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTTGTCC TGATTTATCA AATCAGCCCA
CTCATCCTTT CAACCTGACA AAAACTGGCT CCATTTGCCC TAATTTATCA AATCAGCCCA
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTTGCCC TAATCTACCA AATCAGTCCA
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTTGCCC TAATTTATCA AATCAGTCCA
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTTGCCC TAATTTATCA AATCAGTCCA
```

```
CTCATCCTCT CGACCTGACA AAAACTGGCC CCATTTGCCC TAATTTATCA AATCAGTCCA
CTTATCCTCT CTACCTGACA AAAACTAGCC CCATTCGCCC TAATTTACCA AATCAGCCCA
CTCATCCTCT CCACCTGACA AAAACTAGCC CCATTTGCCC TAATTTACCA AATTAGCCCA

GCAACCTCAT ACCAGAACTA ACAATTACCC TAGGCCTAAT ATCCACAATT ATTGGTGGAT
ATAACCTCAT ACCAGAACTC ATAATTGCCC TAGGACTAAT ATCTACAATT ATTGGTGGCT
ATAA--TAAA CCCAACTCTG ATAATTATGC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACCCTA GTAATTATAT TAGGACTGGC CTCCGCCATC ATTGGTGGAT
GTAA--TGAA CCCAACTCTA ATAATCACAC TAGGCCTTTC CTCCGCCGTC ATCGGCGGAT
GTAA--TGAA CCCAACTCTA ATAATCACAC TAGGCCTTTC CTCCGCCGTC ATCGGCGGAT
GTAA--TGAA CCCAACTCTA ATAATCACAC TAGGCCTTTC CTCCGCCGTC ATCGGCGGAT
ATAA--TAAA CCCAACTCTA GTAATCATAC TAGGCCTAGC CTCCACCATC ATTGGAGGAT
ATAA--TAAA CCCAACTCTA GTAATTACAC TAGGCCTGGC CTCCACTATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTA GTAATTACAC TAGGCCTGGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCGTCCCTA ATAACCATAC TAGGCCTAGC CTCCACCATC ATCGGCGGGT
ATAA--TAAA CCCATCCCTG ATAACCATAC TAGGCCTAGC CTCCACCATC ATCGGCGGGT
ATAA--TAAA TCCATCCCTG GTAATCATAC TAGGCCTAGC CTCCACCATC ATCGGCGGGT
ATAA--TAAA CCCATCCCTG GTAATCATAC TAGGCCTAGC TTCCACCATC ATCGGCGGAT
ATAA--TAAA CCCAACCCTA GTAATCATAT TAAGCCTGGC CTCCACCATC ATCGGCGGAT
ATAA--TAAA CCCAACCCTA GTAGTTATAT TAAGCCTAGC CTCCACCATC GTTGGCGGAT
ATAA--TAAA CCCAACCCTA GTAGTTATAT TAAGCCTAGC CTCCACCATC GTCGGCGGAT
ATAA--TAAA CCCAACCCTA GTAATTATGC TAGGCCTGGC CTCCGCCATC ATTGGTGGAT
ATAA--TAAG CCCAACTCTA CTAATTATAC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTA GTAATTATGC TGGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTG GTAATTATGC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTG GTAATTATAC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTG GTAATTATAC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCAACTCTG GTAAATATGC TAGGCCTAGC CTCCACCATC ATTGGCGGAT
ATAA--TAAA CCCCACCCTT ATAATCACCC TAGGCCTAGC CTCCACCATC GTTGGCGGAT
ATAA--TAAA CCCAACCCTT GTAGTCATCC TAGGCCTAAC CTCCACCATT ATTGGCGGAT

GAGGAGGACT TAATCAGACA CAAATTCGAA AAATCATAGC CTACTCATCC ATTGCCCACT
GAGGTGGACT TAATCAAACT CAAATCCGAA AAATTATAGC CTACTCATCC ATTGCACACT
GAGGCGGCCT AAACCAGACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GAGGCGGCCT AAACCAAACA CCACTACGCA AAAGCCTCGC ATACTCATGA ATCACCCACA
GGGGCGGCCT AAACCAAACG CAACTACGAA AAATCCTAGC ATATTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACG CAACTACGAA AAATCCTAGC ATATTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACG CAACTACGAA AAATCCTAGC ATATTCATCA ATCGCCCACC
GAGGTGGCCT AAATCAGACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCATC
GAGGTGGCCT AAATCAGACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCATC
GAGGTGGCCT AAATCAGACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCATC
GAGGCGGCCT AAACCAAACG CAACTACGAA AGATCCTAGC ATACTCATCA ATCGCCCACC
GAGGCGGCCT AAACCAAACG CAACTACGAA AGATTCTAGC ATACTCATCA ATCGCCCACC
GAGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTGGC ATACTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTGGC ATACTCATCA ATCGCCCACC
GAGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GAGGTGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATATTCATCA ATCGCCCATC
GAGGTGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATATTCATCA ATCGCCCATC
GAGGCGGCCT AAACCAAACA CAACTGCGAA AAATCCTAGC ATACTCATCA ATTGCCCACC
GAGGCGGCCT AAACCAGACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GAGGCGGCCT AAACCAGACA CAGCTACGAA AAATCCTAGC ATACTCATCA ATCGCTCACC
GGGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GGGGCGGCCT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCA ATCGCCCACC
GAGGCGGATT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCC ATCGCTCACC
GAGGCGGTTT AAACCAAACA CAACTACGAA AAATCCTAGC ATACTCATCG ATTGCCCACC

TAGGATGAAT TATTTCCATC ATGCACTTCA TACCTGCCTT AGCCATTATT AATTTAACTA
TAGGCTGAAT CATCTCAATT ATACACTTCA TACCAACCCT AGCAATTATT AACCTAACCA
TGGGTTGAAT AATGATCGTC ATACAGTATT CCCCAAACCT TGCTATCCTA AACCTAATCC
TAGGTTGAAG AATGATTGTT ATGCAGTATT CCCCAAACCT TGCCATCCTG AACCTAATCC
TAGGCTGAAT GATGATCGTC ATACAGTACT CCCCAAACCT CGCCATCCTG AACCTAATCC
TAGGCTGAAT GATGATCGTC ATACAGTACT CCCCAAACCT CGCCATCCTG AACCTAATCC
TAGGCTGAAT GATGATCGTC ATACAGTACT CCCCAAACCT CGCCATCCTG AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAATATT CCCCAAATCT TGCCATTCTA AACCTAATCC
TAGGCTGAAT AATGATTGTC ATACAATATT CCCCAAACCT TGCCATTCTA AACCTAATCC
```

TAGGCTGAAT AATGATTGTC ATACAATATT CCCCAAACCT TGCCATTCTA AACCTAATCC
TGGGTTGGAT AATGATTATT ATACAGTACT CCCCAAACCT TGCCATTCTA AACCTAATCC
TGGGTTGGAT AATGATTATT ATACAGTACT CCCCAAACCT TGCCATTCTA AACCTAATCC
TGGGTTGGAT AATGATTGTT ATGCAGTACT CCCCAAATCT TGCCATTCTA AACCTAACCC
TGGGTTGGAT AATGATTGTT ATACAGTACT CCCCAAATCT TGCCATTCTA AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAATATT CCCCGAATCT TGCCATTCTA AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAGTACT CCCCGAATCT TGCCATTCTA AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAGTATT CCCCGAATCT TGCCATTCTA AACCTAATCC
TAGGCTGAAT AATGATTGTT ATGCAGTATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGCTGAAT AATGATCGTC ATACAGTATT CCCCTAATCT TGCTATCCTA AACCTAATCC
TGGGCTGAAT AATGATTGTT ATACAATATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGTTGAAT AATGATTGTC ATACAGTATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAGTATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGTTGAAT AATGATTGTT ATACAGTATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGTTGAAT AATGATTGTC ATACAGTATT CCCCAAACCT TGCCATCCTA AACCTAATCC
TAGGCTGAAT AATAGTTATC ATACAATACT CCCCCAGCCT CGCCATACTT AACCTAACCC
TAGGCTGAAT AATAATCATT ATACAATACT CCCCTAACCT CGCCCTACTA AACCTAACCC

TTTACATTAT CATAACAACA ACAATATTTA TAACCTTCAA CACATTAAAC TCTACTTCAA
TGTATATCAT CTTAACTACA ACTATATTCA TAATCTTTAA TACCATAAAC TCTACCACTA
TATATATTAC CATAACTACT GCGACCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TAGATATTAC CATAACTACT GCAACCTTCC TAACATTCAC ACATATGGCC TCCACAAGAC
TGTATATCAC TATAACCACC GCAACCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TGTATATCAC TATAACCACC GCAACCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TGTATATCAC TATAACCACC GCAACCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TATACATTAT TATGACCACC GCAACCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATACATTGC TATAACCACC GCAACCTTCC TAACATTCAA AAATGTGACC TCCACAAAAC
TATACATTGC TATAACCACC GCAACCTTCC TAACATTCAA AAATGTGACC TCCACAAAAC
TATATATTAC CATAACTTCC GCGGCCTTCC TGACATTCAA GAATGTGGCC TCCACAAAAC
TATATATTAC CATAACTTCC GCGGCCTTCC TGACATTCAA GAATGTGGCC TCCACAAAAC
TATATATTAC TATGACCTCT GCGGCCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATATATTAC TATGACCTCT GCGGCCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATATATTAC TATGACTTCT GCGGCCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TATATATTAC TATAACTTCT GCGGTCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TATATATTAC TATAACTTCT GCGGTCTTCC TAACATTCAA AAATGTAGCC TCCACAAAAC
TATACATCAC TATAACCACT GCGACCTTCC TAACGTTCAA AAATGTGGCC TCCACAAAAC
TATATATTAC TATAACCACT GCGACCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATACATTAC TATAACCACT GCGACCTTCC TAACATTCAA AAATGTAGCC TCTACAAAAC
TATATATTAC CATAACTACT GCGACCTTCC TAACATTCAA AAACGTGGCC TCCACAAAAC
TATATATTAC CATAACTACT GCGACCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATATATTAC CATAACTACT GCGACCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TATATATTAC CATAACTACT GCGACCTTCC TAACATTCAA AAATGTGGCC TCCACAAAAC
TCTACATTAT CATAACCACC GCAACTTTCC TGACATTCAA AAATGTGACC TCCACAAAAC
TTTACATTAT CATGACCGCC GCAGCCTTCC TGACATTCAA AAATGTAGCC TCCACAAAAC

TTAATACACT GGCCACAAAC TGATCCAAAT TCCCAGTCCT GTCCGCCGCC ACCATGCTTG
TTAACTCCAT AGCCCTAAAC TGATCAAAAT TCCCTATCCT ATCTGCCATT GCAATACTTG
TAAGTACGCT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACGCT GACCCGCACC TCACCAACAA CCCCAATAAC AACCACAAGA ACAATAATAA
TAAGTACACT CGCCCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT CGCCCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT CGCCCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACGCT AACCCTCACC TGATCAAAAA CCCCCGTAAT AACCACGATG GCAATAATAA
TAAGTACACT AACCCTCACC TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT AACCCTCACC TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACCCT AACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT AACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT AACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTATACT AACTCTCACT TGATCAAAAA CCCCCATAAT GACCACAATG GCAATAATAA
TAAGTATACT AACTCTCACT TGATCAAAAA CCCCCATAAT GACCACAATG GCAATAATAA
TAAGTACGCT AGCTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TGAGTACACT AACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACACT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAATGCACT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACGCT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA

TAAGTACGCT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAGTACGCT GACTCTCACT TGATCAAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAACGCACT AGCCCTTACC TGATCGAAAA CCCCCATAAT AACCACAATA GCAATAATAA
TAAACACATT AACCCTCACC TGATCAAAAA CCCCCATAAT GACCGCAATA GCAATAATAA

CACTCCTATC ACTAGGGTGA CCATCACCCG TTGACTCTTT TCGACCAACC ACAAAGACAT
CTCTACTATC ATTAGGGTGA CCATCACCCG CTGACTATTC TCAACAAATC ACAAAGACAT
CACTACTTTC CCTGGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGACAT
CACTACTCTC CCTAAGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CCCTGCTTTC CTTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CCCTGCTTTC CTTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CCCTGCTTTC CTTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTCTC TCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTCTC TCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTCTC TCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC TCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTGGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTGGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CTTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CACTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TTCACCAACC ACAAAGACAT
CCCTACTTTC CCTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT
CCCTCCTTTC CTTAGGGTGG CAATCACCCG TTGATTCTTT TCTACTAACC ACAAAGATAT

TGGCACCCTT TACTTAATTT TTGGTGCTTG GGCCGGAATA GTAGGAACCG CATTAAGTCT
TGGCACCCTT TATTTAATCT TTGGTGCCTG GGCCGGAATA GTAGGAACCG CATTAAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATG GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA ATCGGCACAG CCCTCAGCCT
TGGCACCCTA TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGTCT
TGGCACCCTA TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGTCT
TGGCACCCTA TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGTCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATG GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGTACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGTACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGCCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGCACAG CCCTCAGTCT
TGGCACCCTG TATTTAGTAT TTGGTGCCTG AGCAGGCATA GTCGGTACGG CCCTGAGCCT

ATTAATTCGG GCTGAACTTG GCCAACCCGG GGCCCTAATA GGAGACGACC AGATTTATAA
TTTAATTCGT GCAGAACTTG GCCAACCAGG AGCTCTAATA GGAGATGACC AAATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCAAACTAA GCCAAAACAG TGCCCTGCCT GGCGATGAAC AGATCTACAA
TCTGATCCGC GCCGAACTCA GCCAACCTGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGC GCCGAACTTA GCCAACCTGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGC GCCGAACTTA GCCAACCTGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AGATTTACAA

```
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGACGACC AGATTTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGACGACC AGATTTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCTGG TGCCTTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGACC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCTTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AGATCTACAA
TCTGATCCGT GCCGAACTGA GCCAACCCGG TGCCCTGCTT GGCGATGATC AAATCTACAA
ACTGATTCGT GCCGAACTAA GCCAACCCGG TGCCTTACTT GGCGATGATC AAATCTATAA
TCTGATTCGT GCCGAATTGA GCCAACCCGG TGCCTTACTT GGCGATGATC AAATCTATAA

TGTAATTGTC ACTGCCCACG CATTTGTTAT AATTTTCTTT ATAGTAATAC CAATTATAAT
TGTTATCGTC ACTGCTCATG CATTCGTAAT AATTTTCTTC ATAGTAATAC CAATTATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTAAT GATTTTCTTT ATAGTAATGC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTAAT GATTTTCTTT ATAGTAATGC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTAAT GATTTTCTTT ATAGTAATGC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTC ACCGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATGAT
TGTTATCGTC ACCGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATGAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATCTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATCTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTTAT GATTTTCTTT ATAGTGATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT AATTTTCTTT ATAGTGATAC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATGC CCATTATGAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATGAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTT ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTTATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT
TGTCATCGTC ACAGCCCACG CCTTTGTCAT GATTTTCTTT ATAGTAATAC CCATCATAAT

CGGGGGATTC GGCAACTGAC TTGTACCATT AATAATTGGA GCCCCTGACA TAGCCTTCCC
TGGCGGATTT GGTAACTGAC TGGTGCCACT GATAATTGGG GCCCCCGATA TAGCTTTCCC
TGGCGGATTC GGAAACTGAC TAGTCCCCCT AATAATTGGG GCCCCAGACA TGGCATTTCC
CGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGATA TAGCATTTCC
TGGAGGATTC GGAAACTGAC TAGTCCCCCT GATAATTGGA GCCCCAGACA TAGCATTCCC
TGGAGGATTC GGAAACTGAC TAGTCCCCCT GATAATTGGA GCCCCAGACA TAGCATTCCC
TGGAGGATTC GGAAACTGAC TAGTCCCCCT GATAATTGGA GCCCCAGACA TAGCATTCCC
CGGCGGTTTC GGAAACTGAC TGATTCCCTT AATAATTGGA GCCCCAGATA TGGCATTTCC
CGGCGGTTTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGATA TAGCATTTCC
CGGCGGCTTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGATA TAGCATTTCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TAGTGCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTACCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTACCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTCCCTCT AATAATCGGA GCCCCAGACA TAGCATTCCC
TGGCGGATTC GGAAACTGAT TGGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTTCC
TGGCGGATTC GGAAACTGAC TAGTCCCCCT AATAATTGGA GCCCCAGACA TGGCATTCCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGG GCCCCAGACA TGGCATTTCC
```

TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGG GCCCCAGACA TGGCATTTCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGG GCCCCAGACA TGGCATTTCC
TGGCGGATTC GGAAACTGAC TGGTCCCCCT AATAATTGGG GCCCCAGACA TGGCATTTCC
CGGCGGATTC GGAAATTGAC TGGTCCCCCT GATAATCGGA GCTCCAGATA TAGCCTTCCC
TGGCGGATTC GGAAATTGAC TAGTCCCCCT AATAATTGGA GCCCCAGACA TAGCCTTCCC

TCGTATAAAT AATATAAGTT TCTGACTCCT TCCCCCATCA CTTCTTCTTC TACTTACCTC
GCGTATAAAT AATATAAGCT TCTGACTCCT CCCACCTTCA CTTCTTCTCC TACTAACATC
TCGCATGAAC AATATAAGCT TCTGACTCCT ACCCCCATCT TTCCTGCTCC TTTTAGCCTC
TCGCATGAAC AATATGAGAT GCTGGCGCCT ACCCCCATCA TTCCTGCTCC TTTTGGCCTC
CCGTATGAAC AACATGAGCT TCTGACTTCT GCCCCCGTCC TTTCTGCTCC TTTTAGCCTC
CCGTATGAAC AACATGAGCT TCTGACTTCT GCCCCCGTCC TTTCTGCTCC TTTTAGCCTC
CCGTATGAAC AACATGAGCT TCTGACTTCT GCCCCCGTCC TTTCTGCTCC TTTTAGCCTC
TCGCATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCA TTTCTACTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCA TTTCTGCTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCA TTTCTGCTCC TTTTGGCCTC
TCGTATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTCTGGCCTC
TCGTATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTCTGGCCTC
TCGCATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATAAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TCTTGGCCTC
TCGCATAAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATAAAC AATATGAGCT TCTGGCTCCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGACTTCT ACCCCCATCC TTCCTACTCC TTTTAGCCTC
TCGCATGAAC AACATGAGCT TCTGACTTCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGACTCCT ACCCCCATCC TTCCTACTCC TTTTGGCCTC
TCGCATGAAC AATATGAGCT TCTGACTCCT ACCCCCATCC TTCCTACTCC TTTTAGCCTC
TCGCATGAAC AATATGAGCT TCTGACTCCT ACCCCCATCC TTCCTACTCC TTTTAGCCTC
TCGCATGAAC AATATGAGCT TCTGACTCCT ACCCCCATCC TTCCTACTCC TTTTAGCCTC
TCGCATGAAC AATATGAGCT TCTGACTCCT ACCCCCATCC TTCCTACTCC TTTTAGCCTC
TCGCATGAAT AATATGAGTT TCTGGCTCCT ACCCCCATCC TTCCTTCTCC TTTTGGCCTC
TCGCATGAAT AATATGAGCT TCTGGCTCCT ACCCCCATCT TTCCTTCTCC TTTTGGCTTC

CTCAGCAGTA GAAGCCGGAG TCGGAACCGG ATGAACTGTG TATCCCCCAC TAGCCGGAAA
TTCCGCTGTA GAAGCAGGGG TTGGAACCGG ATGAACAGTA TACCCACCGC TAGCAGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG ATGAACTGTT TACCCTCCGC TGGCGGGAAA
CTCTGGCGTA GAGGCCGGAG CCGGCACAGG GCGAACTGTT TACCCCCCCC TGGCGGGAAA
CTCCGGGGTA GAGGCCGGAG CTGGCACAGG GTGAACCGTC TACCCCCCAC TGGCGGGAAA
CTCCGGGGTA GAGGCCGGAG CTGGCACAGG GTGAACCGTC TACCCCCCAC TGGCGGGAAA
CTCCGGGGTA GAGGCCGGAG CTGGCACAGG GTGAACCGTC TACCCCCCAC TGGCGGGAAA
CTCTGGCGTA GAGGCCGGAG CTGGCACAGG ATGAACCGTC TACCCCCCAC TTGCGGGGAA
CTCTGGCGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTC TACCCCCCAC TGGCGGGGAA
CTCTGGCGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTC TACCCCCCAC TGGCGGGGAA
CTCTGGGGTA GAGGCCGGGG CCGGCACAGG ATGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
TTCTGGGGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGGG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGTACAGG ATGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAAGCCGGGG CCGGCACAGG ATGAACTGTC TATCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCCCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCTCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCTCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACAGG GTGAACTGTT TACCCTCCAC TGGCGGGAAA
CTCTGGGGTA GAGGCCGGAG CCGGCACCGG GTGAACTGTC TACCCCCCTC TGGCAGGAAA
CTCTGGGGTA GAAGCCGGAG CTGGCACCGG ATGAACTGTA TACCCCCCTT TAGCAGGAAA

CCTGGCACAC GCAGGTGCAT CAGTTGACCT TGCTATTTTT TCATTACATT TGGCCGGTGT
TCTAGCACAT GCAGGGGCAT CAGTTGACCT GGCTATCTTT TCACTTCACC TGGCAGGTGC
CCTGGCCCAT GCGGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CCCAACCCCA GCGGGTGCGT CCGTAGGCCC AACCCATTTT TCCCTCCACC TCGCCCGGGT
CCTAGCCCAC GCGGGGGCCT CCGTAGACCT AACCATCTTC TCCCTACACC TAGCCGGGAT
CCTAGCCCAC GCGGGGGCCT CCGTAGACCT AACCATCTTC TCCCTACACC TAGCCGGGAT
CCTAGCCCAC GCGGGGGCCT CCGTAGACCT AACCATCTTC TCCCTACACC TAGCCGGGAT

```
CCTAGCCCAT GCAGGAGCCT CTGTGGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CTTAGCCCAT GCGGGAGCCT CTGTGGACCT AACCATTTTT TCTCTTCACC TGGCCGGGGT
CTTAGCCCAT GCGGGAGCCT CTGTGGACCT AACCATTTTT TCTCTTCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCACC TGGCCGGGGT
CCTAGCCCAT GCGGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCACC TGGCCGGGGT
CCTAGCCCAT GCGGGAGCCT CTGTAGACCT AACCATTTTT TCCCTCCACC TGGCCGGGGT
CCTAGCCCAT GCGGGAGCCT CTGTAGACCT AACCATTTTT TCCCTCCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCCGGAGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTCCATC TGGCCGGGGT
CTTGGCCCAC GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCCGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CCTGGCCCAT GCAGGAGCCT CTGTAGACCT AACCATTTTC TCCCTTCACC TGGCTGGGGT
CCTAGCCCAC GCAGGAGCCT CTGTAGACCT TACCATCTTC TCCCTACACT TGGCTGGGGT
CCTGGCCCAC GCAGGAGCCT CTGTAGACCT TACTATTTTC TCCCTCCACT TGGCTGGGGT

CTCCTCAATT CTTGGGGCCA TTAACTTTAT TACTACGATT ATCAACATAA AACCGCCATC
CTCCTCAATT CTAGGGGCAA TTAACTTCAT TACTACAATT ATTAATATAA AACCCCCATC
TTCGTCCATT TTGGGGGCTA TTAATTTTAT TACTACCATT ATTAACATGA AACCCCCCGC
GTCTTACATT TTTAGGGCCA TTACTTTTAT TTACACGATT TTTAACATGA CGACCCCCGC
TTCGTCCATT TTAGGAGCTA TTAATTTTAT TACCACAATT ATTAATATGA AACCCCCCGC
TTCGTCCATT TTAGGAGCTA TTAATTTTAT TACCACAATT ATTAATATGA AACCCCCCGC
TTCGTCCATT TTAGGAGCTA TTAATTTTAT TACCACAATT ATTAATATGA AACCCCCCGC
ATCATCCATT TTAGGGGCTA TTAACTTTAT TACCACAATC ATTAACATGA AGCCCCCCGC
ATCATCCATT TTAGGGGCTA TTAACTTTAT TACCACAATC ATTAACATGA AACCCCCCGC
ATCATCCATT TTAGGGGCTA TTAACTTTAT TACCACAATC ATTAACATGA AACCCCCCGC
ATCGTCCATT TTAGGAGCTA TTAATTTTAT CACCACAATT ATTAACATGA AACCCCCCGC
ATCGTCCATT TTAGGAGCTA TTAATTTTAT CACCACAATT ATTAACATGA AACCCCCCGC
GTCGTCCATT TTAGGAGCTA TTAATTTTAT TACCACAATT ATCAACATGA AACCCCCCGC
GTCGTCCATT TTAGGAGCTA TTAATTTTAT TACCACAATT ATCAACATGA AACCCCCCGC
GTCGTCCATT TTAGGGGCAA TTAATTTTAT TACCACAATT ATTAACATGA AGCCCCCCGC
GTCGTCCATT TTAGGGGCCA TTAATTTTAT TACCACGATT ATTAACATGA AGCCCCCCGC
GTCGTCCATT TTAGGGGCCA TTAATTTTAT TACCACGATT ATTAACATGA AGCCCCCCGC
TTCGTCCATC TTGGGGGCCA TTAACTTTAT TACCACAATT ATTAACATGA AGCCCCCCGC
TTCGTCCATT CTGGGGGCTA TTAATTTTAT TACCACAATC ATTAACATGA AACCCCCCGC
TTCCTCCATT TTGGGGGCTA TTAATTTTAT TACCACAATT ATTAACATGA AACCTCCCGC
TTCGTCCATT TTGGGGGCTA TTAATTTTAT TACCACCATT ATTAACATGA AACCCCCCGC
TTCGTCCATT TTGGGGGCTA TTAATTTTAT TACCACCATT ATTAACATGA AACCCCCCGC
TTCGTCCATT TTGGGGGCTA TTAATTTTAT TACCACCATT ATTAACATGA AACCCCCCGC
TTCGTCCATT TTGGGGGCTA TTAATTTTAT TACCACCATT ATTAACATGA AACCCCCCGC
GTCGTCTATC CTAGGGGCTA TTAATTTTAT CACTACGATT ATCAACATAA AACCTCCCGC
ATCATCCATC TTAGGGGCCA TTAACTTTAT CACTACAATT ATCAACATGA AACCCCCTGC

TACCTCACAG TACCAAACCC CCCTATTCGT ATGATCAGTG CTAGTTACCG CAGTACTCCT
CACTTCACAA TATCAAACCC CTCTATTTGT ATGATCAGTA TTAGTTACCG CAATTCTTCT
AGTATCCCAA TATCAGACAC CTCTGTTTGT GTGATCTGTA TTAGTCACGG CCGTACTTCT
CATCTCCCAC TATCGGACNN NNNNATTTGT TTGATCCGAA TTAAGCACGG CCATACATCG
AGTCTCCCAA TACCAAACAC CCCTATTTGT GTGATCTGTA TTAGTCACGG CCGTGCTCCT
AGTCTCCCAA TACCAAACAC CCCTATTTGT GTGATCTGTA TTAGTCACGG CCGTGCTCCT
AGTCTCCCAA TACCAAACAC CCCTATTTGT GTGATCTGTA TTAGTCACGG CCGTGCTCCT
AGTATCCCAG TATCAGACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTTCTTCT
AGTATCCCAG TATCAGACGC CCCTATTCGT GTGATCTGTG TTAATCACGG CCGTCCTTCT
AGTATCCCAG TATCAGACGC CCCTATTCGT GTGATCTGTG TTAATCACGG CCGTCCTTCT
AGTATCCCAA TACCAGACAC CTCTATTTGT ATGATCTGTA TTAATCACGG CCGTGCTTCT
AGTATCCCAA TACCAGACAC CTCTATTTGT ATGATCTGTA TTAATCACGG CCGTGCTTCT
AGTATCCCAG TATCAGACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTTCT
AGTATCCCAG TATCAGACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTTCT
AGTCTCCCAA TATCAGACAC CTCTATTTGT ATGATCTGTA TTAATTACGG CCGTACTTCT
AGTCTCCCAA TATCAGACAC CTCTATTTGT ATGATCTGTA TTAATTACAG CCGTACTTCT
AGTCTCCCAA TATCAGACAC CTCTATTTGT ATGATCTGTA TTAATTACAG CCGTACTTCT
AGTATCCCAA TACCAGACAC CTTTATTTGT GTGATCTGTA TTAATCACGG CCGTACTTCT
AGTATCCCAA TATCAGACAC CTCTATTTGT GTGATCTGTG TTAATCACGG CCGTACTTCT
AGTATCCCAA TACCAGACAC CCCTGTTTGT GTGATCTGTA TTAATCACAG CCGTACTTCT
```

```
AGTATCCCAA TATCAGACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTTCT
AGTATCCCAA TATCAAACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTCCT
AGTATCCCAA TATCAAACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTCCT
AGTATCCCAA TATCAGACAC CTCTATTTGT GTGATCTGTA TTAATCACGG CCGTACTCCT
AGTATTTCAG TATCAAACCC CCCTATTCGT GTGATCTGTA CTAGTTACAG CTGTACTCCT
AGTATCCCAA TATCAAACCC CCCTATTCGT TTGGTCTGTG CTAATCACAG CTGTTCTTCT

GTTATTATCT CTACCGGTTT TAGCCGCAGG CATTACAATA CTCCTAACAG ACCGCAACCT
TCTATTATCC CTTCCTGTTT TAGCTGCAGG CATTACAATG CTCCTAACAG ACCGAAACCT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATG CTCCTAACAG ACCGAAATTT
ACTACTATCA CTGCCAGCGC TAGCTGAAGG AATCACCAAA CGCCGAACAG ACCGAAATTT
CCTACTGTCC CTGCCAGTGC TAGCTGCAGG AATCACAATA CTCCTGACAG ACCGAAATTT
CCTACTGTCC CTGCCAGTGC TAGCTGCAGG AATCACAATA CTCCTGACAG ACCGAAATTT
CCTACTGTCC CTGCCAGTGC TAGCTGCAGG AATCACAATA CTCCTGACAG ACCGAAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCAGG GATCACAATA CTCCTAACAG ACCGTAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATA CTCCTAACAG ACCGAAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATA CTCCTAACAG ACCGAAATTT
CCTGCTGTCA CTGCCAGTGC TAGCTGCGGG GATCACAATA CTTCTAACAG ATCGAAATTT
CCTGCTGTCA CTGCCAGTGC TGGCTGCGGG GATCACAATA CTTCTAACAG ATCGAAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCGGG AATCACAATA CTCCTAACAG ACCGGAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCGGG AATCACAATA CTCCTAACAG ACCGGAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCAGG AATTACAATA CTCCTAACAG ACCGGAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCGGG AATTACAATA CTCCTAACAG ACCGGAATTT
CCTACTGTCA CTGCCAGTGC TAGCTGCGGG AATTACAATA CTCCTAACAG ACCGGAATTT
CCTACTCTCA CTGCCAGTGC TAGCTGCGGG AATCACAATG CTCCTAACAG ACCGAAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATA CTCCTAACAG ACCGAAATTT
CCTATTATCG CTACCAGTGC TAGCTGCAGG GATTACAATA CTCCTAACAG ACCGAAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATG CTCCTAACAG ACCGAAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATG CTCCTAACAG ACCGAAATTT
CCTACTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATG CTCCTAACAG ACCGAAATTT
CCTGCTATCA CTGCCAGTGC TAGCTGCAGG GATCACAATA CTCCTAACAG ACCGAAATTT
CCTCCTATCA CTACCAGTTC TAGCTGCAGG AATTACAATA CTGCTGACAG ATCGAAATTT
CCTTCTCTCA CTACCAGTTC TAGCTGCCGG AATTACAATG CTACTAACAG ACCGAAATTT

TAATACTACC TTCTTTGACC CCGCGGGGGG TGGGGACCCA ATCCTCTATC AACATCTGTT
CAATACTACC TTCTTCGACC CAGCGGGTGG TGGAGACCCA ATCCTTTACC AACATTTATT
AAACACCACT TTCTTTGACC CAGNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AAACACCACC TTCTTTGACC CACCCGGAGA ACGAGACCCC AACCTCTACC AGCACCTATT
AAACACCACC TTCTTTGACC CAGCTGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCTGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCTGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CGGCCGGAGG AGGAGACCCC ATTCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG GGGAGACCCC ATTCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG GGGAGACCCC ATTCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CGGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTATC AGCACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCTTTGACC CAGCCGGAGG AGGAGACCCC ATCCTCTACC AACACCTATT
AAACACCACC TTCGTTGAAC CAGCCGGAGG GGGAGACCCC ATCCTCTACC AACACCTATT
AAATACCACA TTCTTTGACC CCGCCGGAGG GGGAGACCCC ATCCTGTATC AGCACCTGTT
AAACACCACA TTCTTTGACC CCGCCGGGGG AGGAGACCCC ATCCTCTATC AACACCTATT

CTGATTCTTC GGCCACCCAG AAGTCTACAT TCTTATCCTC CCAGGCTTCG GCATAATCTC
TTGATTCTTT GGCCACCCCG AAGTATATAT TCTGATTCTC CCGGGGTTCG GCATAATCTC
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATTTC
TTGATTCTTT GGCCACCCAG AGGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATTTC
```

TTGATTCTTT GGCCACCCAG AGGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATTTC
CTGATTTTTT GGTCACCCAG AGGTGTACAT CCTAATTCTA CCGGGATTCG GTATGATCTC
TTGATTCTTC GGTCACCCAG AGGTATACAT CCTGATTCTA CCGGGATTCG GCATGATCTC
TTGATTCTTC GGTCACCCAG AGGTATACAT CCTGATTCTA CCGGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTATACAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AAGTATACAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AAGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTGATTCTA CCGGGGTTCG GCATGATTTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTGATTTTA CCAGGGTTCG GCATGATTTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTGATTCTA CCAGGGTTCG GCATGATTTC
TTGATTCTTT GGCCACCCAG AAGTGTATAT TCTAATTCTA CCGGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTGTACAT TCTAATTCTA CCAGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTGTATAT TCTAATTCTA CCGGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTAATTCTA CCGGGATTCG GCATGATCTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTAATTCTA CCGGGATTCG GCATAATCTC
TTGATTCTTT GGCCACCCAG AGGTATATAT TCTAATTCTA CCGGGATTCG GCATAATCTC
TTGATTCTTT GGCCACCCAG AAGTATATAT TCTAATTCTG CCGGGATTCG GCATGATCTC
CTGATTCTTC GGCCACCCAG AAGTATATAT TCTAATTTTA CCTGGATTTG GCATGATTTC
CTGATTCTTT GGCCACCCAG AGGTATATAT TCTAATTTTA CCAGGGTTCG GCATGATCTC

CCACATCGTA GCCTATTATT CAGGTAAAAA TGAACCCTTC GGCTATATAG GAATAGTATG
TCATATCGTA GCCTATTATT CAGGTAAAAG TGAACCATTC GGATATATAG GAATAGTATG
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CCATATTGTA GCATACTACG CCGGCAAAAA GGAACCTTTT GGCTACATAG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA AGAACCCTTT GGCTACATGG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA AGAACCCTTT GGCTACATGG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA AGAACCCTTT GGCTACATGG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA AGAACCTTTT GGCTACATGG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATGG GAATAGTATG
CCACATTGTG GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATGG GAATAGTATG
CCACATTGTA GCCTACTATG CCGGCAAAAA AGAACCTTTT GGCTATATAG GAATAGTGTG
CCACATTGTA GCTTACTATG CCGGCAAAAA AGAACCTTTT GGCTATATAG GAATAGTGTG
CCACATTGTA GCATACTATG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCACATTGTA GCATACTATG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCACATTGTA GCATACTACG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCACATCGTA GCATACTACG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCACATCGTA GCATACTACG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCATATTGTA GCATACTATG CCGGCAAAAA AGAACCTTTT GGTTACATAG GAATAGTATG
CCATATTGTA GCATACTATG CCGGCAAAAA AGAACCTTTT GGCTACATAG GAATAGTATG
CCATATTGTA GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATAG GAATAGTATG
CCATATTGTG GCATACTATG CTGGCAAAAA GGAACCTTTT GGCTACATAG GGATAGTATG
CCATATTGTG GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATAG GAATAGTATG
CCATATTGTG GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATAG GAATAGTATG
CCATATTGTG GCATACTATG CCGGCAAAAA GGAACCTTTT GGCTACATAG GAATAGTATG
CCACATTGTA GCCTACTATG CAGGTAAAAA AGAACCCTTT GGATACATAG GCATGGTATG
CCACATCGTA GCCTACTATG CAGGTAAAAA GGAACCCTTT GGGTACATAG GCATGGTGTG

AGCAATAATA GCAATTGGGC TCCTAGGATT TATTGTGTGA GCTCATCACA TATTTACCGT
AGCAATAATA GCGATTGGAC TCTTAGGATT TATCGTCTGA GCCCACCACA TATTTACAGT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AGCCATGATG GCCATTGGGC TACTAGGCTT TATCGTAAGA GCTCATCACA TGTTTACAGT
AGCTATGATG GCCATCGGAC TACTAGGCTT TATCGTATGG GCCCATCACA TGTTTACAGT
AGCTATGATG GCCATCGGAC TACTAGGCTT TATCGTATGG GCCCATCACA TGTTTACAGT
AGCTATGATG GCCATCGGAC TACTAGGCTT TATCGTATGG GCCCATCACA TGTTTACAGT
AGCTATAATG GCTATCGGAC TACTAGGCTT TATCGTGTGG GCTCATCATA TATTCACAGT
AGCTATAATG GCTATCGGAC TACTAGGCTT TATCGTGTGG GCTCATCACA TGTTCACAGT
AGCTATAATG GCTATCGGAC TACTAGGCTT TATCGTGTGG GCTCATCACA TGTTCACAGT
GGCCATAATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTTACAGT
GGCCATAATG GCTATTGGGC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTTACAGT
GGCTATAATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTCACAGT
GGCTATAATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTCACAGT
GGCTATAATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTCACAGT
GGCTATGATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TGTTCACAGT
GGCTATGATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TGTTCACAGT
GGCTATAATG GCTATTGGAC TACTAGGCTT TATCGTGTGA GCTCATCACA TATTCACAGT
GGCTATGATG GCCATTGGAC TATTAGGCTT TATCGTATGA GCTCATCACA TATTTACAGT

```
GGCTATGATG GCCATTGGAC TACTAGGCTT TATCGTATGA GCTCATCACA TGTTTACAGT
AGCCATGATG GCCATTGGGC TACTAGGCTT TATCGTATGA GCTCATCACA TGTTTACAGT
AGCTATAATG GCCATTGGGC TACTAGGCTT TATCGTATGA GCTCATCACA TGTTTACAGT
AGCTATAATG GCCATTGGGC TACTAGGCTT TATCGTATGA GCTCATCACA TGTTTACAGT
AGCTATGATG GCTATTGGGC TACTAGGCTT TATCGTATGA GCTCATCACA TGTTTACAGT
GGCTATAATG GCTATCGGCC TGCTAGGATT TATCGTGTGA GCCCACCACA TGTTTACAGT
AGCTATGATG GCTATCGGTC TATTAGGTTT TATCGTGTGA GCCCACCACA TATTTACAGT

CGGCATAGAC GTAGATACTC GAGCCTACTT CACTTCCGCT ACAATAATTA TTGCGATCCC
TGGTATAGAT GTTGACACCC GAGCCTACTT CACATCCGCC ACTATAATTA TTGCTATTCC
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCTTACTT TACCTCCGCC ACTATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCTTACTT TACCTCCGCC ACTATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCTTACTT TACCTCCGCC ACTATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTATTT CACCTCCGCC ACAATAATCA TTGCCATCCC
TGGAATAGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATAGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TCGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TCGCTATCCC
TGGCATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATCA TCGCCATCCC
TGGGATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACGATAATCA TCGCCATCCC
TGGCATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATCA TCGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGGATGGAC GTAGACACAC GGGCCTATTT TACCTCCGCC ACGATAATCA TCCCAATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TCGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATGGAC GTAGACACAC GGGCCTACTT TACCTCCGCC ACAATAATTA TTGCCATCCC
TGGAATAGAC GTAGACACAC GGGCCTATTT CACCTCCGCC ACAATAATCA TTGCCATCCC
CGGAATGGAC GTTGACACAC GGGCCTACTT CACCTCCGCC ACAATAATTA TTGCCATCCC

AACGGGGGTC AAAGTATTTA GCTGACTTGC CACACTACAC GGAGGTGCAA TTAAATGGGA
AACTGGAGTT AAAGTATTTA GCTGACTTGC TACATTACAC GGAGGCGCAA TTAAATGAGA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CACAGGTGTC AAAGTCTTTA GCTGATCAGC CACCCTTCAC GGCGGCTCAA TTAAATGAGA
TACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAC GGCGGCTCAA TTAAATGGGA
TACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAC GGCGGCTCAA TTAAATGGGA
TACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAC GGCGGCTCAA TTAAATGGGA
CACAGGCGTT AAAGTCTTTA GCTGATTAGC TACCCTTCAC GGCGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGAGA
CACAGGCGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGAGA
CACAGGCGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGAGA
CACAGGCGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGGGA
CACAGGCGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGGGA
CACAGGCGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGTTCAA TTAAATGGGA
CACAGGTGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGCTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGCTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGGTTAGC CACCCTCCAT GGCGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTAGC CACCCTTCAC GGCGGCTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAT GGTGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAT GGTGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAT GGTGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAT GGTGGTTCAA TTAAATGAGA
CACAGGTGTC AAAGTCTTTA GCTGATTGGC CACCCTTCAT GGTGGTTCAA TTAAATGAGA
AACAGGGGTC AAAGTCTTTA GCTGATTAGC CACCCTCCAC GGCGGCTCAA TCAAATGAGA
AACGGGAGTT AAAGTTTTTA GCTGATTAGC CACCCTTCAC GGCGGCTCAA TCAAATGAGA

AACCCCTATA TTATGAGCTC TTGGCTTCAT TTTCTTATTT ACAGTCGGCG GACTAACCGG
AACTCCTATG CTATGAGCCC TGGGTTTCAT TTTCCTTTTC ACAGTTGGGG GACTAACTGG
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
TACCCCCCTA CTTTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACGGG
CACCCCACTA CTCTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GTTTAACGGG
```

```
CACCCCACTA CTCTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GTTTAACGGG
CACCCCACTA CTCTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GTTTAACGGG
TACCCCCCTG CTTTGGGCTC TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACAGG
TACCCCCCTG CTTTGAGCCC TAGGCTTTAT CTTCCTATTT ACAGTGGGAG GCTTAACAGG
TACCCCCCTG CTTTGAGCCC TAGGCTTTAT CTTCCTATTT ACAGTGGGAG GCTTAACAGG
CACCCCTCTA CTATGAGCCC TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
CACCCCCCTA CTATGAGCCC TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
CACCCCGCTA CTCTGAGCCC TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACGGG
CACCCCGCTA CTATGAGCCC TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACAGG
CACCCCGCTA CTCTGAGCCC TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACGGG
TACCCCCCTA CTTTGAGCCC TAGGCTTTAT TTTTCTATTC ACAGTGGGGG GCTTAACGGG
TACCCCCTTA CTTTGAGCCC TAGGCTTTAT TTTTCTATTC ACAGTGGGGG GCTTAACGGG
TACCCCTCTA CTCTGAGCCT TAGGCTTTAT TTTCCTATTC ACGGTGGGAG GCTTAACAGG
TACCCCCCTA CTTTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACGGG
TACCCCCCTA CTTTGAGCCT TAGGCTTTAT TTTCCTATTC ACAGTGGGAG GCTTAACAGG
TACCCCTCTA CTTTGAGCCT TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
TACCCCTCTA CTTTGAGCCT TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
TACCCCTCTA CTTTGAGCCT TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
TACCCCTCTA CTTTGAGCCT TAGGCTTTAT CTTCCTATTC ACAGTGGGAG GCTTAACGGG
CACTCCCCTA CTATGAGCCC TAGGCTTTAT CTTCCTATTT ACAGTGGGAG GCTTAACAGG
CACCCCCCTA CTTTGAGCCC TGGGCTTTAT TTTCTTATTT ACGGTAGGAG GCTTAACAGG

TATTATCTTA GCCAACTCAT CACTAGACAT CATACTACAC GATACATATT ACGTAGTTGC
CATTATTCTA GCAAACTCAT CATTAGATAT TATACTTCAC GATACGTATT ATGTAGTTGC
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AATTGTCCTG GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCAT CCCTGGATAT TGTGCTTCAT GACACCTATT ATGTCGTAGC
AATTGTCTTA GCCAACTCAT CCCTAGATAT CGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCAT CCCTAGATAT CGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTATT ATGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTATT ATGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT CGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT CGTACTTCAC GACACCTACT ACGTCGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTGCTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CCCTAGATAT TGTGCTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGACAT TGTACTTCAT GACACCTACT ACGTTGTAGC
AATTGTCCTG GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCGT CTCTAGATAT TGTACTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCAT CCCTAGACAT CGTGCTTCAC GACACCTACT ACGTTGTAGC
AATTGTCTTA GCCAACTCAT CCCTGGACAT TGTACTTCAC GACACCTACT ACGTCGTAGC

ACACTTCCAC TATGTCCTGT CAATAGGAGC CGTCTTTGCC ATTATGGGCG GCTTTGTCCA
ACACTTCCAC TATGTTCTCT CCATAGGAGC CGTATTTGCT ATTATAGGAG GCTTCGTCCA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
ACACTTCCAC TAGGCATTAT CAATCGGCGC TTTGGTCCGC ATCAAAGGAG CCCTCGAACA
ACATTTCCAT TACGTGCTGT CAATGGGGGC TGTGTTTGCC ATTATAGGAG CTTTCGTGCA
ACATTTCCAT TACGTGCTGT CAATGGGGGC TGTGTTTGCC ATTATAGGAG CTTTCGTGCA
ACATTTCCAT TACGTGCTGT CAATGGGGGC TGTGTTTGCC ATTATAGGAG CTTTCGTGCA
ACACTTCCAC TATGTGCTAT CAATGGGGGC TGTGTTCGCT ATCATAGGGG CCTTCGTACA
ACACTTCCAC TATGTACTAT CAATGGGAGC CGTGTTCGCT ATTATAGGGG CCTTCGTACA
ACACTTCCAC TATGTACTAT CAATGGGAGC CGTGTTCGCT ATTATAGGGG CCTTCGTACA
ACATTTTCAT TATGTACTAT CAATGGGGGC TGTATTCGCC ATTATGGGGG CCTTCGTCCA
ACATTTTCAT TATGTACTAT CAATGGGGGC TGTATTCGCC ATTATGGGGG CCTTCGTCCA
ACATTTTCAT TATGTGTTAT CGATAGGAGC TGTATTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAT TATGTATTAT CAATAGGAGC TGTATTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAT TATGTGTTAT CGATAGGAGC TGTATTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAT TATGTGTTAT CAATAGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAT TATGTGTTAT CAATAGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAC TATGTACTAT CAATGGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
```

```
ACATTTTCAC TATGTATTAT CAATGGGGGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACATTTTCAC TATGTACTAT CAATGGGGGC TGTGTTCGCC ATTATAGGAG CCTTCGTACA
ACATTTCCAC TATGTATTAT CAATGGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACATTTCCAC TATGTATTAT CAATGGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACATTTCCAC TATGTATTAT CAATGGGAGC TGTGTTCGCC ATCATAGGGG CCTTCGTACA
ACATTTCCAC TATGTATTAT CAATGGGAGC TGTGTTCGCC ATTATAGGGG CCTTCGTACA
ACACTTTCAC TACGTACTAT CAATGGGGGC CGTATTTGCT ATTATAGGGG CTTTCGTTCA
TCACTTTCAC TATGTACTGT CAATAGGAGC CGTTTTCGCC ATCATGGGAG CTTTCGTCCA

TTGATTCCCC CTGTTCTCAG GCTTTACACT TCACCCTACC TGGACTAAAA TTCACTTCGG
CTGATTCCCA TTATTCTCAG GATATACACT TCATCCGACA TGAACCAAAA TCCACTTCGG
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CTGAGTCCCG CGTCACACAA GTTATACACT ACACGGCACC TGACCCAAAA TCCAATTTGC
CTGATTCCCA CTCTTCACGG GTTACACTCT ACACAGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCA CTCTTCACGG GTTACACTCT ACACAGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCA CTCTTCACGG GTTACACTCT ACACAGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACGG GTTATACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTTACGG GTTATACGCT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTTACGG GTTATACGCT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTCTTCACAG GTTATACACT ACATGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTCTTCACAG GTTATACACT ACATGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCA CTTTTCACAG GTTACACACT ACATGGCACC TGATCCAAAA TCCACTTCGC
CTGATTCCCA CTTTTCACAG GTTACACACT ACATGGCACC TGATCCAAAA TCCACTTCGC
CTGATTCCCA CTTTTCACAG GTTACACACT ACATGGCACC TGATCCAAAA TCCACTTCGC
CTGATTCCCG CTTTTTACAG GTTTTACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTTACAG GTTTTACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACAG GTTATACGCT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCA CTTTTCACAG GTTATACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
TTGATTCCCG CTTTTTACGG GTTATACACT ACACGGCACC TGGTCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACAG GTTATACACT ACACGGCACC TGGTCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACGG GTTATACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACGG GTTATACACT ACACGGCACC TGATCCAAAA TCCACTTTGC
CTGATTCCCG CTTTTCACGG GTTATACACT ACACGGCACC TGGTCCAAAA TCCACTTTGC
CTGATTCCCA CTTTTCACAG GTTACACACT TCACAGCACC TGATCTAAAA TCCACTTTGC
TTGATTCCCG CTTTTCACAG GTTATACGCT TCACAGCACC TGATCTAAAA TCCACTTCGC

AATTATATTT ATTGGGGTAA ACATAACATT CTTCCCTCAG CACTTCCTAG GACTAGCTGG
AATTATATTT ATTGGAGTCA ACCTAACATT CTTTCCTCAA CATTTCCTCG GATTGGCTGG
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CGTAATATTT GTAGGCGTCA ATCTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTCATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTGG GCCTCGCAGG
CGTCATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTGG GCCTCGCAGG
CGTCATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTGG GCCTCGCAGG
CGTAATATTT GTAGGCGTAA ACTTAACATT CTTCCCCCAA CACTTCCTAG GTCTCGCAGG
CGTGATATTT GTAGGCGTAA ACTTGACATT CTTCCCCCAA CACTTCCTAG GTCTCGCAGG
CGTGATATTT GTAGGCGTAA ACTTGACATT CTTCCCCCAA CACTTCCTAG GTCTCGCAGG
CGTAATATTT GTAGGCGTCA ACTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGCGTCA ACTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGCGTCA ACTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGCGTCA ACTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGCGTCA ACTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTC GTAGGCGTCA ACTTAACATT TTTTCCCCAA CATTTCCTAG GCCTCGCAGG
CGTAATATTC GTAGGCGTCA ACTTAACATT TTTTCCCCAA CATTTCCTAG GCCTCGCAGG
TGTAATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
TGTAATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
TGTAATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAG CACTTCCTAG GCCTTGCAGG
TGTAATATTT GTAGGTGTTA ATTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGTGTCA ATCTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
CGTAATATTT GTAGGTGTCA ATCTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
TGTAATATTT GTAGGTGTCA ATTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
TGTAATGTTC GTAGGCGTCA ATTTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG
TGTGATGTTT GTAGGCGTCA ACCTAACATT CTTCCCCCAA CACTTCCTAG GCCTCGCAGG

AATACCACGA CGTTACTCAG ATTACCCAGA TGCATACACC ATATGAAACT -CTTTATCTT
AATACCACGA CGCTACTCAG ACTACCCTGA CGCATATACT TTATGAAACT -CTTTATCCT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN -NNNNNNNNN
AATACCCCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA GCCGCCTCCT
```

```
AATGCCCCGC CGATACTCAG ACTACCCTGA CGCATACGCT CTATGAAACA -CCATCTCCT
AATGCCCCGC CGATACTCAG ACTACCCTGA CGCATACGCT CTATGAAACA -CCATCTCCT
AATGCCCCGC CGATACTCAG ACTACCCTGA CGCATACGCT CTATGAAACA -CCATCTCCT
TATACCCCGC CGATACTCAG ACTACCCAGA CGCATATGCC CTATGAAACA -CCGTTTCCT
TATACCCCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA -CCGTTTCCT
TATACCCCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA -CCGTTTCCT
AATGCCCCGC CGATACTCAG ACTACCCAGA CGCATATGCC CTATGAAACA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCAGA CGCATATGCC CTATGAAACA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTGTGAAACA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTGTGAAACA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCGGA CGCATACGCC CTGTGAAATA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCGGA CGCATACGCC CTGTGAAATA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTACCCGGA CGCATACGCC CTGTGAAATA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCGTACGCC CTATGAAACA -CCGTCTCCT
AATGCCCCGC CGATACTCAG ACTATCCGGA CGCATACGCC CTGTGAAATA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTGTGAAACA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCATATGCC CTATGAAACA -CCGTCTCCT
AATGCCTCGC CGATACTCAG ACTACCCAGA CGCATACGCC CTATGAAACA -CCGTCTCCT
CATACCCCGT CGATACTCAG ACTACCCCGA CGCCTACGCC CTATGAAACA -CTGTCTCCT
TATACCCCGC CGATATTCAG ACTACCCCGA TGCCTATGCC CTATGAAACA -CTGTCTCCT

CTATTGGCTC TATAATCTCC CTTACTGCTG TTATTATATT TTTATTTATC CTCTGAGAAT
CCATCGGATC CATAATTTCC CTTACAGCTG TTATTATATT CCTATTTATT CTTTGAGAAT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CAATCAGATC ACTAATCTCA TTAATTGCCG TAATTATGTT CCTATTTATT TTATGAGAAT
CAATTGGTTC ACTAATTTCA CTAGTTGCTG TGATTATATT CCTATTCATT CTGTGAGAAT
CAATTGGTTC ACTAATTTCA CTAGTTGCTG TGATTATATT CCTATTCATT CTGTGAGAAT
CAATTGGTTC ACTAATTTCA CTAGTTGCTG TGATTATATT CCTATTCATT CTGTGAGAAT
CAATCGGCTC ACTAATCTCA TTAATTGCCG TAATTATGTT CCTATTTATT TTATGAGAAT
CAATCGGCTC ACTAATTTCA TTAATCGCCG TAATTATGTT CCTATTTATT TTATGAGAAT
CAATCGGCTC ACTAATTTCA TTAATCGCCG TAATTATGTT CCTATTTATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTGGTTGCTG TGATTATATT CCTATTTATC TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTGATTGCTG TGATTATATT CCTATTTATC TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTCATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCCG TGATTATATT CCTATTCATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TAATTATATT CCTATTCATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTGTTCATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTGTTCATT TTATGAGAAT
CAATCGGCTC ACTAATCTCA CTGATTGCTG TGATTATATT CCTATTCATT CTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCCG TGATTATATT CCTATTCATC CTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTCATT CTATGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTTATT CTGTGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTTATT CTGTGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTTATT CTGTGAGAAT
CAATCGGCTC ACTAATCTCA TTAGTTGCTG TGATTATATT CCTATTTATT CTGTGAGAAT
CAGTTGGCTC ATTAATCTCG CTTATCGCTG TAATTATGTT CTTATTCATT CTTTGAGAAT
CCATTGGCTC ACTAATCTCA CTTATTGCTG TAATTATGTT CCTATTCATC CTCTGAGAAT

GGCCCACCCC GCACAACTAG GACTACAAGA CGCATCCTCC CCTATTATAG AAGAACTCCT
GGCCCACCCA GCACAACTAG GACTTCAAGA CGCATCCTCC CCTATTATAG AAGAACTTCT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CATGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
```

```
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCTCATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTAATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTTATAG AAGAACTTCT
GGCACATCCA TCACAATTAG GATTCCAAGA CGCGGCCTCA CCTGTTATAG AAGAACTTCT

ACACTTCCAT GACCATGCAT TAATAATCGT CTTTTTAATT AGTACCCTTG TACTTTACAT
CCATTTTCAC GACCACGCTC TAATAATTGT TTTCCTAATT AGCACACTAG TACTTTATAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CGTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTCTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTCTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTCTACAT
CCACTTCCAT GATCACACAC TGATGATTGT TTTCCTAATC AGCACTTTAG TGCTCTACAT
CCACTTCCAT GATCACACAC TAATGATTGT CTTCCTAATC AGCACTTTAG TGCTCTACAT
CCACTTCCAT GATCACACAC TAATGATTGT CTTCCTAATC AGCACTTTAG TGCTCTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTGG TGCTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTGG TGCTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTTTACAT
CCACTTCCAT GATCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTTTACAT
CCACTTCCAT GATCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTTTACAT
CCACTTCCAT GATCACACAC TAATGATTGT CTTCCTAATT AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTCCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTCTAG TACTTTACAT
CCACTTTCAT GACCACACAC TAATGATTGT CTTCCTAATC AGTACTTTAG TACTTTACAT
CCACTTTCAT GACCACACAC TAATGATTGT CTTCCTAATC AGCACTTTAG TACTTTACAT

CATCACAACA ACAGTCTCAA CAAAACTAAC CAACAAGCAC TTACTTGACG CCCAAGAAAT
TATTACAACA ACTGTCTCAA CAAAACTAAC TAATAAACAC CTACTTGATG CTCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTAGCC ATGGTATCAA CTAAGCTAAC AAACAAATAC GTACTGGACT CCCAAGAAAT
TATTGTAGCC ATGGTATCAA CTAAGCTAAC AAACAAATAC GTACTGGACT CCCAAGAAAT
TATTGTAGCC ATGGTATCAA CTAAGCTAAC AAACAAATAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAGTAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAGTAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAGTAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAC GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAGAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAGAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGACT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGATT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGATT CCCAAGAAAT
TATTGTGGCC ATGGTGTCAA CTAAACTAAC AAACAAATAT GTACTGGATT CCCAAGAAAT
TATTGTGGCC ATAGTGTCAA CTAAACTAAC AAACAAATAC GTACTAGACT CCCAAGAAAT
TATTGTGGCC ATAGTGTCAA CTAAACTAAC AAACAAATAC GTACTAGACT CCCAAGAGAT

TGAAATAGTC TGAACAGTTA TACCAGCTTT AGTGTTAATT GCCATTGCCT TGCCGTCACT
TGAAATGGTA TGAACAGTTA TACCGGCTCT AGTATTGATT ACTATTGCCC TACCATCACT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TACCTTCCCT
```

TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGAACAGTGC TCCCAGCAGT AATTCTAATC TTGATTGCCC TGCCCTCCCT
TGAAATTGTA TGAACAGTGC TCCCAGCAGT AATTCTAATC TTGATTGCCC TGCCCTCCCT
TGAAATTGTA TGAACAGTGC TCCCAGCAGT AATTCTAATC TTGATTGCCC TGCCCTCCCT
TGAAATTGTA TGGACAGTGC TCCCAGCAGT AATTCTAATC TTAATTGCCC TGCCCTCCCT
TGAAATTGTA TGGACAGTGC TCCCAGCAGT GATCCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGGACAGTGC TCCCAGCAGT GATCCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGGACAGTGC TCCCGGCAGT AATCCTAATC TTAATTGCCC TACCCTCTCT
TGAAATTGTA TGGACAGTGC TCCCGGCAGT AATCCTAATC TTAATTGCCC TACCCTCTCT
TGAAATTGTA TGGACAGTAC TCCCAGCAGT AATCCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGGACAGTAC TCCCAGCAGT AATCCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATCCTAATC TTAATTGCCC TACCCTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATCCTAATT TTAATTGCCC TGCCCTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATCCTAATT TTAATTGCCC TGCCCTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATCTTAATC TTAATTGCCC TGCCTTCTCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTCTAATC TTAATCGCCC TACCTTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TGCCCTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TACCTTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TACCTTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TACCTTCCCT
TGAAATTGTA TGAACAGTAC TCCCAGCAGT AATTTTAATC TTAATTGCCC TACCTTCCCT
CGAAATTGTG TGAACAGTAC TCCCAGCAGT AATCCTAATT TTAATCGCTC TACCCTCCCT
CGAAATTGTG TGAACAGTAC TCCCAGCAGT AATTCTGTTT TTAATCGCCC TACCCTCCCT

CCGAATCCTT TACCTAATAG ACGAAATCAA TGACCCTCAC CTAACTATCA AAGCCACAGG
TCGAATCCTT TATCTTATAG ATGAAATCAA TGATCCCCAC CTAACCATTA AAGCCACAGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTAACAATTA AAGCTATAGG
TCGAACTCTT TACCTAATAG ACGAGATCAA CGACCCCCAC CTAACGATTA AAGCTATAGG
TCGAATTCTC TACCTAATAG ACGAGATCAA CGACCCCCAC CTAACGATCA AGGCTATGGG
TCGAATTCTC TACCTAATAG ACGAGATCAA CGACCCCCAC CTAACGATCA AGGCTATGGG
TCGAATTCTC TACCTAATAG ACGAGATCAA CGACCCCCAC CTAACGATCA AGGCTATGGG
TCGAATTCTT TACCTAATAG ACGAGATCAA CGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TACCTGATAG ACGAGATCAA CGACCCCCAC CTAACGATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATCAA CGACCCCCAC CTAACGATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATCAA TGACCCACAC CTGACTATTA AAGCCATGGG
TCGAATTCTT TACCTAATAG ACGAGATCAA TGACCCACAC CTGACTATTA AAGCCATGGG
TCGAATTCTT TACCTAATAG ACGAAATCAA TGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATCAA TGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATCAA TGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TATCTAATAG ACGAGATCAA TGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TATCTAATAG ACGAGATCAA TGACCCCCAC CTGACGATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ATGAGATTAA TGACCCCCAC CTGACAATCA AAGCTATGGG
TCGAATTCTT TACCTAATAG ACGAGATCAA CGACCCCCAC CTGACAATTA AAGCTATGGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTGACAATTA AAGCTATGGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTGACAATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTGACAATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTGACAATTA AAGCTATAGG
TCGAATTCTT TACCTAATAG ACGAGATTAA TGACCCCCAC CTGACAATTA AAGCTATAGG
CCGAATTCTC TACCTAATAG ACGAAATTAA CGACCCTCAC CTAACAATCA AAGCTATAGG
CCGAATTCTT TACCTAATAG ACGAAATTAA CGATCCCCAC TTAACAATCA AAGCCATAGG

ACATCAGTGG TACTGAAGCT ATGAGTACAC AGACTACGAG ACATTAAATT TTGACTCATA
ACATCAATGA TACTGAAGTT ATGAATATAC AGATTACGAA ACATTAAACT TTGACTCATA
ACACCAATGA TACTGAAGTT ATGAATACAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACATCAATGA TACTGAAGTT ATGAATATAC AGATTATGAA GACCTGGGAT TTGACTCCTA
ACACCAATGA TACTGAAGCT ATGAGTACAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGCT ATGAGTACAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGCT ATGAGTACAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACATCAATGA TACTGAAGTT ATGAATATAC GGATTATGAA GATCTGGGAT TTGACTCCTA
ACATCAATGA TACTGAAGTT ATGAATATAC AGATTATGAA GACCTGGGAT TTGACTCCTA
ACATCAATGA TACTGAAGTT ATGAATATAC AGATTATGAA GACCTGGGAT TTGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA

```
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
GCACCAATGA TACTGAAGTT ATGAGTATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAATACAC AGACTATGAA GACCTAGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAGTATAC GGACTACGAA GACCTGGGCT TCGACTCCTA
ACACCAGTGA TACTGAAGTT ATGAATATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAATATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAATATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGTT ATGAATATAC GGACTATGAA GACCTGGGCT TCGACTCCTA
ACACCAATGA TACTGAAGCT ATGAATACAC AGACTACGAA AACTTAGGCT TTGACTCCTA
ACACCAATGA TACTGAAGCT ACGAATACAC AGACTACGAA AGCCTGGGCT TTGACTCCTA

TATAATTCCT ACACAAGATC TTCTACCAGG CCAATTCCGC CTACTTGACA CAGACAACCG
CATAATTCCA ACACAAGACC TTCTACCTGG GCAATTCCGC CTCCTTGACA CAGACCACCG
CATAATCCCC ACACAAGACC TCACCCCAGG GCAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATTCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCATCCCTGG ACAATTCCGT CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCATCCCTGG ACAATTCCGT CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCATCCCTGG ACAATTCCGT CTCCTAGAAG CAGACCATCG
CATGATTCCC ACACAAGACC TCTCCCCAGG ACAATTCCGA CTCCTAGAAG CAGACCATCG
CATGATTCCC ACACAAGACC TTTCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATGATTCCC ACACAAGACC TTTCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCACCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCACCG
CATAATCCCC ACACAAGACC TCGCCCCGGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATAATTCCC ACACAAGACC TCGCCCCGGG ACAATTCCGA CTCCTAGAAA CAGACCATCG
CATGATCCCC ACACAAGACC TCGCCCCAGG GCAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG ACAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCGCCCCAGG GCAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCT ACACAAGACC TCGCCCCAGG GCAATTCCGA CTCCTAGAAG CAGACCATCG
CATAATCCCC ACACAAGACC TCGCCCCAGG GCAATTCCGA CTCCTAGAAG CAGACCATCG
TATAATTCCC ACACAAGACC TCACTCCCGG ACAATTCCGA CTACTAGAGA CAGACCACCG
CATAGTCCCC ACACAAGACC TCACCCCAGG CCAATTCCGA CTTCTAGAAA CAGACCATCG

AATAGTTGTC CCCACCAGCT CTCCCGTACG TATATTAATT ACAGCTGAAG ACGTCCTACA
AATAGTTGTC CCAACAGGAT CTCCAGTGCG AATACTTATC ACCGCTGAAG ACGTCCTTCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ACGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ACGTACTCCA
AATAGTGGTG CCCATAGAAT CACCAATTCG AGTTCTAGTA TCCGCAGAAG ATGTACTCCA
AATAGTGGTG CCCATAGAAT CACCAATTCG AGTTCTAGTA TCCGCAGAAG ATGTACTCCA
AATAGTGGTG CCCATAGAAT CACCAATTCG AGTTCTAGTA TCCGCAGAAG ATGTACTCCA
AATAGTAGTA CCCATAGAAT CCCCAATCCG AGTCCTAGTT TCCGCAGAAG ATGTACTCCA
AATAGTAGTA CCCATAGAAT CCCCAATCCG AGTCCTAGTT TCCGCAGAAG ATGTACTCCA
AATAGTAGTA CCCATAGAAT CCCCAATCCG AGTCCTAGTT TCCGCAGAAG ATGTACTCCA
AATAGTAGTC CCAATGGAGT CTCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTACTCCA
AATAGTAGTT CCAATGGAGT CTCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTACTCCA
AATAGTAGTT CCCATAGAAT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTACTCCA
AATGGTAGTT CCCATAGAAT CTCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTACTCCA
AATAGTAGTA CCAATAGAAT CTCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTACTCCA
AATAGTAGTA CCCATAGAGT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTGCTCCA
AATAGTAGTA CCCATAGAGT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ATGTGCTCCA
AATGGTGGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ACGTACTCCA
AATGGTAGTA CCCATGGAAT CCCCAATTCG AGTTCTAGTT TCCGCAGAAG ATGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTC TCCGCAGAAG ACGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTT TCCGCAGAAG ACGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTT TCCGCAGAAG ACGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTT TCCGCAGAAG ACGTACTCCA
AATGGTAGTG CCCATAGAAT CCCCAATTCG AGTTCTAGTT TCCGCAGAAG ACGTACTCCA
AATAGTAGTA CCCATGGAGT CCCCAATTCG AGTTTTAGTC TCTGCAGAAG ATGTGCTCCA
CATGGTAGTA CCCATGGAGT CCCCAATTCG AGTCTTAGTT TCCGCAGAAG ATGTACTCCA

TTCATGAGCT GTGCCGTCTC TAGGCCTAAA AATGGATGCA GTACCCGGAC GACTCAATCA
CTCATGAGCA GTACCATCAT TAGGCCTAAA AATAGATGCC GTACCAGGAC GACTTAATCA
```

CTCCTGAGCA GTGCCAGCCC TAGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TAGGCATCAA AATAGACGCA GTACCCGGAC GCCTAAATCA
CTCCTGAGCA GTGCCGGCCT TAGGCATCAA AATAGACGCA GTGCCAGGGC GCCTGAACCA
CTCCTGAGCA GTGCCGGCCT TAGGCATCAA AATAGACGCA GTGCCAGGGC GCCTGAACCA
CTCCTGAGCA GTGCCGGCCT TAGGCATCAA AATAGACGCA GTGCCAGGGC GCCTGAACCA
CTCCTGAGCG GTACCAGCCC TAGGCATCAA GATAGACGCA GTACCAGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TAGGCATCAA GATAGACGCA GTACCAGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TAGGCATCAA GATAGACGCA GTACCAGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TGGGCATTAA AATAGACGCG GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TGGGCATTAA AATAGACGCG GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TAGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TAGGCATCAA AATAGATGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TGGGCATCAA AATAGACGCA GTGCCTGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TAGGCATCAA AATAGACGCG GTGCCTGGAC GCCTAAATCA
CTCCTGAGCG GTACCAGCCC TAGGCATCAA AATAGACGCG GTGCCTGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TAGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TAGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCA GTGCCAGCCC TGGGCATCAA AATAGACGCA GTACCCGGAC GCCTAAACCA
CTCCTGAGCG GTGCCAGCCC TGGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TGGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TGGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CTCCTGAGCG GTGCCAGCCC TGGGCATCAA AATAGACGCA GTGCCCGGAC GCCTAAATCA
CGCCTGAGCA GTACCCGCCC TAGGCATTAA AATAGACGCA GTACCAGGAC GCCTAAACCA
CGCTTGAGCA GTGCCAGCCC TAGGCATCAA AATAGACGCA GTACCAGGAC GCCTAAATCA

AGCGACATTT ATTGCCACCC GACCAGGAAT CTTCTTCGGT CAGTGTTCAG AAATCTGCGG
AGCCACATTT ATTGCCACTC GACCAGGAAT CTTCTTCGGC CAATGCTCAG AAATTTGTGG
AACAGCCTTT ATCACCTCCC GACCGGGAGT TTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCAGGGGA TTACTACGGC CAATGCTCCG AAATTTACGG
AACAGCTTTT ATTACCTCCC GACCAGGAGT CTATTACGGC CAATGCTCTG AAATCTGCGG
AACAGCTTTT ATTACCTCCC GACCAGGAGT CTATTACGGC CAATGCTCTG AAATCTGCGG
AACAGCTTTT ATTACCTCCC GACCAGGAGT CTATTACGGC CAATGCTCTG AAATCTGCGG
GACGGCCTTT ATTACCTCAC GACCAGGAGT TTATTACGGC CAATGCTCTG AAATTTGCGG
AACGGCCTTT ATCACCTCAC GACCAGGAGT TTATTATGGC CAATGCTCTG AAATTTGCGG
AACGGCCTTT ATCACCTCAC GACCAGGAGT TTATTATGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATCACCTCAC GACCAGGGGT TTACTACGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATCACCTCAC GACCAGGGGT TTACTACGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATCACCTCAC GACCAGGGGT TTACTATGGC CAATGCTCCG AAATTTGCGG
AACAGCCTTT ATTACCTCAC GACCAGGAGT TTACTATGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATTACCTCAC GACCAGGAGT TTACTACGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGGGT TTACTACGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGGGT TTACTACGGC CAATGCTCTG AAATTTGCGG
AACAGCCTTT ATCACCTCAC GACCAGGGGT CTACTACGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCAGGGGT TTATTATGGC CAGTGCTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGGGT TTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGAGT TTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGAGT TTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGAGT CTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATTACCTCAC GACCGGGGGT TTATTATGGC CAATGTTCTG AAATCTGCGG
AACAGCCTTT ATCACCTCAC GACCAGGAGT TTATTACGGT CAATGCTCTG AAATCTGCGG
AACAGCCTTT ATCACCTCAC GACCGGGAAT CTATTATGGT CAATGCTCTG AAATCTGCGG

AGCGAACCAC AGCTTCATGC CTATTACAAT TGAATCGACG CCTGTTAAAT ACTTTGAATC
AGCAAACCAC AGCTTTATAC CTATTGCGAT TGAATCTTCG CCAGTAAAAT ACTTTGAATC
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
GGCTAAAACC AGCTTCATGC CAATTGTAGT CGAAACAGTG GCTCTGGAAG ATTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCGGTC CCCTTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCGGTC CCCTTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCGGTC CCCTTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAGC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCTCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCTCTAGAAC ACTTTGAAAA
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCTCTAGAAC ACTTTGAAAA

```
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCTCTAGAAC ACTTTGAAAA
GGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATAC CAATTGTAGT CGAAGCCGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTTATGC CAATTGTAGT CGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT TGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT TGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT TGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCTAACCAC AGCTTCATGC CAATTGTAGT TGAAGCAGTC CCCCTAGAAC ACTTTGAAAA
AGCCAATCAT AGCTTCATAC CCATCGTCGT TGAAGCAGTC CCCCTAGAGC ACTTCGAAAA
GGCCAATCAC AGCTTCATAC CAATTGTTGT TGAAGCAGTT CCCCTAGAAC ACTTTGAAAA

ATGATCTTCA TCAATACTAG CAGAATCAT- ATGCCACAAC TCAATCCAGA CCCCTGATTC
CTGATCTTCA TCAATACTAG CAGAATCTT- ATGCCACAAC TTAACCCAGA CCCCTGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TTAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAGTCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAGTCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAGTCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCTCAAC TGAATCCCAA CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TAAACCCCAA CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TAAACCCCAA CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAATCCTC ATGCCCCAAC TAAACCCCAA CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAATCCTC ATGCCCCAAC TAAACCCCAA CCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAA CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAA CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAA CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TGAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TTAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TTAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TTAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCCTC ATGCCCCAAC TTAACCCCAG CCCATGATTT
CTGATCTTCA TTAATACTAG AAGAATCTTC ATGCCCCAAC TAAATCCCAG TCCATGATTT
CTGATCTTCA TTAATGCTAG AAGAGTCTTC ATGCCCCAAC TAAACCCCAG CCCATGATTT

TATATCTTAA TCTTCACATG AGCTGTTTTC CTGACCATCC TGCCAAATAA GGTTATTATA
ACCATCTTAA TTTTTACATG AGCCGTATTC TTAACTATTC TACCAAATAA AGTCACCTCA
ATAATCTTAG TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAA TTTTTTCATG ACTTATTTTC CTGATTGTTT TACCACCCAA AGTGCTAGGC
ATAATTTTAA TTTTTTCATG ACTTATTTTC CTGATTGTTT TACCACCCAA AGTGCTAGGC
ATAATTTTAA TTTTTTCATG ACTTATTTTC CTGATTGTTT TACCACCCAA AGTGCTAGGC
ATAATCTTAG TCTTCTCATG ACTTATTTTC TTAATCGTTC TACCACCTAA AGTGTTAGGC
ATAATCTTAG TCTTCTCATG ACTTATTTTC CTAATCGTTC TACCACCTAA AGTGTTAGGC
ATAATCTTAG TCTTCTCATG ACTTATTTTC CTAATCGTTC TACCACCTAA AGTGTTAGGC
ATAATTTTAA TTTTCTCATG ACTTATCTTC CTAATTATTC TACCGCCTAA AGTGCTAGGC
ATAATTTTAA TTTTCTCATG ACTTATCTTC CTAATTATTC TACCGCCTAA AGTGCTAGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTAATTATTT TACCGCCTAA AGTGCTAGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTGATTATTC TACCGCCTAA AGTGCTAGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTAATTGTTT TACCACCTAA AGTGCTGGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTAATTGTTT TACCGCCTAA AGTGCTGGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTAATTGTTT TACCGCCTAA AGTGCTGGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAA TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC CTAATTGTCT TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC TTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC TTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTCTCATG ACTTATTTTC CTAATTGTTC TACCACCCAA AGTGCTAGGC
ATAATTTTAG TTTTTTCATG ACTTGTCTTC CTAATTATTC TACCACCCAA AGTTCTAAGC
ATAATCTTAA TTTTCTCATG ACTTATCTTC CTAATTATTC TACCACCCAA AGTTCTAGGC

CACAAAACCC CAAATGAACC CCTCACTAAA GACCCTTCGA GTCTTCTCAC AGAAATCTGA
```

```
CATAAAACCC CAAACGAACC AACTTCTAAA GACTCATCTA GTTTAATAAC AGAAATCTGA
CATACCTTCA CGAACGAACC CACCCATAAA AATGCGGAAA AGACTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAGCCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCGGAAA AGATTAAACC TGAGCCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAGCCTTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAGCCTTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AAATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AAATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAGCCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCACAAA AATGCAGAAA AGATTAAACC TGAGCCCTGA
CATACCTTCA CGAACGAACC CACCCATAAA AACGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCATAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCATAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CATACCTTCA CGAACGAACC CACCCATAAA AATGCAGAAA AGATTAAACC TGAACCCTGA
CACACCTTCA CAAATGAACC TACCCATAAA AATGCAGAAA AAATTAAACC TGAACCCTGA
CACACCTTCA CAAATGAACC TACCCACAAA AATGCAGAGA AGATTAAACC CGAACCCTGA

CCCTGACCAT GACACTAAAG CTTCTTTGAT CAATTTGCCA GCCAAACCTT CTTAGGTGCC
TCCTGACCAT GACATTAAAG CTTTTTTGAC CAATTTGCCA GCCAATCCTT TTTAGGTATT
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACGCA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTCTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATT
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAGTTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAGTTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAGTTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATACTAAAG CTTTTTCGAT CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATACTAAAG CTTTTTTGAC CAATTCATGA GCCCCACATA CCTGGGAATT
ACCTGACCAT GATACTAAAG CTTTTTTGAC CAATTCATGA GCCCCACATA CCTGGGAATT
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTAGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTAGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTAGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTAGGAATT
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTAGGAATT
ACCTGACCAT GATCCTAAAG CTTCTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATT
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC
ACCTGACCAT GATCCTAAAG CTTTTTTGAC CAATTCATGA GCCCCACACA CCTGGGAATC

CCACTAATTG CCGTTGCTCT TCTACTACCA TGAATATTAT TCCCAGCCCC GTATAAACGA
CCATTAATCG CCATTGCCCT CCTACTCCCA TGAATACTAT TTCCATCACC TTATAAACGA
CCCTTAATTG CTCTAGCACT CAGCCTTCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCCTTAATTG CTCTAGCACT CAGCCTTCCA TGAGTACTTA TTCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTCCCC TGAGTACTCA TCCCCACCCC GACCAACCGG
CCCTTAATTG CCCTAGCACT CAGCCTCCCC TGAGTACTCA TCCCCACCCC GACCAACCGG
CCCTTAATTG CCCTAGCACT CAGCCTCCCC TGAGTACTCA TCCCCACCCC GACCAACCGG
CCCCTAATTG CCCTAGCACT CAGCCTTCCA TGAGTATTAA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTGGCACT CAGCCTTCCA TGAGTACTAA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTGGCACT CAGCCTTCCA TGAGTACTAA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTAGCACT TAGCCTCCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTAGCACT TAGCCTCCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTGGCACT TAGCCTTCCA TGGGTACTTA TCCCCACCCC CACCAACCGA
CCCCTAATTG CCCTGGCACT TAGCCTTCCA TGGGTACTTA TCCCCACCCC CACTAACCGA
```

```
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGAGTACTTA TCCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGGGTACTTA TCCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGGGTACTTA TCCCCACCCC CACTAACCGA
CCCTTAATTG CTCTAGCACT CAGCCTTCCA TGAGTACTCA TCCCCACCCC CACCAACCGA
CCCTTAATTG CTCTAGCACT CAGCCTCCCA TGAGTACTCA TCCCCACCCC CACAAACCGA
CCCTTAATTG CTCTAGCGCT CAGCCTTCCA TGAGTACTTA TCCCCACCCC CACTAACCGA
CCCTTAATTG CTCTAGCACT CAGCCTTCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGAATACTCA TCCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCCTTAATTG CCCTAGCACT CAGCCTTCCA TGAGTACTCA TCCCCACCCC CACTAACCGA
CCTCTAATTG CCCTAGCACT TAGCCTCCCA TGAATCCTCA TCCCCACCCC CACTAACCGA
CCCCTAATTG CCCTAGCACT CAGCCTCCCA TGAATCCTCA TCCCCACCCC CACCAACCGA

TGAATAAATA ACCGACTGAT TACCGTTCAA TCCTGGTTTA TTGCTCGAGC CACTAGTCAA
TGAGTAAATA ATCGCCTAAT TACTATACAA TCCTGATTTA TTTCCCGTTC TACAAGCCAA
TGACTAAACA ACCGCCTCTT AACCCTCCAA GGTTGATTCA TCAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT GACCCTCCAA GGTTGATTTA TTAACCGCTT TACACAACAA
TGACTAAACA ATCGCCTCTT GACCCTCCAA GGCTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ATCGCCTCTT GACCCTCCAA GGCTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ATCGCCTCTT GACCCTCCAA GGCTGATTCA TTAACCGCTT TACACAACAA
TGGCTGAACA ACCGTCTTTT AACCCTCCAA GCTTGATTTA TTAACCGCTT TACACAACAA
TGGCTGAACA ACCGTCTTTT AACTCTCCAA GGTTGATTTA TTAACCGCTT TACACAACAA
TGGCTGAACA ACCGTCTTTT AACTCTCCAA GGTTGATTTA TTAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT GACCCTTCAA GGCTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT GACCCTTCAA GGCTGATTCA TTAACCGCTT TACACAACAA
TGGCTAAACA ACCGCCTCTT AACCCTTCAA GGTTGATTCA TCAACCGCTT TACACAACAA
TGGCTAAACA ACCGCCTCTT AACCCTTCAA GGTTGATTCA TCAACCGCTT TACACAACAA
TGGCTAAACA ACCGCCTCTT GACCCTCCAG GGTTGATTCA TTAACCGATT TACACAACAA
TGGCTAAACA ACCGCCTCTT GACCCTCCAA GGTTGATTCA TTAACCGATT TACACAACAA
TGGCTAAACA ACCGCCTCTT GACCCTCCAA GGTTGATTCA TTAACCGATT TACACAACAA
TGACTAAACA ACCGCCTCTT GACCCTCCAA GGTTGATTTA TTAACCGCTT TACACACCAA
TGGCTAAACA ACCGCCTTTT GACCCTCCAA GGTTGATTCA TTAACCGCTT TACACAGCAA
TGGCTAAACA ACCGCCTCTT GACCCTCCAA GGTTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT AACCCTCCAA GGTTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT AACCCTCCAA GGTTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ACCGCCTCTT AACCCTCCAA GGTTGATTCA TTAACCGCTT TACGCAACAA
TGACTAAACA ACCGCCTCTT AACCCTCCAA GGTTGATTCA TTAACCGCTT TACACAACAA
TGACTAAACA ATCGCCTCTT GACCCTACAA GCCTGATTTA TTAACCGATT CACACAACAA
TGACTAAACA ACCGCCTATT AACCCTCCAA GCCTGATTTA TTAACCGATT TACACAACAA

CTCATACTAC CACTAAATAT AGGCGCACAT AAATGGGCCA TGATCTTAAC CGCCCTCCTC
CTCATATTAC CCCTAAACAT TGGAGCCCAT AAATGAGCCA TAATTTTAAC CGCCCTCCTT
CTCATACTAC CCATCAATCT AGGCGGGCAC AAGTGAGCTG TTCTGTTAAC AGCCCTAATG
CTCATACTAC CCATCAACCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTTATACTGC CCATCAACCT CGGTGGACAC AAATGAGCTG TTCTACTAAC AGCCCTAATA
CTTATACTGC CCATCAACCT CGGTGGACAC AAATGAGCTG TTCTACTAAC AGCCCTAATA
CTTATACTGC CCATCAACCT CGGTGGACAC AAATGAGCTG TTCTACTAAC AGCCCTAATA
CTCATGCTAC CCATCAACCT TGGCGGACAC AAATGAGCTG TTTTATTAAC AGCCCTAATG
CTCATGCTAC CCATCAACCT TGGCGGACAC AAATGAGCTG TTTTATTAAC AGCCCTAATG
CTCATGCTAC CCATCAACCT TGGCGGACAC AAATGAGCTG TTTTATTAAC AGCCCTAATG
CTCATACTAC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTCTTAAC AGCCCTAATA
CTCATACTAC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTCTTAAC AGCCCTAATA
CTCATACTGC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTCCTAAC AGCCCTAATA
CTCATACTGC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTCCTAAC AGCCCTAATA
CTCATACTAC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTCATACTAC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTCATACTAC CCATTAACCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTCATATTAC CCATCAATCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTCATACTGC CCATCAATCT CGGCGGACAC AAATGAGCTG TTCTGTTGAC AGCCCTAATA
CTCATACTAC CCATCAATCT CGGCGGACAC AAATGAGCTG TTCTGTTAAC AGCCCTAATG
CTCATACTAC CCATCAATCT CGGCGGGCAC AAATGAGCTG TTCTGTTAAC AGCCCTAATA
CTCATACTAC CCATCAATCT CGGCGGACAC AAATGAGCTG TTCTATTAAC AGCCCTAATG
CTCATACTAC CCATCAATCT CGGCGGGCAC AAATGAGCTG TTCTATTAAC AGCCCTAATG
CTCATACTAC CCATCAATCT CGGCGGGCAC AAATGAGCTG TTCTGTTAAC AGCCCTAATG
CTCATGCTTC CTATCAATCT CGGCGGCCAC AAATGAGCTG TCTTACTGAC AGCCCTAATG
CTAATACTCC CCATCAACCT CGGCGGGCAC AAATGAGCCA TACTATTAAC AGCCTTGATA
```

CTATTTTTAA TAACCTTAAA TCTTCTAGGC CTTCTGCCGT ATACTTTTAC TCCGACCACT
CTCTTCCTAA TTACATTAAA CCTTTTAGGC CTCCTACCAT ACACATTTAC CCCAACCACC
CTACTCTTAA TCACACTTAA CCTACTCGGC CTCCTCCCCT ACACCTTCAC CCCTACAACC
TTACTCTTAA TTACACTCAA GCTACTCGGT CTCCTCCCCT ACACCTTCAC CCCTACAACT
CTACTTTTAA TTACACTCAA CCTGCTAGGC CTCCTCCCCT ATACCTTTAC CCCTACAACT
CTACTTTTAA TTACACTCAA CCTGCTAGGC CTCCTCCCCT ATACCTTTAC CCCTACAACT
CTACTTTTAA TTACACTCAA CCTGCTAGGC CTCCTCCCCT ATACCTTTAC CCCTACAACT
CTACTTTTAA TTACACTCAA TCTACTCGGG CTCCTCCCCT ACACCTTTAC TCCTACAACT
CTACTTTTAA TTACACTCAA CCTACTCGGG CTCCTCCCTT ACACCTTTAC TCCTACAACT
CTACTTTTAA TTACACTCAA CCTACTCGGG CTCCTCCCTT ACACCTTTAC TCCTACAACT
CTACTTTTAA TTACACTCAA CCTACTTGGC CTCCTCCCCT ACACCTTCAC CCCTACAACC
CTACTTTTAA TTACACTCAA CCTACTTGGC CTCCTCCCCT ACACCTTCAC CCCTACAACC
CTACTTTTAA TCACACTCAA CCTACTCGGC CTCCTCCCCT ACACCTTTAC CCCTACAACT
CTACTTTTAA TCACACTCAA CCTACTCGGC CTCCTCCCCT ACACCTTTAC CCCTACAACT
CTACTTTTAA TTACACTTAA TCTACTCGGC CTCCTCCCCT ACACCTTTAC CCCCACAACT
CTACTTTTAA TTACACTTAA CCTACTCGGC CTCCTCCCCT ACACCTTTAC CCCCACAACT
CTACTTTTAA TTACACTTAA CCTACTCGGC CTCCTCCCCT ACACCTTTAC CCCCACAACT
CTACTCTTAA TCACACTCAA CCTACTCGGC CTCCTCCCCT ACGCCTTCGC CCCTACAACT
CTACTCTTAA TCACACTCAA CCTACTCGGC CTCCTCCCCT ACACCTTCAC CCCTACAACT
TTACTCCTAA TCACACTCAA CCTACTCGGC CTCCTCCCCT ATACCTTCAC CCCTACAACT
CTACTCTTAA TCACACTTAA CCTACTCGGC CTTCTCCCCT ACACCTTCAC CCCTACAACT
CTACTCTTAA TTACACTTAA CCTGCTCGGC CTCCTCCCCT ACACCTTCAC CCCTACAACT
CTACTCTTAA TTACACTTAA CCTACTCGGC CTCCTCCCCT ACACCTTCAC CCCTACAACT
CTACTCTTAA TTACACTTAA CCTACTCGGC CTCCTCCCCT ATACCTTCAC CCCTACAACT
CTACTTTTGA TTACACTCAA TCTACTGGGC CTACTTCCAT ATACCTTTAC TCCTACAACA
CTACTCTTAA TCACACTCAA CCTACTAGGC CTACTTCCAT ACACCTTTAC CCCTACAACA

CAATTATCCA TAAATATAGC CCTGGCCGTA CCATTATGAC TGGCCACTGT ATTAATTGGC
CAACTATCTA TAAATATAGC ACTCGCTGTA CCATTATGAT TAGCTACCGT CCTAATTGGT
CAACTCTCCC TCAACATGGG CTTACCCGTC CCCCTGTGAC TAGCCACCGT AATTATTGGC
CAACTCTCCC TCAACATAGG ACTCGCCGTC CCCCTGTGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TTAACATAGG ACTCGCCGTC CCCCTGTGGC TAGCTACCGT GATTATTGGC
CAACTCTCCC TTAACATAGG ACTCGCCGTC CCCCTGTGGC TAGCTACCGT GATTATTGGC
CAACTCTCCC TTAACATAGG ACTCGCCGTC CCCCTGTGGC TAGCTACCGT GATTATTGGC
CAGCTCTCTC TTAATATAGG ACTCGCCGTC CCTCTGTGAT TGGCTACCGT AATTATTGGC
CAACTCTCTC TTAATATAGG ACTCGCCGTC CCCCTGTGAT TGGCTACCGT AATTATTGGT
CAACTCTCTC TTAATATAGG ACTCGCCGTC CCCCTGTGAT TGGCTACCGT AATTATTGGT
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAT TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAT TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATAGG ACTCGCCGTC CCCCTGTGAT TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATAGG ACTCGCCGTC CCCCTGTGAT TAGCCACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAT TAGCTACCGT AATCATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATAGG ACTCGCCGTC CCCCTGTGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG GTTCGCCGTC CCCCTGTGGC TAGCTACCGT AATTATTGGT
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTGTGAC TAGCTACCGT AATTATTGGT
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTATGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTATGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTATGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG ACTCGCCGTC CCCCTATGAC TAGCTACCGT AATTATTGGC
CAACTCTCCC TCAACATGGG GTTCGCCGTC CCCCTATGAC TGGCCACCGT GATTATTGGC
CAACTCTCCC TCAACATAGG ATTTGCCGTC CCCCTATGAC TAGCCACCGT AGTCATTGGC

ATACGAAATC AGCCAACACA TTCTTTAGCT CACCTCCTCC CAGAAGGAAC GCCAACCCCT
ATACGAAATC AACCAACCCA CTCTTTGGCT CACCTTCTAC CAGAAGGTAC ACCAACCTTA
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCGTTCCC
ATACGAAACC AACCCACTGC CGCCCTGGGC CACCTGCTGC CAGAAGGAAC CCCCGTTCCT
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTAC CAGAAGGAAC CCCTATCCCT
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTAC CAGAAGGAAC CCCTATCCCT
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTAC CAGAAGGAAC CCCTATCCCT
ATACGAAACC AGCCCACCGC CGCCCTAGGC CACCTTCTGC CAGAAGGAAC CCCCGTCCCC
ATACGAAACC AGCCCACCGC CGCCCTAGGC CACCTTCTGC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AGCCCACCGC CGCCCTAGGC CACCTTCTGC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTAC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTAC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCATTCCC

```
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGTTAC CAGAAGGAAC CCCCGTTCCC
ATGCGAAACC AACCCACCGC CGCCCTAGGC CACCTATTAC CAGAAGGAAC CCCTGTTCCC
ATGCGAAACC AACCCACCGC CGCCCTAGGC CACCTATTAC CAGAAGGAAC CCCTGTTCCC
ATACGAAACC AACCCACTGC CGCCCTGGGC CACCTGCTGC CAGGAGGAAC CCCCGTTCCT
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGGAC CCCCGTTCCC
ATACGAAACC AACCTACCGC CGCCCTAGGC CACCTACTGC CAGAAGGGAC CCCCGTTCCC
ATACGAAACC AACCCACCGC TGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCATTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCGTTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCGTTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTGCTGC CAGAAGGAAC CCCCGTTCCC
ATACGAAACC AACCCACCGC CGCCCTAGGC CACCTACTAC CAGAGGGAAC CCCAGTGCCC
ATGCGAAACC AACCCACCGC TGCCCTAGGC CATCTACTAC CAGAAGGAAC CCCAGTACCC

CTAATTCCCA TTCTAATTAT TATCGAGACA ATTAGCCTAT TTATTCGACC CTTGGCCCTT
CTAATCCCAG TCTTAATCAT TATCGAAACG ATTAGCCTAT TCATCCGACC ATTAGCCCTA
CTCATCCCTG TCTTAATTAT TATCGAAACA ATCAGCCTAT TTATCCGCCC ACTAGCGCTA
CTTATCCCAG TCTTAATTCT TATCGAAACA ATCAGCTTAT TTATCCGACC CCTAGCACTA
CTCATTCCCG TCCTAATTAT CATCGAAACA ATCAGCCTAT TTATCCGTCC CCTGGCACTA
CTCATTCCCG TCCTAATTAT CATCGAAACA ATCAGCCTAT TTATCCGTCC CCTGGCACTA
CTCATTCCCG TCCTAATTAT CATCGAAACA ATCAGCCTAT TTATCCGTCC CCTGGCACTA
CTCATCCCGG TCTTAATCAT TATCGAAACA ATCAGCTTAT TTATTCGCCC CCTAGCATTA
CTCATCCCGG TCTTAATCAT TATCGAAACA ATCAGCTTAT TTATCCGCCC CCTAGCGTTA
CTCATCCCGG TCTTAATCAT TATCGAAACA ATCAGCTTAT TTATCCGCCC CCTAGCGTTA
CTCATCCCAG TCTTAATTGT TATCGAAACA ATCAGCCTAT TTATTCGCCC CCTGGCACTA
CTCATCCCAG TCTTAATTGT TATCGAAACA ATCAGCCTAT TTATTCGCCC CCTGGCACTA
CTCATCCCAG TCTTAATCAT TATCGAAACA ATCAGCCTAT TTATTCGCCC CCTGGCATTA
CTCATCCCGG TCTTAATCAT TATCGAAACA ATCAGCCTAT TTATTCGCCC CCTGGCATTA
CTCATCCCAG TCTTAATCAT TATCGAAACA ATCAGCCTAT TTATTCGCCC CCTAGCATTA
CTCATCCCAG TCTTAATCAT TATCGAAACA ATCAGCCTAT TCATTCGCCC CCTAGCATTA
CTCATCCCAG TCTTAATCAT TATCGAAACA ATCAGCCTAT TCATTCGCCC CCTAGCATTA
CTTATCCCAG TCTTAATCAT TATCGAAACA ATCAGCCTAT TTATCCGACC CCTAGCACTA
CTCATCCCAG TCTTAATCGT TATCGAAACA ATCAGCCTAT TCATCCGCCC CCTGGCACTA
CTCATCCCAG TTTTAATTAT TATCGAAACG ATCAGCCTAT TCATCCGCCC CCTAGCACTA
CTCATCCCTG TCTTAATTAT TATCGAAACA ATCAGCCTAT TTATCCGCCC ACTGGCACTA
CTCATCCCTG TCTTAATTAT TATCGAAACA ATCAGCCTAT TTATCCGCCC ACTGGCACTA
CTCATCCCTG TCTTAATTAT TATCGAAACA ATCAGCCTAT TTATCCGCCC ACTGGCACTA
CTCATCCCTG TCTTAATTAT TATCGAAACA ATCAGCCTAT TTATCCGCCC ACTGGCACTA
CTCATTCCAG TCTTGATTGT CATCGAAACA ATTAGCCTAT TTATCCGCCC CCTGGCCTTA
CTCATCCCAG TTTTAATCAT TATTGAAACA ATTAGCCTAT TTATCCGCCC CCTGGCCCTA

GGAGTTCGAC TAACAGCAAA TCTAACCGCA GGCCACCTAT TAATTCAACT AATTTCTACA
GGCGTTCGAC TAACTGCAAA CCTTACCGCA GGACACCTAC TTATTCAATT AATTTCTACA
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TGATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCGAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTCCGGC TCACAGCAAA CCTGACTGCA GGACATCTGT TAATTCAACT GATTGCCACC
GGCGTCCGGC TCACAGCAAA CCTGACTGCA GGACATCTGT TAATTCAACT GATTGCCACC
GGCGTCCGGC TCACAGCAAA CCTGACTGCA GGACATCTGT TAATTCAACT GATTGCCACC
GGCGTTCGAC TTACAGCAAA CCTAACTGCA GGCCATCTCT TAATTCAACT GATTGCCACC
GGCGTTCGAC TTACAGCAAA CCTAACTGCA GGCCATCTCT TAATTCAACT GATTGCCACC
GGCGTTCGAC TTACAGCAAA CCTAACTGCA GGCCATCTCT TAATTCAACT GATTGCCACC
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGCCATCTGT TAATTCAACT GATCGCCACC
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGCCATCTGT TAATTCAACT GATCGCCACC
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGCCATCTAT TAATTCAACT GATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGCCATCTAT TAATTCAACT GATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTGACTGCA GGCCACCTAC TAATTCAACT GATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCG GGCCACCTAT TAATTCAACT GATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCG GGCCACCTAT TAATTCAACT GATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGACATCTAT TAATTCAACT GATTGCCACT
GGCGTTCGAC TTACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTTCGAC TCACGGCAAA CCTAACTGCA GGTCATCTAT TAATTCAACT AATTGCCACT
GGCGTCCGAC TCACAGCAAA CCTAACTGCA GGCCATCTAT TAATCCAACT AATTGCTACC
GGTGTCCGAC TTACAGCAAA TCTAACTGCA GGCCACTTAT TAATCCAACT AATCGCCACC
```

```
GCAACCTTTG TTATACTGTC AACTATACCA ACAATTGCCG TACTTACACT TACTGTACTA
GCAACCTTTG TTATATTATC TATTATACCA ACAGTTGCTA TATTAACACT TATTGTATTA
GCCGCCTTCG TACTCCTTCC AATAATACCG GCCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCCTTCG TACTCCTCCC AATAATACCA GCCGTGGCTA TTTTGACATC AACAGTGCTA
GCCGCCTTCG TACTCCTCCC AATAATGCCA ACCGTGGCTA TTTTAACATC AACGGTGCTA
GCCGCCTTCG TACTCCTCCC AATAATGCCA ACCGTGGCTA TTTTAACATC AACGGTGTTA
GCCGCCTTCG TACTCCTCCC AATAATGCCA ACCGTGGCTA TTTTAACATC AACGGTGTTA
GCCGCATTTG TCCTCCTTCC AATAATACCG GCCGTAGCTA TCTTAACATC AACAGTACTA
GCCGCATTTG TTCTCCTTCC AATAATGCCG GCCGTAGCCA TCTTAACATC AACAGTACTA
GCCGCATTTG TTCTCCTTCC AATAATGCCG GCCGTAGCCA TCTTAACATC AACAGTACTA
GCCGCATTTG TACTCCTTCC AATAATGCCT ACCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCATTTG TACTCCTTCC AATAATGCCT ACCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCATTTG TACTCCTTCC AATAATACCG ACCGTAGCTA TCTTAACATC AATAGTGCTG
GCCGCATTTG TACTCCTTCC AATAATGCCG ACCGTAGCTA TTTTAACATC AATAGTGCTG
GCCGCATTTG TACTCCTCCC AATAATACCG ACCGTAGCTA TTTTAACATC AACAGTGCTG
GCCGCATTTG TACTCCTCCC AATAATACCG ACCGTAGCTA TTTTAACATC AACGGTGCTG
GCCGCATTTG TACTCCTCCC AATAATACCG ACCGTAGCTA TTTTAACATC AACAGTGCTG
GCCGCCTTCG TACTCCTCCC AATAATACCA GCCGTGGCTA TTTTGACATC AACAGTGCTA
GCCGCCTTCG TACTCCTTCC AATAATACCG ACCGTAGCTA TTTTAACATC AACAGTGCTG
GCCGCCTTCA TACTCCTTCC AATAATACAT GCCGTAGCTA TTTTAACATC AACGGTGTTA
GCCGCCTTCG TACTCCTTCC AATAATACCG GCCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCCTTCG TACTCCTTCC AATAATACCG GCCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCCTTCG TACTCCTTCC AATAATACCG GCCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCCTTCG TACTCCTTCC AATAATACCG GCCGTAGCTA TTTTAACATC AACAGTGCTA
GCCGCATTCG TACTCCTTCC AATAATGCCA ACCGTGGCCA TTTTAACATC AGTAGTCCTA
GCCGCATTCG TACTCCTTCC GATAATGCCA ACCGTGGCCA TCTTAACATC AGTAGTCCTA

GCCCTACTAA CCATTTTAGA AATTGCTGTC GCAATAATCC AAGCTTACGT ATTTGTTCTT
GCCCTATTAA CTATTCTAGA AATCGCTGTT GCCATAATCC AAGCATATGT ATTTGTTCTT
TTTTTACTAA CCCTGCTAGA AGTGGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTACTAA CTCTACTGGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTCCTACTAA CCCTACTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTC
TTCCTACTAA CCCTATTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTC
TTCCTACTAA CCCTATTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTC
TTCCTATTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AGGCATACGT ATTCGTCCTT
TTCCTATTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AGGCATACGT ATTCGTCCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTCGTTCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTCGTTCTT
TTCCTACTAA CCCTGCTAGA AGTGGCCGTA GCAATAATTC AAGCATACGT ATTCGTTCTT
TTCCTACTAA CCCTGCTAGA GGTAGCCGTA GCAATAATTC AAGCATACGT ATTCGTTCTT
TTCCTACTAA CCCTGCTAGA GGTAGCCGTA GCAATAATTC AAGCATACGT ATTCGTTCTT
TTTCTACTAA CTCTACTGGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATCC AAGCATACGT ATTTGTTCTT
TTTCTACTAA CCCTACTAAA AGTAGCTGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTGCTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTACTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTTCTGCTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTTCTT
TTCCTACTAA CCCTGCTAGA AGTAGCCGTG GCAATAATCC AAGCATACGT ATTTGTTCTC
TTTCTCCTAA CCCTGCTAGA AGTAGCCGTA GCAATAATTC AAGCATACGT ATTTGTCCTC

TTATTAAGCC TGTACCTACA AGAAAATGTC TAATGGCCCA TCAAGCACAC GCATACCACA
CTATTAAGCC TCTATCTACA AGAAAACGTT TAATGGCCCA TCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACTTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TTTACCTACA AGAAAACGTT TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TTTACCTACA AGAAAACGTT TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TTTACCTACA AGAAAACGTT TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTATCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
```

```
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTT TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACTTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTT TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGTC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTAAGCC TCTACCTACA AGAAAACGTC TAATGGCCCG CCAAGCACAC GCATATCACA
CTACTGAGCC TCTATCTACA AGAGAACGTC TAATGGCCCA CCAAGCACAC GCATATCACA

TGGTTGACCC AAGCCCATGA CCCCTTACTG GGGCAGTCGC CGCTTTGCTA TTAACTTCAG
TGGTTGACCC AAGCCCATGA CCCTTAACTG GGGCAGTTGC CGCACTCCTA TTAACCTCCG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GAGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CTGCGCCTGA GCCGTAACAG GCGCAGTAGC TGCCCTGCTG ACCACATCAC
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCAG
TGGTCGACCC CAGCCCCTGA CCCTTAACCG GCGCAGTAGC CGCCCTCCTG ATGACATCAG
TGGTCGACCC CAGCCCCTGA CCCTTAACCG GCGCAGTAGC CGCCCTCCTG ATGACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATGACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATGACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTAGC TGCCCTCCTG ATAACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GCGCAGTAGC TGCCCTCCTG ATAACATCGG
TGGTCGACCC CAGCCCCTGA CCCCTAACCG GTGCAGTAGC TGCCCTCCTG ATGACATCAG
TGGTCGACCC CAGCCCCTGA CCCCTGACCG GCGCAGTTGC CGCCCTCCTG ATAACGTCAG

GACTCGCAGT TTGATTTCAC TTTAAGTCAA TAACCCTACT AGCAATAGGC CTACTACTAA
GATTAGCAGT ATGATTCCAC TTCAAAACAA TAACCTTATT AGCAATAGGC CTCCTATTAA
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACCCGCACC CTCATTCCAA TTCAACACCA CAGCACTAAT AACAATAACA CTCAACCTAC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCATCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCATCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCATCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT GACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAT TTTAATTCCA CAGTACTTAT GACAATAGGA CTCACCCTAC
GACTTGCAGT CTGATTCCAT TTTAATTCCA CAGTACTTAT GACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACGATAGGA CTCACCCTCC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACGATAGGA CTCACCCTCC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTAC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTAC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTCC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATGGGA CTTACCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATGGGA CTTACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTGCTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAT TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCACCCTGC
GACTTGCAGT CTGATTCCAC TTTAACTCCA CAGTACTTAT AACAATAGGA CTCGCCTTGC
```

```
GACTCGCAGT CTGATTCCAC TTCAACTCCA CAGTACTTAT AACAATAGGA CTCATCTTGC

TAATTCTTAC TATAATTCAA TGATGACGAG ACATTATCCG TGAAGGCACA TTTCAAGGAC
TAATTTTAAC TATAATTCAA CGATGACGAG ATATTATTCG CGAGGGAACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATGTACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATATCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATATCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATATCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGGC
TTCTCCTAAC CATGTACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATGTACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATCATTCG AGAAGGCACA TTTCAGGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGGC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGGC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAAGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAGGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAGGGCACA TTTCAAGGAC
TTCTCCTAAC CATATACCAA TGATGACGAG ATATTATTCG AGAGGGCACA TTTCAAGGAC
TTCTCTTAAC CATGTACCAA TGATGACGAG ACATTATTCG AGAAGGCACA TTTCAAGGAC
TCCTTCTAAC CATATACCAA TGATGACGAG ACATCATTCG AGAAGGCACA TTCCAAGGAC

ATCACACCCC CCCAGTCCAA AAAGGCCTGC GTTACGGAAT AATCTTATTT ATTACATCCG
ATCATACTCC CCCTGTACAA AAAGGTCTAC GCTACGGAAT AATTCTATTC ATTACATCCG
ACCACACACC CCCTGTCCAA AAAGGGCTTC GATACGGAAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ACCACACGCC TCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTGTTT ATTACCTCAG
ACCACACGCC TCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTGTTT ATTACCTCAG
ACCACACGCC TCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTGTTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ACCACACACC TCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGGCTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ATCACACACC CCCTGTCCAA AAAGGGCTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ATCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ATCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATTCTATTT ATTACCTCAG
ATCACACACC CCCTGTCCAA AAAGGACTTC GATACGGAAT AATCCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATATGGGAT AATCCTATTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTCC GGTACGGGAT AATTCTATTT ATTACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGGAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGGCTCC GATACGGGAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGGCTTC GATACGGGAT AATTCTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAAGGACTTC GATACGGGAT AATTCTATTT ATCACCTCAG
ACCATACACC CCCTGTCCAA AAAGGACTCC GATATGGAAT AATTTTATTT ATCACCTCAG
ACCACACACC CCCTGTCCAA AAGGGGCTCC GATATGGAAT AATCCTATTT ATTACTTCAG

AAGTGTTCTT CTTCCTGGGA TTTTTCTGAG CCTTCTACCA CTCAAGCCTA GCCCCCACCC
AAGTATTCTT TTTCCTAGGC TTCTTCTGAG CATTTTATCA TTCAAGTTTA GCCCCTACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGCCTA GCCCCGACAC
AAGTCTTCTA TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGGG CATTTTACCA TGCGAGCCTG GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGGG CATTTTACCA TGCGAGCCTG GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGGG CATTTTACCA TGCGAGCCTG GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CCTTTTACCA CGCAAGCCTG GCCCCAACTC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CCTTTTACCA CGCAAGCCTA GCCCCAACTC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CCTTTTACCA CGCAAGCCTA GCCCCAACTC
AAGTTTTCTT TTTCCTAGGC TTCTTCTGAG CATTCTACCA CGCAAGCCTG GCCCCAACAC
```

```
AAGTTTTCTT CTTCCTAGGC TTCTTCTGAG CATTCTACCA CGCAAGCCTG GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACGC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACGC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTTTACCA CGCAAGCCTA GCCCCAACAC
AAGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTTTATCA CGCAAGCCTA GCCCCGACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGTCTG GCCCCGACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGTCTG GCCCCTACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGTCTG GCCCCTACAC
AAGTTTTCTT TTTCCTGGGC TTTTTCTGAG CATTTTACCA CGCAAGTCTG GCCCCGACAC
AAGTTTTCTT CTTTCTAGGC TTTTTCTGAG CATTCTACCA CTCGAGCTTG GCACCCACAC
AGGTTTTCTT TTTCCTAGGC TTTTTCTGAG CATTCTACCA CTCAAGCCTG GCCCCAACGC

CAGAACTAGG AGGCATTTGA CCGCCGACAG GCATCATCCC TTTAGACCCA TTCGAAGTCC
CAGAACTCGG GGGCATTTGA CCACCCACTG GCATTACACC TTTAGACCCG TTTGAAGTAC
CAGAACTAGG GGGATGCTGA CCCCCAACTG GCATCATCAC CTTGGACCCC TTTGAAGTAC
CAGAACAAGC AGGATCCTGA CCCCCAACTG GCATTATTAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGGTGCTGA CCCCCAACTG GCATTATCAC CTTAGATCCC TTCGAAGTAC
CAGAACTAGG CGGGTGCTGA CCCCCAACTG GCATTATCAC CTTAGATCCC TTCGAAGTAC
CAGAACTAGG CGGGTGCTGA CCCCCAACTG GCATTATCAC CTTAGATCCC TTCGAAGTAC
CAGAACTAGG TGGGTGCTGA CCCCCAACTG GCATTACCAC TTTAGACCCC TTTGAAGTGC
CAGAGCTAGG CGGGTGCTGA CCCCCAACTG GCATTACCAC TTTGGACCCC TTTGAAGTGC
CAGAGCTAGG CGGGTGCTGA CCCCCAACTG GCATTTCCAC TTTGGACCCC TTTGAAGTGC
CAGAACTAGG CGGGTGCTGA CCCCCAACCG GCATTATCAC CTTAGACCCA TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACCG GCATTATTAC CTTAGACCCA TTTGAAGTAC
CAGAACTAGG CGGGTGCTGA CCCCCAACCG GCATTATCAC CTTAGACCCA TTTGAAGTGC
CAGAACTAGG CGGGTGCTGA CCCCCAACCG GCATTATCAC CTTAGACCCA TTTGAAGTGC
CAGAACTAGG CGGGTGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCA TTTGAAGTAC
CAGAACTAGG TGGATGCTGA CCCCCAACTG GCATTATTAC CTTAGACCCA TTTGAAGTAC
CAGAACTAGG TGGATGCTGA CCCCCAACTG GCATTATTAC CTTAGACCCA TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACCG GCATTATTAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAAGTAC
CAGAGCTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAAGTAC
CAGAACTAGG CGGATGCTGA CCCCCAACTG GCATTATCAC CTTAGACCCC TTTGAGGTAC
CAGAGCTAGG GGGATGCTGA CCTCCCACCG GCATCATCCC ACTAGACCCC TTTGAAGTAC
CAGAACTAGG GGGATGCTGG CCTCCCACCG GCATCATTCC ACTAGACCCC TTTGAAGTAC

CCCTCCTAAA TACAGCAGTC TTACTAGCCT CTGGCGTTAC TGTCACATGG GCCCATCACA
CACTTTTAAA CACAGCAGTT CTCCTTGCTT CTGGCGTTAC TGTTACATGA GCCCATCACG
CTCTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA CACAGCAGTA CTGTTAGCCT CTGGCGTCAC GGTCACTTGA GCGCACTCTA
CACTACTTAA TACAGCAGTA CTGCTGGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGCTGGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGCTGGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CGCTACTTAA CACAGCAGTA CTACTAGCCT CCGGCGTCAC AGTCACTTGA GCTCACCACA
CACTACTCAA CACAGCAGTA CTATTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTCAA CACAGCAGTA CTATTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACTTGG GCCCACCACA
CACTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTACTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGCTAGCCT CCGGCGTCAC AGTCACTTGA GCTCACCACA
CACTACTTAA TACAGCAGTA CTGCTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA CACAGCAGTA CTGTTAGCCT CCGGCGTCAC GGTCACTTGA GCCCACCACA
CACTACTTAA CACAGCAGTA CTATTAGCCT CCGGCGTCAC AGTCACTTGA GCCCACCACA
CTCTACTTAA TACAGCAGTA CTGTTGGCCT CCGGTGTCAC AGTCACTTGA GCCCACCACA
CACTACTTAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACCTGA GCCCACCACA
CACTACTCAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACCTGA GCCCACCACA
CACTACTCAA TACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACCTGA GCCCACCACA
CACTACTTAA CACAGCAGTA CTGTTAGCCT CCGGCGTCAC AGTCACCTGA GCCCACCACA
```

```
CACTACTCAA CACAGCAGTA CTACTAGCCT CCGGCGTTAC AGTCACCTGA GCCCACCACA
CCCTACTCAA CACAGCAGTA CTACTAGCCT CTGGCGTTAC AGTCACTTGA GCCCACCACA

GCCTCATGGA AGGAAAACGA ACAGAAGCTA CACAGGCCCT TTCCCTAACC ATTTTACTAG
GCCTAATAGA AGGAAAACGA ACTGAAGCTA CACAAGCGCT TACCTTAACT ATTATTCTTG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTACTGG
GCATCATATG CCGCCGACGA AGACAAAACA TTTAAGCACT AGCCCTAACA ATCCAATTAG
GCATCATGGA ACGCGAACGC AAACAAACCA TCCAAGCGCT AACCCTGACA ATCCTATTAG
GCATCATGGA ACGCGAACGC AAACAAACCA TCCAAGCGCT AACCCTGACA ATCCTATTAG
GCATCATGGA ACGCGAACGC AAACAAACCA TCCAAGCGCT AACCCTGACA ATCCTATTAG
GCATCATAGA ACGCGAACGC AAACAAACCA TCCAGGCATT AACCCTAACA GTCCTACTAG
GCATCATAGA ACGCGAGCGC AAACAAACCA TCCAGGCATT AACCCTAACA GTCCTACTAG
GCATCATAGA ACGCGAGCGC AAACAAACCA TCCAGGCATT AACCCTAACA GTCCTACTAG
GCATCATAGA ACGCGAGCGC AAACAAACCA TCCAAGCGTT GGCCCTAACA ATCCTACTAG
GCATCATAGA ACGCGAGCGC AAACAAACCA TCCAAGCGTT GACCCTAACA ATCCTACTAG
GCATCATAGA ACGCGAACGC AAACAAACCA TCCAAGCATT AACCCTGACA ATCCTATTAG
GCATCATAGA ACGCGAACGC AAACAAACCA TCCAAGCATT AACCCTGACA ATCCTATTAG
GCATCATAGA ACGCGAACGC AAACAAACCA TCCAAGCATT AACCCTGACA ATCCTATTAG
GCATCATAGA ACGCGAACGT AAACAAACCA TCCAAGCATT AACCCTGACA ATCCTATTAG
GCATCATAGA ACGCGAACGT AAACAAACCA TCCAAGCATT AACCCTGACA ATCCTATTAG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCGCT AGCCCTAACA ATCCTATTAG
GCATCATAGA ACGCGAACGC AAACAAACCA TTCAAGCACT AACCCTAACA ATCCTGTTAG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTATTAG
GCATTATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTATTAG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTATTAG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTATTAG
GCATCATAGA GCGCGAACGC AAACAAACCA TTCAAGCACT AGCCCTAACA ATCCTATTAG
GCATCATAGA ACGTGAACGC AAACAAACCA TCCAAGCTCT GACCCTGACC ATCCTGCTAG
GTATTATAGA ACGTGAACGC AAACAAACCA TTCAAGCACT GACCCTAACC ATCCTACTAG

GCCTGTACTT CACTGCCCTC CAAGCAATAG AATATTATGA AGCCCCATTT ACTATTGCAG
GCCTATACTT TACTGCCCTT CAAGCAATAG AATATTACGA AGCCCCATTC ACTATCGCAG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCTCCATTT ACCATCGCCG
AACTCTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATTTCTG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCTCCATTC ACCATCGCTG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCTCCATTC ACCATCGCTG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCTCCATTC ACCATCGCTG
GTCTTTACTT CACAGCCCTC CAAGCAATAG AGTACTACGA AGCCCCATTC ACCATTGCTG
GACTCTACTT CACAGCCCTC CAAGCAATAG AGTACTACGA AGCCCCATTC ACCATTGCTG
GACTCTACTT CACAGCCCTC CAAGCAATAG AGTACTACGA AGCCCCATTC ACCATTGCTG
GGTTTTACTT TACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATCGCTG
GGTTTTACTT TACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATCGCTG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATCGCTG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATCGCTG
GATTCTACTT CACAGCCCTC CAAGCAATAG AATACTACGA AGCCCCATTC ACCATTGCTG
GGTTCTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCCCCATTC ACCATTGCTG
GGTTCTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCCCCATTC ACCATTGCTG
GATTTTACTT CACAGCTCTT CAAGCAATAG AATACTACGA AGCCCCATTT ACCATCGCCG
GATTCTACTT CACAGCCCTC CAAGCAATAG AGTACTACGA AGCTCCATTT ACCATCGCTG
GGTTCTACTT CACAGCCCTT CAAGCAATAG AATACTACGA AGCTCCATTT ACCATCGCCG
GGTTTTATTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCTCCATTT ACCATCGCCG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCTCCATTT ACCATCGCCG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCTCCATTT ACCATCGCCG
GGTTTTACTT CACAGCCCTC CAAGCAATAG AATATTACGA AGCTCCATTT ACCATCGCCG
GGTTCTACTT TACAGCCCTG CAAGCGATAG AATACTACGA AGCCCCATTC ACCATCGCCG
GATTCTACTT TACAGCCCTA CAAGCAATAG AGTACTATGA AGCCCCATTC ACTATTGCTG

ACGGCGTTTA TGGGACAACA TTCTTTGTTG CCACGGGCTT CCACGGGCTT CATGTCATCA
ACGGTGTTTA CGGAACTACC TTCTTCGTCG CTACAGGTTT CCACGGCCTC CATGTTATTA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATTA
ACGGAGTATA CAGCTCCACC TTCCTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGGGTGTA CGGCTCCACC TTCTTTGTTG CAACCGGGTT CCACGGACTT CATGTCATCA
ACGGGGTGTA CGGCTCCACC TTCTTTGTTG CAACCGGGTT CCACGGACTT CATGTCATCA
ACGGGGTGTA CGGCTCCACC TTCTTTGTTG CAACCGGGTT CCACGGACTT CATGTCATCA
ACGGAGTATA TGGCTCCACC TTCTTTGTAG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
```

```
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGATT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGATT CCACGGACTT CACGTCATCA
ACGGAGTATA CGGCTCCACC TTCTTTGTCG CAACCGGATT CCACGGACTT CACGTCATCA
ATGGGGTATA CGGCTCCACC TTCTTCGTCG CAACCGGGTT CCACGGGCTT CACGTCATCA
ATGGAGTGTA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTTATCA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ATGGGGTATA CGGCTCCACC TTCTTTGTCG CAACCGGGTT CCACGGACTT CACGTCATCA
ACGGAGTGTA CGGCTCCACC TTCTTCGTCG CAACAGGATT TCATGGACTC CACGTCATCA
ACGGGGTGTA TGGCTCCACC TTCTTCGTCG CAACCGGATT CCACGGACTC CATGTCATCA

TTGGATCCAC CTTCCTAATA GTATGCCTAC TACGACAAAT CTTATACCAT TTCACCTCTT
TTGGTTCAAC CTTCTTAGCA GTATGCCTCC TACGACAAAT CCTATATCAC TTTACCTCTT
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCCG
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACAGCTG
TTGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAT TTTACATCTG
TTGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAT TTTACATCTG
TTGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAT TTTACATCTG
TTGGCTCCAC CTTCCTAGCG ATCTGCCTCC TCCGACAAGT TCAATACCAC TTTACATCTG
TCGGCTCCAC CTTCCTAGCA ATCTGTCTCC TCCGACAAAT TCAGTACCAC TTTACATCTG
TCGGCTCCAC CTTCCTAGCA ATCTGTCTCC TCCGACAAAT TCAGTACCAC TTTACATCTG
TCGGCTCTAC CTTCCTAGCG ATTTGTCTCC TCCGACAAAT CCAATATCAC TTTACATCTG
TCGGCTCTAC CTTCCTAGCG ATTTGTCTCC TCCGACAAAT CCAATATCAC TTTACATCTG
TCGGCTCTAC CTTCCTAGCG ATTTGCCTCC TCCGACAAAT CCAATATCAC TTTACGTCTG
TCGGCTCTAC CTTCCTAGCG ATTTGCCTCC TCCGACAAAT CCAATATCAC TTTACATCTG
TCGGCTCCAC CTTCCTAGCG ATCTGTCTCC TCCGACAAAT CCAGTACCAC TTTACATCTG
TCGGCTCCAC CTTCCTAGCG ATCTGTCTCC TCCGACAAAT CCAATACCAC TTTACATCTG
TCGGCTCCAC CTTCCTAGCG ATCTGTCTCC TCCGACAAAT CCAATACCAC TTTACATCTG
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATACCAC TTCACATCTG
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCTG
TCGGCTCCAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCTG
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATACCAC TTCACATCTG
TCGGCTCTAC CTTCCTGGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCTG
TCGGCTCTAC CTTCCTGGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCTG
TCGGCTCTAC CTTCCTAGCG GTCTGCCTTC TCCGACAAAT CCAATATCAC TTCACATCTG
TCGGCTCCAC TTTCTTAGCG GTATGTCTTC TCCGACAAAT CCAGTACCAC TTTACATCTG
TCGGCTCTAC CTTTTTAGCA GTCTGCCTCC TCCGACAAAT CCAGTATCAC TTTACATCTG

CCCACCACTT TGGCTTTGAA GCAGCTGCAT GATACTGACA TTTCGTAGAC GTAGTTTGAC
CCCACCATTT TGGCTTCGAA GCTGCTGCAT GATATTGACA TTTTGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTG TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATATTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATATTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATATTGACA CTTCGTAGAT GTAGTTTGAC
AGCACCACTT CGGATTTGAA GCCGCCGCAT GATACTGACA CTTTGTAGAC GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTTGTAGAT GTAGTCTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTTGTAGAT GTAGTCTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTCTGAC
AGCACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTCTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGGTTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTCTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
```

AACACCACTT TGGATTTGAA GCCGCCGCAT GATACTGACA CTTCGTAGAT GTAGTTTGAC
AACACCACTT CGGATTCGAA GCTGCCGCAT GATACTGACA TTTTGTAGAT GTAGTATGAC
AACACCATTT CGGATTTGAA GCCGCCGCAT GATATTGACA CTTTGTAGAC GTAGTCTGAC

TCTTCCTTTA TGTTTCAATC TATTGATGAG GATCTTAATG AACCTAATCT TAACAAT---
TTTTCCTTTA CGTATCTATC TACTGATGAG GATCATAATG AATCTAATCC TAATTATAAT
TATTCTTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGAATCCAAT TAGTCATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTGATTA CAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTGATTA CAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTGATTA CAGCAGTCCT
TATTCCTATA CGTCTCCATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCCATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCCATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TACTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTGTA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTGTA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTGTA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TGTTCCTGTA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TGTTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TGGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTATA CGTCTCTATT TATTGATGAG GATCATAATG AACCTAATTA TAGCAGTCCT
TATTCCTCTA CGTCTCTATC TATTGATGAG GATCATAATG AATCTAGCTA TAACAATCCT
TTTTCCTCTA CGTCTCTATC TATTGATGAG GATCATAATG AACCTAACCA TAGCTGTCCT

---AATTCTA ATTTCCTCAT TAATCTCGAC TATTCTGGCT ATTATTGCAT TCTGACTCCC
CTTAATCTCC TCTATAATC- -----TCAAT TATCCTAGCT ATTATCGCAT TCTGGTTACC
AGCAATTGCA ACCACCCTA- -----TCCTG CATCCTAGCA ATCGTTGCAT TCTGGCTACC
AGCAATTGCA GTCACCCTA- -----ACCTG CATCCTATAA GTCATTGCAT TCTGCCTGCC
AGCAATCGCA ATCACCCTA- -----TCCTG TGTCCTAGCA GTAATTGCAT TCTGACTACC
AGCAATCGCA ATCACCCTA- -----TCCTG TGTCCTAGCA GTAATTGCAT TCTGACTACC
AGCAATCGCA ATCACCCTA- -----TCCTG TGTCCTAGCA GTAATTGCAT TCTGACTACC
AGCAATTACA GTTATCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA ATTACCCTA- -----TCCTG CATCCTAGCA GTCATTGCAT TCTGACTGCC
AGCAATTGCA ATCACCCTA- -----TCCTG CATCCTAGCA GTCGTCGCGT TCTGGTTACC
AGCAATTGCA ATCACCCTA- -----TCCTG CATCCTAGCA GTCGTCGCGT TCTGGTTACC
GGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
GGCAATTGCA ATCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CGTCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG TGTCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG TGTCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTG- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGATTACC
AGCAATCGCA ATCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGCTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA GTCACCCTA- -----TCCTG CATCCTAGCA GTCGTTGCAT TCTGGTTACC
AGCAATTGCA ATCGCCCTA- -----TCCTG CATCTTAGCA ACTGTTGCAT TCTGACTGCC
GACAATTACA ATCATTTTA- -----TCCTG TATCCTAGCA ACCATTGCAT TCTGATTACC

GCAAATAAAC CCAGACTCAG AAAAACTATC GCCGTACGAA TGTGGCTTTG ACCCCCTCGG
TCAAATAAAT CCGGACATAG AAAAATTATC TCCATATGAA TGCGGCTTCG ACCCCCTTGG
ACAAATAAAC CCTGACTCAG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
CCAAAAGAGT CCCGACTCCG AAAAACTTTC CCCCTATGAA TGCGCCTTTG ACCACCTCGG
ACAAATAAAT CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAT CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAT CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATATGC CCCGACTCCG AAAAACTTTC CCCCTATGAA TGCGGCTTTG ACCCCCTTGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG

ACAAATGTGT CCCGACTCCG AAAAACTTTC CCCCTATGAA TGCGGCTTTG ACCCCCTCGG
ACAGACAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAGATAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAGATAAAC CCTGACTCCG AAAAACTTTC CCCCTATGAA TGCGGCTTTG ACCCCCTCGG
ACAGATAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAGATAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAGATAAAC CCTGACTCCG AAAAACTTTC CCCCTACGAA TGCGGCTTTG ACCCCCTCGG
ACAAATAAAT CCCGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAA TGCGGCTTCG ACCCCCTCGG
ACAAATAAAT CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAG TGTGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
ACAAATAAAC CCTGACTCCG AAAAACTCTC CCCCTACGAG TGCGGCTTTG ACCCCCTCGG
TCAAATAAAT CCTGACTCCG AAAAACTTTC TCCATATGAA TGCGGCTTTG ACCCCCTTGG
CCAAATGAAC CCCGACTCCG AAAAACTTTC TCCATATGAA TGCGGTTTTG ACCCCCTCGG

ATCTGCCCGC CTCCCATTTT CTATCCGATT CTTTCTTATT GCCATCCTAT TTCTCCTATT
ATCAGCCCGT CTCCCATTCT CAATTCGATT TTTTCTCGTT GCCATCCTAT TCCTTCTATT
GTCCGCCCGC CTTCCTTTTT CACTGCGATT TTTTTTAGTG GCAATCTTAT TCCTCCTGTT
GTCCGCCCGA CTCCCCTTTT CACTCCGATT TTTTTTAGTG GCAATCTTCT TTCTACTATT
ATCTGCCCGC CTTCCTTTCT CACTGCGATT TTTCCTGGTA GCAATCTTAT TTCTCCTATT
ATCTGCCCGC CTTCCTTTCT CACTGCGATT TTTCCTGGTA GCAATCTTAT TTCTCCTATT
ATCTGCCCGC CTTCCTTTCT CACTGCGATT TTTCCTGGTA GCAATCTTAT TTCTCCTATT
GTCCGCCCGA CTCCCCTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTGCTATT
GTCTGCCCGC CTTCCTTTTT CACTGCGATT CTTTTTAGTG GCAATCTTAT TCCTCCTATT
GTCCGCCCGA CTCCCCTTTT CACTCCGATT TTTCTTAGTG GCAATCTTAT TTCTACTATT
ATCCGCCCGC CTTCCTTTTT CACTACGATT CTTCCTGGTA GCAATCTTAT TTCTTCTATT
ATCCGCCCGC CTTCCTTTTT CACTACGATT CTTCCTGGTA GCAATCTTAT TTCTTCTATT
GTCCGCCCGC CTTCCCTTTT CACTGCGGTT TTTCCTAGTG GCAATCTTAT TTCTTCTATT
GTCCGCCCGC CTTCCCTTTT CACTGCGATT TTTCCTAGTG GCAATCTTAT TTCTTCTATT
GTCCGCCCGC CTTCCCTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTTCTATT
GTCTGCCCGC CTTCCCTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTCCTATT
GTCTGCCCGC CTTCCCTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTCCTATT
ATCTGCCCGC CTTCCTTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTACTATT
GTCCGCCCGC CTTCCTTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TTCTCCTATT
ATCTGCCCGC CTTCCTTTTT CACTGCGATT TTTCTTAGTG GCAATCTTAT TCCTCTTATT
ATCTGCCCGC CTTCCTTTTT CTCTGCGATT CTTTTTAGTA GCAATCTTAT TCCTCCTATT
GTCTGCCCGC CTTCCTTTTT CACTGCGATT CTTTTTAGTG GCAATCTTAT TCCTCCTATT
GTCTGCCCGC CTTCCTTTTT CACTGCGATT CTTTTTAGTG GCAATCTTAT TCCTCCTATT
ATCTGCCCGC CTTCCTTTTT CACTGCGATT CTTTTTAGTG GCAATCTTAT TCCTCCTATT
ATCCGCCCGC CTCCCCTTCT CACTACGATT CTTCCTGGTG GCAATCTTAT TTCTGTTGTT
ATCTGCCCGT CTCCCCTTCT CACTACGATT CTTCCTAGTG GCAATTCTAT TTTTACTATT

TGACCTAGAA ATCGCCCTGC TACTCCCACT CCCCTGAAGC ACTCACCTAG AC---CCCAT
TGACCTAGAA ATTGCTCTTC TCCTACCACT ACCCTGAAGC ATTCACTTAG AC---CCGAC
TGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
TGACTTAAGA ATGGCACTAT AACTCCCCCT ACCCTGGGCA GACCAACTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAC TACTTCCCCT ACCCTGGGGA GATCAGCTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAC TACTTCCCCT ACCCTGGGGA GATCAGCTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAC TACTTCCCCT ACCCTGGGGA GATCAGCTAG CCTCTCCCAC
CGACTTAGAA ATCGCACTAT TGCTCCCCCT ACCCTGGGGG GACCAACTGG CCTCTCCCAC
CGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
CGACTTAGAA ATCGCACTAT TGCTCCCCCT ACCCTGGGGG GACCAACTAG CCTCTCCCAC
TGACCTAGAA ATCGCACTAC TACTCCCCCT ACCATGAGGA GATCAACTAG CCTCTCCCAC
TGACCTAGAA ATCGCACTAC TACTCCCCCT ACCATGAGGA GATCAACTAG CCTCTCCCAC
TGACTTAGAA ATTGCACTAT TACTCCCCCT ACCGTGAGGA GATCAGCTAG CCTCTCCCGC
TGACTTAGAA ATTGCACTAT TACTCCCCCT ACCGTGAGGA GATCAGCTAG CCTCTCCCAC
TGACTTAGAA ATCGCACTAT TGCTCCCCCT ACCATGAGGA GATCAGCTAG CCTCTCCCAC
TGATTTAGAA ATCGCACTAT TGCTCCCCCT ACCGTGAGGA GATCAGCTAG CCTCCCCCAC
TGATTTAGAA ATCGCACTAT TGCTCCCCCT ACCGTGAGGA GATCAGCTAG CCTCCCCCAC
TGACCTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAGCTAG CCTCTCCCAC
TGACTTAGAA ATTGCACTAT TACTCCCCCT ACCGTGGGGA GATCAGCTAG CCTCTCCCAC
TGACTTAGAA ATTGCACTAT TACTTCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC

```
CGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
CGACTTAGAA ATTGCACTAT TACTCCCCCT ACCATGGGGA GATCAACTAG CCTCTCCCAC
CGACCTAGAA ATTGCATTAC TACTCCCCCT ACCCTGAGGA GACCAACTAG CCTCCCCCAC
TGATCTGGAA ATCGCATTAC TGCTCCCCCT ACCCTGAGGA GATCAACTAG CCTACCCCAC

ACTTATATTA ACATGGGCAT TTGTTATTAT TATGCTCTTA ACCTTCGGTT TAGTTTACGA
TCTCGTACTA ACATGAGCAT TCACCATCAT CGTCCTCCTA ACAATTGGCC TAATTTATGA
CGCCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CACCGCCCTA CTCTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCCTTTATGA
CATTACCCTA TTTTGAGCAA CAACCATTCT AATCCTATTA GCCCTAGGCC TCGTTTATGA
CATTACCCTA TTTTGAGCAA CAACCATTCT AATCCTATTA GCCCTAGGCC TCGTTTATGA
CATTACCCTA TTTTGAGCAA CAACCATTCT AATCCTATTA GCCCTAGGCC TCGTTTATGA
CACCGCCCTA CTCTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCGTTTATGA
CGCCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CGCCACCCTA CTCTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCGTCTATGA
CATTGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCATTTATGA
CATTGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCATTTATGA
CATCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCATTTATGA
CATCGCCCTA CTTTGAACAA CAACCATTTT AAGCCTGTTA ACCCTGGGCC TCATTTATGA
CACCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCGTTTATGA
CACCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCGTTTATGA
CACCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTGTTA ACCCTAGGCC TCGTTTATGA
CACCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTACTA ACCCTAGGCC TCATTTATGA
CGCCGCCCTG CTTTGAGCAA CAGCCATTTT AATCCTATTA ACCCTAGGCC TCGTTTATGA
CGCCGCCCTA CTATGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CGCCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CGCCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CGCCGCCCTA CTTTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CGCCGCCCTA CTCTGAGCAA CAACCATTTT AATCCTATTA ACCCTAGGCC TCATTTATGA
CACCGCCCTA TTATGAGCAA TAGCCATCTT AGTCCTACTA ACCCTAGGTC TCATCTATGA
CACCGCCCTA TTATGAGCAA TAACCATTTT AATCCTACTA ACCCTGGGTC TTGTCTATGA

ATGACTTCAA GGAGGATTAG AATGAGCAGA AATGACCCAT ATCTTATTTA CCATCTCCAC
ATGACTACAA GGAGGCCTAG AATGAGCAGA AATGACCCAT ATTTTATTTA CCCTAACTGC
ATGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ACGAACCCAA GGAGGGCTTG AATGCCCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACTCAA GGGGGACTCG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACTCAA GGGGGACTCG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACTCAA GGGGGACTCG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGGACCCAA GGAGGACTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
GTGGACCCAG GGAGGACTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
GTGAACCCAA GGAGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
GTGAACCCAA GGGGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
GTGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
GTGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGACTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGAGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGGGGGCTTG AATGGGCCGA AATGACCCCT GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGACTTG AATGGGCTGA AATGACCCCA GTACACTTCA GCTTCAGCTC
ATGAACCCAA GGAGGGCTAG AATGAGCTGA AATGACCCCA GTACACTTCA GCTTCAGCTC

TGCATTCATA CTCGGACTTT CAGGTTTAAC CTTTAACCGC ACACACCTAC TTTCTGCCCT
CGCCTTTACC CTGGGACTCT CAGGCCTAAC TTTTAACCGA ACACACCTAC TATCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTGAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTGAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTGAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
```

CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGACTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGACTAAC CTTCCACCGA ACCCACCTCC TCTCTGCACT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCTGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATA TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTGAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
CGCCTTCATG TTAGGACTAA TAGGATTAAC CTTCCACCGA ACCCACCTCC TCTCCGCCCT
TGCCTTCATA CTAGGACTTA TAGGATTGAC CTTCCATCGA ACCCACCTTC TCTCTGCCCT
CGCCTTCATA CTAGGACTTA TGGGACTAAC CTTTCACCGA ACCCACCTTC TCTCTGCCCT

TCTATGTCTT GAAGGAATAA TATTATCCTT ATTCATCGCC CTCGCAATAT GGTGCTCCCA
CTTATGTCTT GAAGGAATAA TATTATCCTT GTTTATTGCC TTAACAATAT GATGTTCTCT
TCTCTGCCTA GAAGGAATGA TACTATCTCT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTGTCTTT ATTTATCGCC CTCTCCCTCT GATCACTTCA
TCTCTGCCTA GAGGGAATAA TATTATCTTT ATTTATTGCC CTCTCCCTCT GGTCCCTTCA
TCTCTGCCTA GAGGGAATAA TATTATCTTT ATTTATTGCC CTCTCCCTCT GGTCCCTTCA
TCTCTGCCTA GAGGGAATAA TATTATCTTT ATTTATTGCC CTCTCCCTCT GGTCCCTTCA
TCTCTGCCTA GAAGGGATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTG GAGGGGATAA TACTGTCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTG GAGGGGATAA TACTGTCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTTT GTTTATTGCC CTCTCCCTCT GATCTCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTTT GTTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTTT GTTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTGTCTTT ATTTATCGCC CTCTCCCTCT GATCACTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATCGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATAA TACTATCTTT ATTTATTGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTCT ATTTATCGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTCT ATTTATCGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTCT ATTTATCGCC CTCTCCCTCT GATCCCTTCA
TCTCTGCCTA GAAGGAATGA TACTATCTCT ATTTATCGCC CTCTCCCTCT GATCCCTTCA
CCTCTGCCTA GAAGGAATAA TATTATCCCT ATTTATTGCC CTCTCCCTCT GGTCCCTCCA
TCTCTGCCTA GAAGGAATAA TACTATCTCT ATTTATTGCT CTCTCCCTCT GGTCCCTTCA

AAATGAAACT ATAATATTCT CCTCCGCCCC TCTTCTACTT CTAGCCCTCT CGGCATGTGA
AAATGACATA ATTATATTCT CCTCCGCACC ACTCCTACTT TTAGCCCTAT CGGCATGCGA
ACTAGAATCC ACCACCTATG CCACCGCCCC AATACTACTA CTAGCATTCT CGGCATGCGA
ACTAGAATCC ACCGCCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ACTAGAGTCT ACCGCCTACG CCGCCGCCCC AATACTGCTA CTAGCATTCT CAGCATGCGA
ACTAGAGTCT ACCGCCTACG CCGCCGCCCC AATACTGCTA CTAGCATTCT CAGCATGCGA
ACTAGAGTCT ACCGCCTACG CCGCCGCCCC AATACTGCTA CTAGCATTCT CAGCATGCGA
ACTAGAATCC ACCACCTATG CCACCGCCCC AATGCTACTA CTGGCATTCT CAGCATGTGA
ACTAGAATCC ACCACCTATG CCGCTGCCCC AATACTACTA CTAGCATTCT CAGCATGTGA
ACTAGAATCC ACCACCTATG CCGCTGCCCC AATACTACTA CTAGCATTCT CAGCATGTGA
ACTAGAATCA ACTACCTACG CCACCGCCCC CATACTGCTA CTAGCATTCT CAGCATGTGA
ACTAGAATCT ACTACTTACG CCACCGCCCC GATACTGCTA CTAGCATTCT CAGCATGTGA
ACTAGAATCT ACTACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CAGCATGTGA
ACTAGAATCT ACTACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CAGCATGTGA
ACTAGAATCT ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CAGCATGTGA
ATTAGAATCT ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CAGCATGCGA
ATTAGAATCT ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CAGCATGCGA
ACTAGAATCC ACCGCCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTACTA CTAGCATTTT CGGCATGTGA
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA

```
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ACTAGAATCC ACCACCTACG CCACCGCCCC AATACTGCTA CTAGCATTCT CGGCATGTGA
ATTAGAATCA ACCACCCATG CCACCGCCCC CATACTCCTA CTTGCATTTT CGGCATGCGA
ATTAGAATCA ACTACCTATG CCACCGCCCC CATACTCCTA CTAGCATTTT CAGCATGCGA

AGCCGGCCTT GGCCTAAGCT TACTAGTCGC CACCTCTCGC ACACACGGGT CTGATCATCT
AGCAGGCCTA GGCCTAAGCT TATTAGTTGC CACATCCCGC ACCCACGGAT CTGATCACCT
AGCCGGAGCA GGCCTGGCCC TCCTTGTAGC TACCACCCGT ACACATGGCA CAGACCACCT
AGCCGGAGCG GGCTTGGCCC TCCTTGTTGC TGCCACGCGT ACACACGGCA CAGATCACCT
GGCCGGAGCA GGCTTGGCCC TCCTTGTAGC CGCCACGCGC ACACACGGCA CAGACCACCT
GGCCGGAGCA GGCTTGGCCC TCCTTGTAGC CGCCACGCGC ACACACGGCA CAGACCACCT
GGCCGGAGCA GGCTTGGCCC TCCTTGTAGC CGCCACGCGC ACACACGGCA CAGACCACCT
AGCCGGAGCG GGTTTAGCCC TCCTTGTAGC TACCACACGA ACACACGGCA CAGACCACCT
GGCCGGAGCG GGTTTAGCCC TCCTTGTAGC TACCACACGA ACACACGGCA CAGACCACCT
GGCCGGAGCG GGTTTAGCCC TCCTTGTAGC TACCACACGA ACACACGGCA CAGACCACCT
AGCCGGAGCG GGCTTGGCCC TCCTTGTAGC TACTACGCGT ACGCACGGGA CAGACCACCT
AGCCGGAGCG GGCTTGGCCC TCCTTGTAGC TACTACGCGT ACGCACGGCA CAGACCACCT
AGCCGGGGCA GGCTTGGCCC TCCTTGTAGC TACTACACGC ACACACGGCA CAGACCACCT
AGCCGGGGCA GGCTTGGCCC TCCTTGTAGC TACTACACGC ACACACGGCA CAGACCACCT
GGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACTACACGT ACACACGGCA CAGACCACCT
GGCCGGAGCG GGCTTGGCCC TCCTTGTAGC TACTACACGT ACACACGGCA CAGACCACCT
GGCCGGAGCG GGCTTGGCCC TCCTTGTAGC TACTACACGT ACACACGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TGCCACGCGT ACACACGGCA CAGACCACCT
AGCTGGGGCA GGCTTGGCCC TCCTTGTAGC TACCACACGT ACACACGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACCACGCGT ACACACGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACCACACGT ACACATGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACCACGCGT ACACATGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACCACGCGT ACACATGGCA CAGACCACCT
AGCCGGAGCA GGCTTGGCCC TCCTTGTAGC TACCACGCGA ACACATGGCA CAGACCACCT
AGCCGGAGCA GGCCTGGCCC TCCTCGTTGC CGCCACGCGA ACGCACGGCA CAGATCACCT
AGCTGGAGCA GGCCTGGCCC TCCTTGTCGC CACCACACGC ACACACGGCA CAGATCACCT

TAAAAACCTT AATCTCCTAC AATGCTAAAA CTACTAATCC CAACTATTAT ACTGCTTCCA
ACAAAACCTT AACCTCTTAC AATGCTAAAA TTATTATTCC CAACTATTAT ATTATTCCCA
CCAAAATCTA AATCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA GTTTTAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA GTTTTAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA GTTTTAATCC CAACACTAAT GCTATTTCCA
CCAAAACCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACGCTAAT GCTATTCCCA
CCAAAACCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTAAT GCTATTCCCA
CCAAAACCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATTC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATTC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTGTTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTGTTCCCA
CCAAAATCTA AACCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTGAT GCTATTCCCA
CCAAAATCTA AATCTCCTAC AATGCTAAAA ATTTTAATCC CAACGCTAAT GCTATTCCCA
TCAAAATCTA AATCTCCTAC AATGCTAAAA ATTTTAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AATCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AATCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTCCCA
CCAAAATCTA AATCTCCTAC AATGCTAAAA ATTATAATCC CAACACTAAT GCTATTTCCA
CCAAAATCTA AATCTCTTAC AATGCTAAAA ATTATAATCC CAACAATAAT GCTATTTCCA
CCAAAACCTA AACCTACTAC AATGCTAAAA ATCTTAATCC CGACATTAAT GCTATTCCCC
CCAAAACCTA AACCTGCTAC AATGCTAAAA ATTTTAATCC CGACACTAAT GCTGTTTCCC

ACGATTTGAA TAACCTCCCC TAAATGATTA TGACCCACCA CTACAACACA CAGTTTAATC
ATAATCTGAA TACTAAATTT CAAATGATTA TGACCAATCA CAACTGCTCA CAGCCTAATC
ACGACCTGAC TTGTAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAGCCACAA CAACAGCCCA AGCCTTAACC
ACGACTTGAC TTGTAGCCCC AAAATGACTC TGAACCACAA CAACGGCCCA AGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTC TGAACCACAA CAACGGCCCA AGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTC TGAACCACAA CAACGGCCCA AGCCCTAGTC
```

```
ACAACTTGAC TTGTTACCCC AAAATGATTA TGAACCACAA CAACGGCCCA GGCCCTAATT
ACGACCTGAC TTGTTACCCC AAAATGATTA TGAACCACAA CAACGGCCCA AGCCCTAATT
ACGACTTGAC TTGTTACCCC AAAATGATTA TGAACCACAA CAACGGCCCA AGCCCTAATT
ACGACCTGAC TTGTGACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGCT
ACGACCTGAC TTGTGACCCC AAAGTGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA GGCCCTAGTC
ACGACTTGAC TTGTAACCCC AAAATGACTT TGAACCACGA CAACAGCCCA GGCCCTAGTT
ACGACTTGAC TTGTAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA GGCCCTAGTT
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCTTGACC
ACGACCTGAC TTGTGGCCCC AAAATGACTT TGAACTACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGTAACCCC AAAGTGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTAGTC
ACGACCTGAC TTGCAACCCC AAAATGACTT TGAACCACAA CAACAGCCCA AGCCCTGGTC
ACAACATGAC TTGCAACACC AAAATGACTA TGAGCCACAA CAACAGCCCA AGCCTTAATC
ACAACATGAA TCACAACCCC AAAATGACTG TGAACCACAA TAACAGCCCA AGCCTTAGCC

ATCGCTACCC TATCTCTTAC GTTATTTAAG CACCACTCTA CAACCCCTTG ATCAAATCTT
ATTGCCACCT TATCACTCTC ACTATTTAAA TGCTATTCCA CAACCCAATG ATCAAACCTT
ATTGCCACCG CAAGCCTGAC TCTACTTAAC TNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
ATTGCCACCG CAAGCCTAAC CCTACTTAAC TGAAACTCAG AAACCGGCTG AACCTCCTCA
ATTGCCACCA CAAGTCTGAC TCTACTTAAC TGAAACTCAG AAACCGGCTG AACCTCCTTA
ATTGCCACCA CAAGTCTGAC TCTACTTAAC TGAAACTCAG AAACCGGCTG AACCTCCTTA
ATTGCCACCA CAAGTCTGAC TCTACTTAAC TGAAACTCAG AAACCGGCTG AACCTCCTTA
ATTGCCACCG CAAGCCTGAC CCTACTTAAC TGAAACTCAG AGACCGGTTG AACCTCCTCA
ATTGCCACCA CAAGCCTGAC TTTACTTAAC TGAAACTCAG AGACCGGTTG AACCTCCTCA
ATTGCCACCA CAAGCCTGAC TTTACTTAAC TGAAACTCAG AGACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTAAC CCTACTTAAC TGAGACTCAG AAACCGGTTG AGCTTCCACC
ATTGCCACCG CAAGCCTAAC CCTACTTAAC TGAGACTCAG AAACCGGTTG AGCCTCATCA
ATTGCCACCG CAAGCCTGAC CTTACTTAAC TGAAATTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC CTTACTTAAC TGAAATTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TTTACTTAAC TGAAATTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TTTACTTAAC TGAAATTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TTTACTTAAC TGAAATTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC CCTACTTAAC TGAAACTCAG AAACCGGCTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TCTACTTAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC CCTACTTAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TCTACTCAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TCTACTTAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TCTACTTAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTGAC TCTACTTAAC TGAAACTCAG AAACCGGTTG AACCTCCTCA
ATTGCCACCG CAAGCCTAAC CCTACTCAAC TGGAGCTCAG AAACTGGCTG AACTTCCACT
ATTGCTACCG CAAGCCTAAC CCTGCTTAAC TGAGACTCGG AAACTGGTTG AACTTCCACC

AACTTCATAA TAGCCACTGA CACAATCTCC ACACCATTAA TCGTCCTTAC CTGCTGACTA
AACTCTTTAA TAGCCACTGA CATGATTTCT ACCCCTCTAA TTGTCTTAAC CTGCTGATTA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AATCCCTACC TGGGCGCAGA CCCACTCTCC ACACCCCTGC TCGTCTTAAC ATGCTGACTT
AACCCATACC TGGGCACAGA CCCGCTTTCC ACGCCCCTGC TTGTCCTAAC ATGCTGACTT
AACCCATACC TGGGCACAGA CCCGCTTTCC ACGCCCCTGC TTGTCCTAAC ATGCTGACTT
AACCCATACC TGGGCACAGA CCCGCTTTCC ACGCCCCTGC TTGTCCTAAC ATGCTGACTT
AACCCCTACC TAGGCACGGA CCCACTCTCC ACCCCCCTGC TCGTCTTAAC ATGCTGACTC
AACCCCTATC TAGGCACGGA CCCACTCTCC ACCCCCCTGC TCGTCTTAAC ATGCTGACTC
AACCCCTATC TAGGCACGGA CCCACTCTCC ACCCCCCTGC TCGTCTTAAC ATGCTGACTC
AACCCCTACC TAGGCACAGA CCCCCTCTCC ACCCCCCTGC TCGTCTTAAC ATGCTGACTT
AACCCCTACC TAGGCACGGA CCCCCTCTCC ACCCCCCTGC TCGTCTTAAC ATGCTGACTT
AACCCCTACC TAGGCACGGA CCCCCTCTCC ACCCCCCTGC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TGGGCACGGA CCCCCTCTCC ACCCCCCTGC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TAGGCACAGA CCCCCTATCC ACCCCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TAGGCACAGA CCCCTTATCC ACCCCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TAGGCACAGA CCCCTTATCC ACCCCCCTAC TCGTCTTGAC ATGCTGACTT
AATCCCTACC TGGGCGCAGA CCCACTCTCC ACACCCCTGC TCGTCTTAAC ATGCTGACTT
AACCCCTACC TGGGCACAGA CCCGCTCTCC ACACCCCTAC TCGTCTTAAC GTGCTGACTT
AACCCCTACC TGGGCACAGA CCCGCTTTCC ACACCTCTAC TCGTCTTGAC ATGCTGACTT
```

```
AACCCCTACC TGGGCACAGA CCCGCTCTCC ACACCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TGGGCACAGA CCCGCTCTCC ACACCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TGGGCACAGA CCCGCTCTCC ACACCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TGGGCACAGA CCCGCTCTCC ACACCCCTAC TCGTCTTGAC ATGCTGACTT
AACCCCTACC TAGGCACGGA TCCACTCTCC ACCCCCCTAC TAGTCCTAAC ATGCTGACTA
AACCCCTACC TGGGCACAGA TCCGCTCTCC ACCCCTCTGC TCGTCCTAAC ATGCTGACTA

CTCCCACTTA TGATCCTCGC TAGCCAAAAC CACATCTCAT CAGAACCTAT TAACCGCCAA
CTCCCACTTA TAATTCTTGC CAGCCAAAAC CATATCTCCA CAGAACCAAT TAACCGCCAA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CTCCCCCTAA TAATTTTAGC AAGCCAAAAC CACATCTCCC CAGAGCCTAT TAGCCGCCAA
CTCCCCTTAA TAATTTTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CGGCCGCCAA
CTCCCCTTAA TAATTTTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CGGCCGCCAA
CTCCCCTTAA TAATTTTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CGGCCGCCAA
CTCCCCCTAA TAATCCTAGC AAGCCAAAAT CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCCTAA TGATCCTAGC AAGCCAAAAT CACATCTCCT CAGAACCAAT CAGCCGCCAA
CTCCCCCTAA TGATCCTAGC AAGCCAAAAT CACATCTCCT CAGAACCAAT CAGCCGTCAA
CTCCCCCTAA TAATCCTAGC AAGCCAAAAC CACATCTCCC CAGAACCCAT CGGCCGCCAA
CTCCCCCTAA TGATCCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CGGCCGCCAA
CTCCCCTTAA TGATCCTAGC AAGCCAAAAC CACATCTCCC CGGAACCAAT CAGCCGCCAA
CTCCCCTTAA TGATCCTAGC AAGCCAAAAC CACATCTCCC CGGAACCAAT CAGCCGCCAA
CTCCCCTTAA TGATCCTAGC AAGCCAAAAC CACATCTCCC CGGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAGCCTAT TAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTGGC AAGTCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAGCCGCCAA
CTCCCCTTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCAAT CAACCGCCAA
CTACCCCTAA TGATTCTAGC AAGCCAAAAC CACATTTCTC CAGAGCCCGT TAGTCGTCAA
CTACCCCTAA TAATTCTAGC AAGCCAAAAC CACATCTCCC CAGAACCCAT TAATCGCCAA

CGAAGTTATA TTATACTATT AGTATCATTA CAAACATTAC TGATTATAGC CTTTAGCGCC
CGAGCCTACA TCACACTACT AGTATCTCTA CAAACCCTAC TTATTATAGC CTTCAGCGCC
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CGAGCCTACA TCACCCTACT AGCATCCCTC CAACTATTCT TAATTATGGC CTTCGGGGCC
CGAACCTACA TCACCCTTCT AGTGTCCCTC CAACTATTCC TAATCATGGC CTTCGGAGCC
CGAACCTACA TCACCCTTCT AGTGTCCCTC CAACTATTCC TAATCATGGC CTTCGGAGCC
CGAACCTACA TCACCCTTCT AGTGTCCCTC CAACTATTCC TAATCATGGC CTTCGGAGCC
CGAACCTACA TTACCCTACT AACGTCCCTC CAACTATTCT TAATTATGGC CTTCGGGGCC
CGAATCTACA TTACCCTGCT AACGTCCCTC CAACTATTCT TAATTATGGC CTTCGGGGCC
CGAACCTACA TTACCCTGCT AACGTCCCTC CAACTATTCT TAATTATGGC CTTCGGGGCC
CGAACCTACA TTACCCTTCT AGTGTCCCTC CAACTATTTT TAATTATAGC CTTCGGAGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAACTATTTT TAATTATAGC CTTCGGAGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAGCTATTTT TAATTATAGC CTTCGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAGCTATTTT TAATTATAGC CTTCGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAGCTATTTT TAATTATAGC CTTCGGAGCC
CGAACCTACA TTACCCTACT AGTGTCCCTC CAACTATTTT TAATTATAGC CTTCGGAGCC
CGAACCTACA TTACCCTACT AGTGTCCCTC CAACTATTTT TAATTATAGC CTTCGGAGCC
CGAACCTACA TCACCCTACT AGTATCCCTC CAACTATTCC TAATTATAGC CTTTGGGGCC
CGAACCTACA TTACCCTACT AGTGTCCCTC CAACTATTCC TAATTATAGC CTTCGGGGCC
CGAACCTATA TTACCCTATT AGTATCCCTC CAACTATTCT TAATTATAGC CTTTGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAACTATTCC TAATTATAGC CTTTGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAACTATTCC TAATTATAGC CTTTGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAACTATTCC TAATTATAGC CTTTGGGGCC
CGAACCTACA TCACCCTACT AGTGTCCCTC CAACTATTCC TAATTATAGC CTTTGGGGCC
CGAACCTACA TCACCCTTCT AGTGTCTCTC CAACTATTCT TAATTATAGC CTTCGGGGCC
CGAACCTACA TTACCCTTCT AGTGTCCCTC CAGCTGTTCT TAATCATAGC CTTCGGGGCC

ACAGAAATTA TTCTATTTTA TATTATATTT GAAGCCACCT TAATCCCCAC CTTAATTATT
ACTGAAATTA TTTTATTTTA CATTATATTT GAAGCCACCC TAATCCCCAC CCTTATTATT
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAAGCTACAC TAATCCCAAC TCTAATAATT
ACCGAAATTA TCCTATTTTA TATCATATTT GAAGCCACGC TAATTCCAAC CCTAATTATT
ACCGAAATTA TCCTATTTTA TATCATATTT GAAGCCACGC TAATTCCAAC CCTAATTATT
```

ACCGAAATTA TCCTATTTTA TATCATATTT GAAGCCACGC TAATTCCAAC CCTAATTATT
ACTGAAATTA TCCTATTCTA CATCATGTTT GAAGCTACAC TAATCCCAAC TCTAATTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAGGCTACAC TAATCCCAAC TCTAATTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAGGCTACAC TAATCCCAAC TCTAATTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATTA TCCTGTTCTA TATCATGTTT GAGGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATTA TCCTGTTCTA TATCATGTTT GAGGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATTA TCCTATTCTA CATTATGTTC GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAGATTA TCCTGTTCTA CATTATGTTC GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATTA TCCTATTCTA CATCATGTTT GAGGCTACAC TAATCCCAAC TCTAATTATT
ACCGAAATCA TCCTGTTTTA TATCATGTTT GAAGCCACAC TAATCCCGAC CCTAATTATT
ACCGAAATTA TCCTGTTCTA CATCATGTTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAGATTA TCCTGTTCTA TATCATGTTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATCA TCCTATTCTA CATCATATTT GAAGCCACAC TAATCCCTAC CCTAATTATT
ACCGAAATCA TCCTATTTTA TATCATATTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATCA TCCTATTTTA TATCATATTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATCA TCCTATTCTA TATCATATTT GAAGCCACAC TAATCCCAAC CCTAATTATT
ACCGAAATCA TCTTATTCTA CATCATATTT GAAACCACAC TAATCCCCAC CCTGATTATT
ACCGAAATTA TTTTATTTTA TATCATATTT GAAGCCACAC TGATCCCCAC CTTAATTATT

ATTACTCGCT GAGGAAACCA AGCAGAACGC CTCAACGCAG GAACCTACTT CCTATTTTAT
ATCACACGAT GAGGAAATCA AGCAGAACGT CTTAACGCAG GAACCTACTT TTTATTCTAT
ATTACACGAT GGGGAAACCA AACCGAGCGA CTCAACGCAG GCACCTATTT TCTGTTTTAT
ATCAAACGAA GAGGGAACCA AACCGAAAGA CTAAACGCAA GAACCTATTT TCTATTCTAC
ATTACGCGAT GAGGAAACCA AACCGAACGA CTTAACGCAG GCACCTACTT TCTATTCTAC
ATTACGCGAT GAGGAAACCA AACCGAACGA CTTAACGCAG GCACCTACTT TCTATTTTAC
ATTACGCGAT GAGGAAACCA AACCGAACGA CTTAACGCAG GCACCTACTT TCTATTTTAC
ATCACACGAT GAGGGAACCA AACCGAACGA CTTAACGCAG GCACCTATTT TCTATTCTAC
ATCACACGAT GAGGGAACCA AACTGAACGA CTTAACGCAG GCACCTATTT CCTATTCTAC
ATCACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT CCTATTCTAC
ATTACGCGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GAACCTATTT TCTATTCTAC
ATTACGCGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GAACCTATTT TCTATTCTAC
ATCACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GAACCTACTT TCTATTCTAC
ATCACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GAACCTACTT TCTATTCTAC
ATTACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GAACCTACTT TCTATTCTAC
ATCACACGAT GGGGAAACCA AACTGAACGA CTTAACGCAG GAACCTATTT TCTATTTTAC
ATCACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT CCTATTCTAC
ATTACACGAT GGGGAAACCA AACTGAACGG CTTAACGCAG GCACCTACTT TCTGTTCTAT
ATTACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTCTAC
ATTACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTCTAC
ATTACACGAT GGGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTTTAT
ATTACACGAT GGGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTTTAT
ATTACACGAT GGGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTTTAT
ATTACACGAT GAGGAAACCA AACTGAACGA CTTAACGCAG GCACCTATTT TCTATTTTAT
ATCACACGAT GAGGAAACCA AACCGAACGT CTCAACGCAG GCACCTACTT CCTATTCTAC
ATCACACGAT GAGGAAACCA AACTGAACGA CTCAACGCAG GTACCTACTT CCTGTTCTAC

ACCTTGGCTG GCTCACTCCC ACTGCTAATT GCCCTCCTCT ATCTATATAA CATCACCGGC
ACACTCGCCG GCTCCCTCCC CCTGCTGATC GCCCTCCTCT ACCTTTATAA CACGGCCGGT
ACCCTAGCCG GATCCCTCCC CCTGCTAGTT GCCCTACTAA TTCTGCAAAA AGAACTAGGC
ACTTTNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNGGC
ACNNNNGCCG GATCGCTCCC TCTACTAGTT GCCCTATTAA TCCTGCAAAA AGAACTAGGC
ACTGCGGCCG GATCGCTCCC TCTACTAGTT GCCCTACTAA TCCTGCAAAA AGAACTAGGC
ACTCTGGCCG GATCGCTCCC TCTACTAGTT GCCCTACTAA TCCTGCAAAA AGAACTAGGC
ACTTTGGCCG GATCACTTCC TCTACTGGTT GCCCTACTGA TTCTGCAAAA AGAGCTAGGC
ACCTTGGCCG GGTCACTTCC TCTACTGGTT GCCCTACTGA TTCTGCAAAA AGAGCTAGGC
ACCTTGGCCG GGTCACTTCC TCTACTGGTT GCCCTACTGA TTCTGCAAAA AGAGCTAGGC
ACCCTGACCG GATCACTTCC CCTGCTAGTT GCCCTATTAG TCCTGCAAAA AGAGCTAGGC
ACCCTGNCCG GATCACTTCC TCTGCTAGTT GCCCTATTAG TCCTGCAAAA AGAGCTAGGC
ACCCTGGCCG GATCACTTCC TCTGCTAGTT GCCCTATTAA TCCTGCAAAA AGAACTAGGT
ACCTTAGCCG GATCACTCCC TCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAACTAGGT
ACCCTGGCCG GATCGCTTCC TCTGCTAGTT GCCCTATTAA TCCTGCAAAA AGAACTAGGC
ACCCTGGCCG GATCGCTTCC TCTGCTAGTT GCCCTATTAA TCCTGCAAAA AGAACTAGGC
ACCTTGGCCG GATCGCTTCC TCTGCTAGTT GCCCTATTAA TCCTGCAAAA AGAACTAGGC
ACCCTAGCCG GGTCCCTCCC TCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAACTAGGC
ACCCTAGCCG GATCCCTCCC TCTGCTAGTT GCCCTATTAA TCCTACAAAA AGAACTAGGC

```
ACCCTAGCCG GATCCCTTCC TCTGCTAGTT GCCCTACTAG TTCTGCAAAA AGAACTAGGC
ACC---GCCG GATCCCTCCC CCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAGCTAGGC
ACCCTAGCCG GATCCCTCCC TCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAGCTAGGC
ACCTTGGCCG GATCCCTCCC TCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAGCTAGGC
ACCCTAGCCG GATCCCTCCC TCTGCTAGTT GCCCTACTAA TCCTGCAAAA AGAACTAGGC
ACACTAGCCG GGTCCCTCCC CCTACTAGTT GCCCTACTGA TCCTACAAAA AGAAATAGGC
ACACTAGCCG GATCCCTCCC CCTACTAGTC GCCCTGCTAA TTCTCCAAAA AGAATTAGGC

TCACTATCAA TCCT-----C TCAATAAACT TAATCTTTAT TCACCCTAAT ACCTGAGCCG
TCGCTATCAA TACT-----C TCAATCAACC TTATTTCTAT TTATCCTGAT AATTGAGCCC
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACTTGAGCCG
ACCCACTCAA TAGTGAGTAT CAAACATAAA CAA--CCCAC CCCCCTATGC GCATAAGCCG
TCCCTTTCAA TACTAATTAT TCAATACATA CAA--CCCAC CCCCCTATGT ACCTGAGCTG
TCCCTTTCAA TACTAATTAT TCAATACATA CAA--CCCAC CCCCCTATGT ACCTGAGCTG
TCCCTTTCAA TACTAATTAT TCAATACATA CAA--CCCAC CCCCCTATGT ACCTGAGCTG
TCCCTTTCAA TATTGGTTAT CCAATATATA CAG--CCCGC CCCTCTATGT GCGTGAGCCG
TCCCTTTCAA TATTGATTAT CCAATATATA CAG--CCCGC CCCTCTATGC ACATGAGCCG
TCCCTTTCAA TATTGATTAT CCAATATATA CAG--CCCGC CCCTCTATGC ACATGAGCCG
TCCCTTTCAA TACTGATTAT TCAATACATA CAA--CCTGC CCCCCTATGC ACCTGAGCCG
TCCCTTTCAA TACTGATTAT TCAATACATA CAA--CCTGC CCCCCTATGC ACCTGAGCCG
TCCCTTTCAA TACTGATCAT TCAATATATA CAA--CCTGC CCCCCTATGC ACCTGAGCCG
TCCCTTTCAA TACTGATCAT TCAATATATA CAA--CCTGC CCCCCTATGC ACCTGAGCCG
TCCCTTTCAA TACTAATTAT TCAATATATA CAA--CCTGC CCCCTTATGT ACCTGAGCCG
TCCCTTTCAA TACTAATTAT TCAATATATT CAA--CCTGC CCCCTTATGC ACCTGAGCCG
TCCCTTTCAA TACTAATTAT TCAATATATT CAA--CCTGC CCCCTTATGC ACCTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACCTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCCCTATGC ACCTGAGCCG
TCCCTTTCAA TATTAGTTAT TCAATATATA CAA--CCCGC CCCACTGTGC ACTTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACTTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACTTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACTTGAGCCG
TCCCTTTCAA TATTGATTAT TCAATATATA CAA--CCCGC CCCGCTATGC ACTTGAGCCG
TCCCTATCAA TACTAATTAT CCAATACATA CAA--CCTGC CCCCCTATGC ACCTGAGCCG
TCGCTCTCAA TACTAATTAT CCAATACATG CAG--CCTGC CCCCCTGTGC ACCTGAGCCG

ACACCTTTCT ATGAGCCGCC TGCTTGATCG CCTTCTTAGT AAAAATACCT CTGTATGGAG
ACACCTTCCT ATGAGCTGCT TGCCTTATCG CTTTCCTAGT AAAAATACCA TTATATGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ATAAAATCTG ATGAGCGGCC TGCTTAATTG CCTTCCTAGT AAAAATACCC CTATACGGAG
ACAAAATCTG GTGGGCGGCC TGCTTAATTG CCTTTCTAGT AAAAATACCC CTATATGGAG
ACAAAATCTG GTGGGCGGCC TGCTTAATTG CCTTTCTAGT AAAAATACCC CTATATGGAG
ACAAAATCTG GTGGGCGGCC TGCTTAATTG CCTTTCTAGT AAAAATACCC CTATATGGAG
ATAAAATCTG ATGAGCGGCC TGCTTAATTG CCTTCCTAGT AAAGATACCC CTGTACGGAG
ATAAAATCTG ATGAGCGGCC TGCTTAATTG CCTTCCTAGT AAAAATACCC CTATACGGAG
ATAAAATCTG ATGAGCGGCC TGCTTAATTG CCTTCCTAGT AAAAATACCC CTATACGGAG
ACAAAATCTG ATGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATACGGAG
ACAAAATCTG ATGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATATGGAG
ACAAAATCTG ATGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATACGGGG
ACAAAATCTG ATGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATACGGGG
ACAAGATCTG GTGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATACGGGG
ACAAAATCTG GTGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATATGGGG
ACAAAATCTG GTGGGCAGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATATGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAATTGCCC CTATACGGGG
ACAAAATCTG ATGAGCGGCC TGCTTAATCG CCTTCCTAGT AAAAATACCC CTATACGGGG
ACAAAGTCTG ATGAGCGGCC TGCTTAGTCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ACAAAATCTG ATGGGCGGCC TGCTTAATCG CCTTCCTAGT GAAAATACCC CTATACGGGG
ATAAAATCTG ATGAGCAGCC TGCTTAATTG CCTTCCTAGT AAAAATACCC CTCTACGGAG
ACAAAATCTG ATGGACAGCC TGCCTAGTTG CTTTCCTGGT AAAAATACCC CTCTACGGAG

TACATTTATG ACTCCCCAAA GCCCATGTAG AAGCCCCAAT TGCCGGATCA ATAATTCTTG
TTCACCTATG ACTCCCAAAA GCTCATGTAG AAGCTCCAAT CGCCGGATCA ATGATCCTTG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCACACGTAG AAGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCACACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
```

```
CCCACCTCTG ACTCCCAAAA GCACACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCACACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG GCTCCCAAAA GCACACGTAG AGGCCCCAAT TGCAGGATCT ATGGTCCTAG
CTCACCTCTG ACTTCCAAAA GCACACGTAG AAGCCCCAAT TGCAGGATCT ATAGTCCTGG
CTCACCTCTG ACTTCCAAAA GCACACGTAG AAGCCCCAAT TGCAGGATCT ATAGTCCTGG
CCCACCTCTG ACTTCCAAAA GCACACGTAG AAGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTTCCAAAA GCACACGTAG AAGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTTCCAAAA GCACACGTAG AGGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTTCCAAAA GCACACGTAG AGGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCCAAA GCACACGTAG AAGCCCCAGT TGCAGGATCT ATAGTTCTGG
CCCACCTCTG ACTCCCCAAA GCACACGTAG AAGCCCCAGT TGCCGGATCT ATAGTTCTGG
CCCACCTCTG ACTCCCCAAA GCACACGTAG AAGCCCCAGT TGCCGGATCT ATAGTTCTGG
CCCACCTCTG ACTCCCAAAA GCCCACGTAA AGGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCCAAA GCACACGTAG AGGCCCCAGT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
CCCACCTCTG ACTCCCAAAA GCGCACGTAG AGGCCCCAAT TGCAGGATCC ATAGTCCTGG
TCCACCTCTG ATTACCAAAA GCACACGTAG AAGCCCCAGT TGCAGGATCT ATAGTCCTAG
TCCACCTCTG ACTACCAAAA GCACACGTAG AAGCCCCCAT TGCAGGATCC ATAGTTTTAG
```

```
CCGCAGTACT CCTAAAATTA GGCGGGTATG GAATAATTCG AATAACCATT ATTCTTGAAC
CTGCAGTACT TCTGAAACTA GGCGGATATG GAATAATTCG AATAACTATT ATATTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGTGGCTACG GCATAATACG AATGATCGTT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGTGGCTACG GCATAATACG AATGATCGTT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGTGGCTACG GCATAATACG AATGATCGTT ATGCTAGAAC
CTGCCGTACT ACTAAAACTC GGCGGCTACG GTATAATACG AATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTC GGCGGCTACG GCATGATACG GATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTC GGCGGCTACG GCATGATACG GATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGCGGATACG GCATAATACG AATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGCGGATACG GCATAATACG AATAATTATT ATGCTAGAGC
CTGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATCGTT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGTGGCTACG GCATGATACG AATAATTGTT ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGTGGCTACG GCATGATACG AATAATTGTC ATGCTAGAAC
CTGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTGTT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATAATTATT ATGCTAGAAC
CCGCCGTACT ACTAAAACTT GGCGGCTACG GCATAATACG AATGATTATT ATGCTAGAAC
CTGCCGTACT ACTAAAACTA GGCGGTTATG GCATGATACG AATGATTATT ATACTAGAGC
CCGCCGTACT ACTAAAACTA GGCGGCTACG GCATAATACG AATGATTGTC ATGCTAGAGC
```

```
CCGCAACAAA ATCACTAGCC TACCCATTCA TTATTTTAGC ATTATGGGGA ATCATTATAA
CGGCAACCAA ATCACTAGCC TACCCATTTA TTATTCTCGC ACTTTGAGGG ATCATTATAA
CGACATCCAA AAATCTCGCG TACCCATTTA TTATTCTAGC TTTATGAGGC ATTATTATGA
CAACATCCAA AAATCTCGCG TACCCATTTA TTATCCTAGC TTTATGAGGC ATTATCATGA
CAACATCCAA AAATTTCGCG TACCCATTTA TCATCCTAGC CCTATGGGGC ATTATTATGA
CAACATCCAA AAATTTCGCG TACCCATTTA TCATCCTAGC CCTATGGGGC ATTATTATGA
CAACATCCAA AAATTTCGCG TACCCATTTA TCATCCTAGC CCTATGGGGC ATTATTATGA
CAGCATCTAA AAATTTAGCA TACCCATTTA TCATCCTGGC CCTATGGGGC ATCATCATAA
CAACATCTAA AGATTTAGCA TACCCATTTA TCATCCTGGC CCTATGGGGC ATCATCATAA
CAACATCTAA AGATTTAGCA TACCCATTTA TCATCCTAGC CCTATGGGGC ATCATCATAA
CATCATCCAA AAATCTCACG TACCCATTTA TTATCCTAGC TTTATGGGGT ATTATTATGA
CATCATCCAA AAATCTCACG TACCCATTTA TTATCCTAGC TTTATGGGGT ATTATTATGA
CAGCATCCAA AAATCTCGCG TACCCATTTA TCATCCTGGC TTTATGGGGT ATTATCATGA
CGGCATCCAA AAATCTCGCG TACCCATTTA TCATCCTGGC TTTATGGGGT ATTATCATGA
CAGCATCCAA AAATCTCACG TACCCATTTA TTATCCTAGC TTTATGGGGT ATCATCATGA
CAGCATCCAA AAATCTCACG TACCCATTTA TTATCCTGGC TTTATGGGGT ATCATCATGA
CAGCATCCAA AAATCTCACG TACCCATTTA TTATCCTGGC TTTATGGGGT ATCATCATGA
CGGCATCCAA AAATCTCGCG TACCCATTTA TCATCCTGGC TTTATGGGGT ATTATCATGA
```

CGACATCCAA AAATCTCGCG TACCCATTTA TTATCCTAGC TTTATGAGGC GTTATCATGA
CAGCATCCAA AAATCTCGCG TACCCATTTA TTATCCTAGC TTTATGGGGT ATTATCATGA
CAACATCCAA AAATCTCGCG TACCCATTTA TTATTCTAGC TTTATGGGGC ATTATCATGA
CAACATCCAA AAATCTCGCG TACCCATTTA TTATTCTAGC TTTATGAGGC ATTATCATGA
CAACATCCAA AAATCTCGCG TACCCATTTA TTATTCTAGC TTTATGAGGC ATTATCATGA
CAACATCCAA AAATCTCGCG TACCCATTTA TTATTCTAGC TTTATGGGGC ATTATCATGA
CTGCATCCAA AAACCTTGCA TACCCATTTA TTATCCTGGC CCTATGAGGC ATTATCATGA
CGACATCCAA GAACCTTGCA TACCCGTTCA TTATCTTAGC TCTATGAGGT ATTATCATAA

CAGGATCCAT TTGTATACGA CAATCTGATA TAAAATCTTT AATCGCCTAC TCATCCGTCA
CAGGATCCAT TTGCATGCGA CAATCTGACA TAAAATCCTT AATCGCCTAC TCATCCGTCA
CCGGGTCAAT TTGTCTGCGT CAAACGGACC TAAAATCCCT AATTGCATAT TCATCAGTAA
CCGGGTCAAT TTGTCTGCGA CAGACAGACC TAAAATCCCT AATTGCATAT TCATCAGTAA
CAGGATCGAT CTGTTTACGA CAGACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CAGGATCGAT CTGTTTACGA CAGACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CAGGATCGAT CTGTTTACGA CAGACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CCGGGTCAAT TTGCCTCCGA CAAACAGACC TAAAATCCCT AATTGCATAT TCGTCAGTAA
CCGGGTCAAT TTGCCTCCGA CAAACAGACC TAAAATCCCT AATTGCATAT TCATCAGTGA
CCGGGTCAAT TTGCCTCCGA CAAACAGACC TAAAATCCCT AATTGCATAT TCATCAGTGA
CCGGGTCAAT TTGTCTACGA CAGACAGACC TAAAATCCCT AATCGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAGACAGACC TAAAATCCCT AATCGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAGACAGACC TAAAATCCCT AATCGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAGACAGACC TAAAATCCCT AATCGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAAACAGACC TAAAATCCCT AATCGCATAT TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAAACAGACC TAAAATCCCT AATCGCATAT TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAAACAGACC TAAAATCCCT AATCGCATAT TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAGACAGACC TAAAATCCCT AATCGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTGCGA CAGACAGACC TAAAATCCCT AATTGCATAT TCATCAGTAA
CCGGGTCAAT TTGTCTACGA CAAACAGACC TAAAGTCCCT AATTGCATAT TCATCCGTAA
CCGGGTCAAT TTGTCTGCGA CAAACAGACC TAAAGTCCCT AATTGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTGCGA CAAACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTGCGA CAAACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CCGGGTCAAT TTGTCTGCGA CAAACAGACC TAAAATCCCT AATTGCATAC TCATCAGTAA
CCGGATCAAT TTGTCTACGA CAGACAGATC TAAAATCCCT AATTGCTTAC TCTTCAGTAA
CTGGGTCGAT TTGCTTACGA CAAACAGACC TAAAATCTCT AATTGCCTAC TCATCAGTAA

GCCACATGGG ACTGGTTGCC GCCGGCATTC TTATTCAAAC CCCCTGAGGC TTTACAGGAG
GCCACATAGG ATTAGTTGCC TCTGGCATTT TAATTCAAAC CCCATGAGGC TTCACAGGAG
GTCATATAGG ACTGGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTTACCGGAG
GCCACATAGG ACTGGTAGTA GCAGGGATCC TCATCCAAAC CCCCTGAGAC TTAACCGGAG
GCCACATAGG ACTAGTGGTA GCAGGGATCC TCATCCAGAC CCCATGAGGC TTCACCGGAG
GCCACATAGG ACTAGTGGTA GCAGGGATCC TCATCCAGAC CCCATGAGGC TTCACCGGAG
GCCACATAGG ACTAGTGGTA GCAGGGATCC TCATCCAGAC CCCATGAGGC TTCACCGGAG
GCCATATAGG GTTGGTAGTA GCAGGGATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GTCACATAGG GCTAGTAGTA GCAGGAATCC TTATCCAAAC CCCCTGAGGC TTTACCGGGG
GTCACATAGG GCTAGTAGTA GCAGGAATCC TTATCCAAAC CCCCTGAGGC TTTACCGGGG
GCCACATGGG ATTGGTGGTA GCAGGAATCC TCATCCAAAC CCCATGAGGC TTCACCGGGG
GCCACATGGG GTTAGTGGTA GCAGGAATCC TCATCCAAAC CCCATGAGGC TTCACCGGGG
GCCACATGGG GTTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GCCACATGGG ATTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GCCACATAGG GTTAGTAGTA GCAGGAATCC TCATCCAAAC CCCATGAGGC TTCACCGGGG
GCCACATGGG GTTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GCCACATGGG GTTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GCCACATGGG ATTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGGG
GCCACATAGG ACTGGTAGTA GCAGGGATCC TCATCCAAAC CCCCTGAGGC TTTACCGGGG
GCCACATAGG ACTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTTACCGGAG
GCCACATAGG ACTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGGGGC TTTACCGGAG
GCCACATAGG ACTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTTACCGGAG
GCCACATAGG ACTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTTACCGGAG
GCCACATAGG ACTAGTAGTA GCAGGAATCC TCATCCAAAC CCCCTGAGGC TTTACCGGAG
GCCACATGGG ACTAGTAGTA GCGGGAATCC TCATCCAAAC CCCCTGAGGC TTCACCGGAG
GCCACATAGG ACTAGTAGCA GCAGGAATCC TCATCCAAAC TCCTTGGGGC TTCACTGGAG

CGATTATCTT AATAATTGCC CACGGACTAA CCTCCTCAGC TCTATTTTGC CTTGCCAATA
CTATCATTTT AATAATTGCC CACGGACTAA CCTCATCAGC CCTATTTTGC CTTGCTAATA
CCATTATTCT GATGATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGC TTAGCAAACA
ACATTATCCT AATAATCCCA AACGCCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA

CCATTATTCT GATAATCGCA CACGGCCTGG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTGG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTGG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT AATGATCGCA CACGGCTTAG CATCCTCCGC ATTATTCTGC TTAGCAAACA
CCATTATTCT AATGATCGCA CACGGCTTAG CATCCTCCGC ATTATTCTGC TTAGCAAACA
CCATTATTCT AATGATCGCA CACGGCTTAG CATCCTCCGC ATTATTCTGC TTAGCAAACA
CTATTATCCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGC CTGGCAAACA
CTATTATCCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGC CTGGCAAACA
CTATTATTCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGT TTAGCAAACA
CTATTATTCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT AATAATCGCA CACGGCCTAG CATCATCCGC ACTATTCTGT TTAGCAAACA
CCATTATTCT AATAATCGCA CACGGCCTAG CATCATCCGC ACTATTCTGT TTAGCAAACA
CTATTATTCT GATAATCGCA CACGGCCTAG CATCATCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTGGCAAACA
CCATCATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CCATTATTCT GATAATCGCA CACGGCCTAG CATCCTCCGC ATTATTCTGT TTAGCAAACA
CTATCATTTT AATGATCGCA CACGGATTAG CATCTTCCGC ATTATTCTGC CTAGCAAACA
CCATCATTCT GATAATCGCA CATGGACTAG CATCCTCCGC ATTATTCTGT CTAGCAAACA

CTTCCTATGA ACGAACGCAC TCTCGAACTC TCCTACTAGC CCGAGGAATG CAAATTATTA
CCACATACGA GCGAACCCAC TCCCGAACAC TCCTATTAGC ACGAGGTATA CAAGTTGCCA
CTAACTACGA ACGCCTTCAT AGCCGAACCC TGCTTCTTGC ACGGGGAATA CAAGCCGTAC
CTAATTATGA ACGCCTTCAC AGCCGAACCC TATTGCTTGC CCGAGGAATA CAAGCCGTCC
CTAACTATGA ACGCCTTCAC AGCCGGACCC TACTTCTTGC GCGAGGAATA CAAACCATCC
CTAACTATGA ACGCCTTCAC AGCCGGACCC TACTTCTTGC GCGAGGAATA CAAACCATCC
CTAACTATGA ACGCCTTCAC AGCCGGACCC TACTTCTTGC GCGAGGAATA CAAACCATCC
CTAATTATGA ACGCCTTCAC AGCCGAACCC TATTGCTTGC CCGAGGAATA CAAGCCGTCC
CTAATTATGA GCGCCTTCAC AGCCGAACCC TATTGCTTGC CCGAGGAATG CAAGCCGTCC
CTAATTATGA GCGCCTTCAC AGCCGAACCC TATTGCTTGC CCGAGGAATG CAAGCCGTCC
CTAACTATGA ACGCCTCCAC AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCATCC
CTAACTATGA ACGCCTCCAC AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCATCC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCGTCC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCGTCC
CTAACTATGA ACGACTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCATTC
CTAACTATGA ACGACTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCATTC
CTAACTATGA ACGACTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCATTC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TACTTCTTGC ACGAGGAATA CAAGCCGTCC
CCAACTATGA ACGCCTTCAT AGCCGAACCC TGCTCCTTGC ACGAGGGATA CAAGCCGTAC
CTAACTATGA ACGCCTTCAT AGCCGAACTC TGCTTCTTGC ACGAGGAATA CAAGCCGTAC
CTAACTATGA ACGCCTTCAT AGCCGAACTC TGCTTCTTGC ACGAGGAATA CAAGCCGTAC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TGCTTCTTGC ACGAGGAATA CAAGCCGTAC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TGCTTCTTGC ACGAGGAATA CAAGCCGTAC
CTAACTATGA ACGCCTTCAT AGCCGAACCC TGCTTCTTGC ACGAGGAATA CAAGCCGTAC
CCAATTATGA GCGCCTCCAC AGTCGAACCC TGCTTCTCGC ACGAGGAATA CAAGCCATCC
CTAACTATGA ACGCCTCCAT AGCCGAACCC TTCTCCTCGC ACGAGGGATA CAAGCCATCC

TACCGCTCAT AACAACCTGA TGACTTATCA TTAGCCTATC AAATATAGCT CTCCCCCCTC
TACCTCTCAT AGCAACCTGA TGATTCATTA TAAGCCTAGC AAACATAGCC CTACCGCCTC
TCCCCCTAAT AGCCACATGA TGATTTATCG CCAACCTAGC CAACCTGGCC CTCCCTCCTC
TCCCCCTAAT GGCCACATGA TGATTTATTG CCAACCTAGC TAACCTAGCC CTCCCCCCTC
TCCCCCTAAT AGCCACATGA TGATTTATTG CTAATCTAGC CAACCTGGCC CTCCCTCCCC
TCCCCCTAAT AGCCACATGA TGATTTATTG CTAATCTAGC CAACCTGGCC CTCCCTCCCC
TCCCCCTAAT AGCCACATGA TGATTTATTG CTAATCTAGC CAACCTGGCC CTCCCTCCCC
TCCCCCTAAT GGCCACATGA TGATTTATTG CCAACCTAGC TAACCTAGCC CTCCCCCCTC
TCCCCCTAAT GGCCACATGA TGATTTATTG CCAACCTAGC TAACCTAGCC CTCCCCCCTC
TCCCCCTAAT GGCCACATGA TGATTTATTG CCAACCTAGC TAACCTAGCC CTCCCCCCTC
TCCCCCTAAT AGCTACATGA TGATTCATCG CCAACCTGGC CAACCTGGCC CTCCCTCCCC
TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTGGC CAACCTGGCC CTCCCTCCCC
TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTAGC CAACCTGGCC CTCCCCCCTC
TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTAGC CAACCTGGCC CTCCCCCCTC
TCCCCCTAAT AGCCACATGA TGGTTCATCG CTAACCTAGC CAACCTGGCC CTCCCTCCTC
TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTAGC CAACCTGGCC CTCCCCCCTC
TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTAGC CAACCTGGCC CTCCCCCCTC

TCCCCCTAAT AGCCACATGA TGATTCATCG CCAACCTAGC CAACCTGGCC CTCCCCCCTC
TCCCCCTAAT AGCCACATGA TGATTTATCG CCAACTTAGC CAACCTGGCC CTCCCTCCTC
TCCCCTTAAT AGCCACATGA TGATTTATTG CCAACCTAGC CAACCTAGCC CTCCCTCCCC
TACCCCTAAT AGCCACATGA TGATTTATCG CCAACCTAGC CAACCTGGCC CTCCCTCCCC
TACCCCTAAT AGCCACATGA TGATTTATCG CCAACCTAGC CAACCTGGCC CTCCCTCCTC
TACCCCTAAT AGCCACATGA TGATTTATCG CCAACCTAGC CAACCTGGCC CTCCCTCCTC
TACCCCTAAT AGCCACATGA TGATTTATCG CCAACCTAGC CAACCTGGCC CTCCCTCCTC
TCCCCCTGAT AGCTGCATGA TGATTCATCG CCAACCTGGC CAACTTAGCC CTCCCTCCCC
TCCCCCTAAT AGCCGCATGA TGATTTATTG CTAACCTGGC CAACCTAGCC CTCCCTCCCC

TTCCTAACCT AATGGGAGAA CTTATAATTA TAGTATCTAT ATTTAACTGA TCTATCTGAA
TCCCCAACCT AATAGGCGAA CTAATAATTA TAATTTCGAT ATTTAACTGA TCTAATTGAA
TCCCCAACCT AATAGGAGAA CTAGTTATTA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATCA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATGGGAGAG CTAGTTATCA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCCAACCT AATGGGAGAG CTAGTTATCA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCCAACCT AATGGGAGAG CTAGTTATCA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATCA TCACCTCAAT ATTTAACTGA TCAAGCTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATCA TTACCTCAAT ATTTAACTGA TCAAGCTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATCA TTACCTCAAT ATTTAACTGA TCAAGCTGAA
TCCCCAACCT AATAGGAGAA CTAGTTATTA TTTCCTCAAT ATTCAACTGA TCAAGCTGGA
TCCCCAACCT AATAGGAGAA CTAGTTATTA TTTCCTCAAT ATTCAACTGA TCAAGCTGGA
TCCCTAACCT AATAGGAGAA CTAGTTATTA TTTCCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCTAACCT AATAGGAGAA CTAGTTATTA TTTCCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCTAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCTAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCTAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCTAACCT AATAGGAGAA CTAGTCATCA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCTAACCT AATAGGAGAA CTAATTATTA TTTCCTCAGT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATAGGAGAA CTAGTTATTA TCACCTCAAT ATTTAACTGA TCAAACTGAA
TCCCCAACCT AATGGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT AATAGGAGAA CTAGTCATTA TTACCTCAAT ATTCAACTGA TCAAACTGAA
TCCCCAACCT TATAGGAGAA CTAATTATCA TCACCTCAAT ATTCAACTGA TCAAGCTGAA
TCCCCAATCT AATGGGAGAG TTAATCATTA TCACCTCAAT ATTCAACTGA TCAAGCTGAA

CTATTCTACT AACCGGCACC GGAACCCTTA TTACAGCTAG CTACTCACTT TACCTCTTTA
CTATTATACT AACCGGTGTC GGAACCCTAA TTACAGCCAG CTATTCACTC TACCTATATA
CAATCATCCT TACAGGGGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATATATC
CAATTATCCT CACAGGAGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATTGTCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTATT
CAATTGTCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTATT
CAATTGTCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTATT
CAATTTTCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATATACC
CAATTATCCT CACAGGAGGG GGGACCCTTA TTACCGCCAG CTACTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGGACCCTTA TTACCGCCAG CTACTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGGACCCTTA TTACCGCCAG TTACTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGGACCCTTA TTACCGCCAG TTATTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGAACCCTGA TTACCGCCAG CTATTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTATTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGA GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTACC
CAGTTATCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATGTACC
CAATTATCCT CACAGGAGGG GGAACTCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATCATTCT CACGGGAGGG GGAACCCTCA TTACCGCCAG CTACTCCCTC TATATGTATC
CAATTATCCT CACAGGAGGG GGAACCCTTA TTACCGCCAG CTACTCCCTC TACATATATC
CAATTATCCT CACAGGAGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATTATCCT CACAGGAGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATTATCCT CACGGGAGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATTATCCT CACAGGAGGA GGAACCCTCA TTACCGCCAG CTACTCCCTC TACATGTATC
CAATTATCCT TACAGGGACC GGAACCCTTA TCACCGCCAG CTATTCTCTC TATATGTATC
CAATTGTCCT TACAGGAACT GGAACCCTCA TTACCGCCAG CTACTCCTTA TACATGTTCC

TTTCCTCCCA GCGAGGACCA GCCTCAAAT- --AACACCCC CATAGAGCCT TCACATACCC
TATCTTCACA ACGAGGACCT ATCCCGAACA ACCTAACCCC TATAGAATCC TCTCATACCC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATTATGGC AGTTGAACCA TCCCACACAC

```
TAATAACACA ACGAGGCCCA GTGTCCACCT TAATTATAGC AGTAGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TTATCATAGC AGTAGAGCCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TTATCATAGC AGTAGAGCCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TTATCATAGC AGTAGAGCCA TCCCACACAC
TAATGACACA ACGAGGCCCA GTGTCCACCC TAATCATGGC AATCGAACCA TCTCACACAC
TAATGACACA ACGAGGCCCA GTGTCTACCT CAATCATGGC AATCGAACCA TCTCACACAC
TAATGACACA ACGAGGCCCA GTGTCTACCT CAATCATGGC AATCGAACCA TCTCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAACCC TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAACCC TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAACCT TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAACCT TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCC TAATCATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAGCCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTATCCACCT TAATCATGGC AGTTGAGCCA TCCCACACAC
TAATAACACA ACGGGGTCCA GTGTCCACCT TAGTTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCG GTATCCACCT TTATCATGGC AGTTGAACCG TCCCACACAC
TAATAACACA ACGAGGCCCA GTGTCCACCT TAATTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTGTCCACCT TGATTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTGTCCACCT TAATTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTGTCCACCT TAATTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCA GTGTCCACCT TAATTATGGC AGTTGAACCA TCCCACACAC
TAATAACACA ACGAGGCCCT GTAACCACCT TAATCACAGC AATTGAACCA TCCCACACAC
TAATGACCCA ACGAGGCCCC ATAACTACCC TAATCACAGC AATTGAGCCA TCCCACACAC

GAGAACATCT CTTATTTACC CTACATATTA TTCCCATTAT ACTTCTAATA ATTAAGCCTG
GAGAACATCT CCTGCTAACC CTACATATCA TCCCAATTAT TCTCCTAATA ATTAAGCCTG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCCATTAT TTTACTGATG CTAAAGCCAG
GAGAACACCT ACTCATAGAA CTACACCTCA TCCCCATTAT TCTACTAATA CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCCATCGC CCTACTAATG CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCCATCGC CCTACTAATG CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCCATCGC CCTACTAATG CTAAAACCAG
GAGAACACCT ACTTATAGCA CTACACCTCA TCCCCATTAT TCTATTGATA CTAAAGCCAG
GAGAACACCT ACTTATAGCA CTACACCTCA TCCCCATTAT TCTATTGATA CTAAAGCCAG
GAGAACACCT ACTTATAGCA CTACACCTCA TCCCCATTAT TCTATTGATA CTAAAGCCAG
GAGAACACCT ACTCATGGCA CTACACCTCA TCCCTATTAT TTTACTGATA CTAAAGCCAG
GAGAACACCT ACTCATGGCA CTACACCTCA TCCCTATTAT TTTACTGATA CTAAAGCCAG
GAGAACACCT ACTCATGGCA CTACACCTCA TCCCCATTAT TTTACTAATA CTAAAGCCAG
GAGAACACCT ACTCATGGCA CTACACCTCA TCCCTATTAT TTTACTAATA CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCTATCAT TTTACTAATA CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCTATCAT TTTACTAATA CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCTATCAT TTTACTAATA CTAAAACCAG
GAGAACACCT ACTCATAGCA CTGCACCTTA TCCCCATTAT TCTACTAATG CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCTATTAT TCTACTGATG CTAAAACCAG
GAGAACACCT ACTCATAGCA CTACACCTTA TCCCCATTAT TCTACTGATG CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCCATTAT TCTACTAATG CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCCATTAT TCTACTGATG CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCCATTAT TCTACTAATG CTAAAGCCAG
GAGAACACCT ACTCATAGCA CTACACCTCA TCCCCATTAT TCTACTGATG CTAAAGCCAG
GAGAACACCT ACTCATGGCG CTACACCTAA TCCCCATCAT CATACTGACA CTAAAACCAG
GAGAACACCT ACTCATGGCA CTACACCTAA TCCCCATTAT CATACTGACA CTAAAACCAG

AACTTATCTG AGGCTGATGC TGGTATGTCC ACCCCTCAAT TATCACAAAT ATTC----CT
AACTCATCTG AGGATGATGT TGATATGTCT ATTACTCAAT TATCACAAAT ATTCCTAACT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TTACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTTATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AGCTCACATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTTATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTTATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
```

```
AACTCATATG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AGCTCATATG AGGCTGATGT TTCTATGTGT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCAAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTA----- ---------- ---------- ----------
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCTAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCTAC ACTCATCTTT
AACTCATGTG AGGCTGATGT TTCTATGCAT ---------- --TCACCTAC ACTCATCTTT
AGCTCATGTG AGGTTGATGC TTCTATGCAT ---------- --CCACCAAC ACTTATCTTT
AGCTCATATG AGGCTGATGT TTCTATGTAT ---------- --TCACCAAC ACTCATCTTT

AACATGCCTC TCCT--TGAC AGTTGTTATC CTGGCTCTGC CTATCATTTT ATCTACTGTA
TGTCTATCCC TAACAATCAT TATCCTCATT CTA------C CTATTATAAT ATCATTGCTA
AGCTCAACCC TCCTAATCAT TTTTATTCTT CTAACATACC CCCTCATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATTAT TATAACCCTC CTAACATATC CCCTCATCGT ATCCCTCAGC
AGCTCAACCC TCCTAATAAT TTTTATCCTT CTAACATACC CCCTTATTGT ATCCACCAAC
AGCTCAACCC TCCTAATAAT TTTTATCCTT CTAACATACC CCCTTATTGT ATCCACCAAC
AGCTCAACCC TCCTAATAAT TTTTATCCTT CTAACATACC CCCTTATTGT ATCCACCAAC
AACTCAACCC TCCTAATTAT TTTTGTCCTC CTGATATTTC CCCTCATCGT ATCCCTCAAC
AGCTCAACCC TCTTAATTAT TTTTGTCCTC TTAATATTTC CCCTCATCGT ATCCCTCAAC
AGCTCAACCC TCTTAATTAT TTTTGTCCTC TTAATATTTC CCCTCATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATTAT TTTTACCCTC CTAACATATC CCCTCATTGT ATCCCTCAAC
AGCTCAACCC TCCTAATTAT TTTTACCCTC CTAACATATC CCCTCATTGT ATCCCTCAAC
AGCTCAACCC TCCTAATTAT TTTTACCCTC CTAACATATC CCCTCATTGT ATCCCTCAGC
AGCTCAACCC TCCTAATTAT TTTTACCCTC CTAACATATC CCCTCATTGT ATCCCTCAGC
AGCTCAACCC TCCTAATTAT TTTTACCCTC CTAACATATC CCCTCATTGT ATCCCTCAGC
AGCTCAACCC TCCTAATTAT TTTTATCCTC CTAACATATC CCCTCATTGT ATCCCTCAAC
AGCTCAACCC TCCTAATTAT TTTTATCCTC CTAACATATC CCCTCATTGT ATCCCTCAAC
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTAACATACC CCCTTATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTAACATACC CCCTTATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTAACATATC CCCTTATCGT ATCCCTCAAC
---------- ---------- ---------- ---------- ---------- ----------
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTGACATATC CCCTTATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTGACATATC CCCTTATCGT ATCCCTCAAC
AGCTCAACCC TCCTAATCAT TTTCATTCTT CTAACATACC CCCTTATCGT ATCCCTCAAC
AACTCAACCC TCCTCATTAT TCTTTTCCTC CTAACGTACC CCCTCATCGT CTCCCTCAAT
AACTCAGCCC TCCTCATTAT TTTCATCCTC CTAGCATACC CCCTCTTCGT CTCCCTCAAC

CTTAAACCGT CAAACAAC-- -----TGACC GCATTATGTT AAAAACGCCG TAAAATTATC
ACAAAACCAT CAAACAAC-- -----TGACT ACACCATGTA AAAAACGCTG TAAAACTATC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCGG TTCAAACAGC
CCCAACCCTC TCAACAAACA ATGGGCAACT AC-CCCTGTC AAAACCGCAG TCCAAACCGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGTCCCC TCAACAAAAA ATGAGCAATT AC-TCACGTC AAAACCGCAG TTCAAACGGC
CCCAGCCCCC TCAACAAAAA ATGAGCAATT TC-TCACGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCCC TCAACAAAAA ATGAGCAATT TC-TCACGTC AAAACCGCAG TTCAAACAGC
CCCAACCCTC TAGACAAAAA ATGGGCAACT AC-TCATGTC AAAACCGCAG TTCAAACAGC
CCCAACCCTC TAGACAAAAA ATGGGCAACT AC-TCATGTC AAAACCGCAG TTCAAACAGC
CCCAACCCTC TCAACAAAAA ATGGGCAACT AC-CCATGTC AAAACCGCAG TCCAAACGGC
CCCAACCCTC TCAACAAAAA ATGGGCAACT AC-CCATGTC AAAACCGCAG TCCAAACGGC
CCCAACCCTC TCAACAAAAA ATGGGCAACT AC-CCATGTC AAAACCGCAG TTCAAACGGC
CCCAACCCTC TCAACAAAAA ATGGGCAATT AC-CCATGTC AAAACCGCAG TTCAAACGGC
CCCAACCCCC TCAACAAAAA ATGGGCAATT AC-CCATGTC AAAACCGCAG TTCAAACGGC
CCCAGCCCTC TTAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TACAAACAGC
---------- ---------- ---------- ---------- ---------- ----------
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCAG TTCAAACAGC
CCCAGCCCTC TCAACAAAAA ATGGGCAACC AC-CCATGTC AAAACCGCGG TTCAAACAGC
CCCAGCCCTC TTAACAAAAA ATGAGCGACC AC-CCATGTT AAAACCGCAG TTCAAACGGC
CCTAACCCTC TTAACAAAAA ATGAGCGACC AC-CCATGTC AAAACCGCAG TTCAAGCGGC

ATTTTTAGTA AGCTTAGCCC CGTCACTAAC CTGCCTAAAC ATAAATCTTC AATCCTTCAC
TTTTCTAGTA AGCCTAATTC CATCCATTAC CTGCTTAAGC ATAAATCTTC AATCCTTTAC
```

```
CTTCTATGCT AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTTATTAC
CTTCTACGCC AGCCTCCTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATTAC
CTTCTATGTT AGCCTGCTCC CCCTTGCAAT ATTCTTCGAC CAGGGTATAG AAGTCATTAC
CTTCTATGTT AGCCTGCTCC CCCTTGCAAT ATTCTTCGAC CAGGGTATAG AAGTCATTAC
CTTCTATGTT AGCCTGCTCC CCCTTGCAAT ATTCTTCGAC CAGGGTATAG AAGTCATTAC
CTTCTATCTC AGCCTACTCC CCCTTGCAGT ATTTTTCGAT CAAGGCATAG AAGTCATTAC
CTTCTATGTC AGCCTATTTC CCCTTGCAGT ATTTTTTGAC CAAGGCATAG AAGTCATTAC
CTTCTATGTC AGCCTATTCC CCCTTGCAGT ATTTTTTGAC CAAGGCATAG AAGTCATTAC
CTTTTATGCC AGCCTGCTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATCAC
CTTTTATGCC AGCCTGCTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATCAC
CTTTTATGCC AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATTAC
CTTTTATGCC AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATTAC
CTTTTATGCC AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AAGTCATTAC
CTTTTATGCC AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAGGGTATAG AAGTCATTAC
CTTTTATGCC AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGTATAG AAGTCATTAC
CTTCTATGCA AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATGG AAGTCATTAC
CTTCTATGCT AGCCTGCTCC CCCTTGCAGT ATTCTTTGAC CAGGGCCTAG AAGTCATTAC
CTTCTATGCT AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATGG AAGTCATTAC
---------- ---------- ---------- ---------- ---------- ------TCAC
CTTCTATGCT AGCCTACTCC CCCTTGCAGT ATTTTTTGAC CAAGGCATAG AAGTCATCAC
CTTCTATGCT AGCCTACTCC CCCTTGCAGT ATTTTTTGAC CAAGGCATAG AAGTCATCAC
CTTCTATGCT AGCCTACTCC CCCTTGCAGT ATTCTTTGAC CAAGGCATAG AGGTCATTAC
CTTCTACATC AGCCTGCTCC CCCTCGCAAT ATTCTTTGAT CAAGGCATAG AAACCATCAC
CTTCTACATC AGCCTGCTCC CCCTCGCAAT ATTCTTCGAC CAAGGCATAG AAACCATCAC

TATCTACTAT CGATGATTTC ATATCTCACC AATAGAAATT AATATTAGCT TTCAGTTTGA
CATCTACTAC CGATGATTCT ACATCTCACC AACAGAAATT AATATTAGCT TCCAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGG CATTGAATAA ACATCGCTAC CTTTGACATT AGCATCAGCT TTAAATTCGA
TACTAACTGG CATTGAATAA ACATCGCTAC CTTTGACATT AACATCAGCT TTAAATTCGA
TACTAACTGG CATTGAATAA ACATCGCTAC CTTTGACATT AACATCAGCT TTAAATTCGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ACATTGCCAC CTTTGACATT AATATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AATATCAGCT TCAAATTTGA
CACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
CACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGG CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAATTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAATTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAATTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
AACTAACTGA CATTGAATGA ACATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATC AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
GGCTATTTAA CATTTAATAG ATTTCGTAAC CGTTGACATT AACTCCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
TACTAACTGA CATTGAATAA ATATTGCCAC CTTTGACATT AACATCAGCT TCAAATTTGA
CACCAACTGA CACTGAATAA ACATCGCCAC CTTCAACATC AACGTCAGCT TCAAATTTGA
CACCAACTGA CATTGAATAA ACATCGCCAC TTTTGACATC AACATCAGCT TCAAATTTGA

CCTATACTCT ATAGTTTTTA TAACGATCGC ACTATACGTA ACCTGATCGA TTCTAGAATT
TCAATACTCC ATAATCTTTA TAACAATTGC ACTTTACGTA ACCTGATCTA TTCTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAGTT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTATGTA ACTTGATCAA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTATGTA ACTTGATCAA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTATGTA ACTTGATCAA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTGGC CCTCTACGTA ACCTGATCCA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTGGC CCTCTACGTA ACCTGATCCA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTGGC CCTCTACGTA ACCTGATCCA TCTTAGAATT
CCAATACTCA ATTATCTTTA CACCAGTGGC CCTCTACGTA ACTTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTTA CACCAGTAGC CCTCTACGTA ACTTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACCTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACCTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACCTGATCAA TCCTAGAATT
```

```
CCAATACTCA ATTATCTTCA CACCCGTAGC CCTCTACGTA ACCTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTCA CACCCGTAGC CCTCTACGTA ACCTGATCAA TCCTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGGTCAA TTTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTGGC CCTCTACGTA ACTTGATCAA TCTTAGAATT
CCAATACTCG ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAATT
CCAATACGCA ATTAACTTCT CACCCGTAGC CCTCGACGTA ACTTGATCGA TCCTATATTT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAATT
CCAATACTCA ATTATCTTTA CACCCGTAGC CCTCTACGTA ACTTGATCAA TTTTAGAATT
CCAATACTCA ATTATTTTTA CACCCGTCGC CCTCTACGTT ACATGATCAA TCCTAGAATT
CCAATACTCA ATTATTTTTA CACCCGTCGC CCTCTACGTT ACATGATCTA TCCTAGAATT

TGCCATCTAC TACATACAAA CGGACATTTT AATTAACCGA TTTTTCAAAT ATCTCCTAAT
TGCTATTTAT TATATACATA CAGACATTAT AATCAATCGA TTCTTTAAAT ATTTATTAAC
TGCCTCATGA TACATACACT CGGACCCCAA CATAAACCGA TTCTTCAAAT ATCTGCTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ATCTACTCCT
TGCCTCATGG TACATACACT CAGACCCTAA TATAAACCGG TTCTTCAAAT ACCTACTCCT
TGCCTCATGG TACATACACT CAGACCCTAA TATAAACCGG TTCTTCAAAT ACCTACTCCT
TGCCTCATGG TACATACACT CAGACCCTAA TATAAACCGG TTCTTCAAAT ACCTACTCCT
CGCCTCATGA TATATGCACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ATTTACTCTT
TGCCTCATGG TATATACACT CAGACCCCAA CATAAACCGA TTCTTTAAGT ACTTACTCCT
TGCCTCATGG TATATACACT CAGACCCCAA CATAAACCGA TTCTTTAAGT ACTTACTCCT
TGCCTCATGA TATATACACT CAGACCCCGA CATAAACCGA TTCTTTAAAT ACCTGCTCCT
TGCCTCATGA TATATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTGCTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCCT
TGCCTCATGG TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ACCTGCTCCT
TGCCTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ATTTACTCCT
TGCCTCATGG TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAGT ATCTACTCCT
TGCCT----- ---------- ---------- ---------- ---------- ----------
TGCCTCGTGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ATCTGCTCCT
TGCCTCGTGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ATCTGCTCCT
TGCCTCGTGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTCAAAT ACCTGCTCCT
CGCTTCATGA TACATACACT CAGACCCCAA CATAAACCGA TTCTTTAAAT ACCTACTCTT
CGCTTCGTGA TATATACACT CAGACCCCAA TGTAAACCGA TTCTTTAAAT ACCTGCTCTT

TTTTTTAATC GCCATACTAA TCCTAGTAAC TGCCAACAAT ATATTTCAAC TCTTTATTGG
ATTCCTTATT GCCATAATTA TTCTAGTTAC AGCTAATAAC ATATTCCAAC TTTTTATTGG
ATTTTTGATT GCCATAATCA CCCTAGTCAC AGCCAACAAC ATATTTCAGC TGTTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TGTTCATCGG
ATTCCTAATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTTATTGG
ATTCCTAATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTTATTGG
ATTCCTAATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTTATTGG
ATTCTTAATC GCCATGATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTCATCGG
ATTCTTAATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAGC TATTCATCGG
ATTCTTAATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAGC TATTCATCGG
ATTCCTGATC GCCATAACTA CCCTAGTTAC ATCCAACAAC ATATTCCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTTAC ATCCAACAAC ATATTCCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTCAC ATCCAACAAC ATATTCCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTCAC ATCCAACAAC ATATTCCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTTAC AGCCAACAAC ATATTTCAAC TATTCATCGG
ATTCCTGATC GCCATAATTA CCCTAGTTAC AGCCAACAAC ATATTTCAAC TATTCATCGG
ATTCCTGATT GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTTCAGC TGTTCATTGG
ATTCCTGATC GCTATAATTA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TGTTCATCGG
ATTTCTGATT GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTTCAAC TGTTTATCGG
---------- ---------- ---------- ---------- ---------- ----------
ATTCTTGATT GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTTCAGC TGTTCATCGG
ATTCTTGATT GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTTCAGC TGTTCATCGG
ATTCTTGATT GCCATAATTA CCCTAGTCAC AGCCAACAAC ATATTTCAGC TGTTCATCGG
ATTCTTAATC GCCATAATTA CCCTAGTCAC AGCTAACAAT ATATTCCAAC TATTCATCGG
ATTCCTAGTC GCCATAATCA CCCTAGTCAC AGCCAACAAC ATATTCCAAC TATTTATTGG

GTGAGAGGGC GTAGGAATTA TATCATTCCT TTTAATTGGG TGATGGTACG GCCGCGCCGA
```

```
CTGGGAGGGG GTGGGTATTA TATCATTCCT ATTAATTGGA TGATGGTATG GCCGTGCTGA
CTGAGAAGGG GTGGGCATTA TATCCTTCCT ATTAATCGGA TGATGGTACG GACGCGCGGA
CTGAGAAGGG GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGAGGA
CTGAGAAGGA GTAGGCATCA TATCTTTTCT ACTAATCGGG TGATGATATG GACGCACGGA
CTGAGAAGGA GTAGGCATCA TATCTTTTCT ACTAATCGGG TGATGATATG GACGCACGGA
CTGAGAAGGA GTAGGCATCA TATCTTTTCT ACTAATCGGG TGATGATATG GACGCACGGA
CTGAGAGGGA GTAGGCATTA TGTCTTTCCT ACTAATTGGA TGATGATACG GACGTGCAGA
CTGAGAGGGG GTGGGCATTA TATCTTTCCT ACTAATTGGA TGATGATACG GACGTGCAGA
CTGAGAGGGG GTGGGCATTA TATCTTTCCT ACTAATTGGA TGATGATACG GACGTGCAGA
TTGAGAGGGA GTGGGCATTA TATCTTTCCT ACTGATCGGG TGATGATACG GACGCGCAGA
TTGAGAGGGA GTGGGCATTA TATCTTTCCT ACTGATCGGG TGATGATACG GACGCGCAGA
CTGAGAAGGG GTAGGCATTA TATCTTTCCT ACTAATCGGA TGATGATACG GACGCGCGGA
CTGAGAAGGG GTAGGCATTA TATCTTTCCT ACTAATCGGG TGATGATACG GACGCGCGGA
TTGAGAGGGA GTAGGCATTA TATCCTTCCT GCTAATCGGA TGATGATACG GACGTACGGA
TTGAGAAGGA GTAGGCATTA TATCTTTCCT GCTAATCGGA TGATGATACG GACGTGCGGA
TTGAGAAGGA GTAGGCATTA TATCTTTCCT GCTAATCGGA TGATGATACG GACGTGCGGA
CTGAGAAGGA GTAGGCATCA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGCAGA
CTGAGAAGGG GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGCGGA
CTGAGAAGGG GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGTGCGGA
---------- ---------- ---------- ---------- ---------- ----------
CTGAGAAGGA GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGCGGA
CTGAGAAGGA GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGCGGA
CTGAGAAGGA GTGGGCATTA TATCTTTCCT ATTAATCGGG TGATGATACG GACGCGCAGA
CTGAGAAGGA GTAGGCATTA TGTCCTTCCT GCTAATCGGA TGATGATACG GACGTGCAGA
CTGGGAAGGA GTAGGCATCA TATCCTTTCT ACTAATCGGA TGATGATACG GGCGCGCAGA

CGCCAACATG GCCGCACTCC AGGCAGTGGT CTATAACCGC GTAGGAGACA TCGGACTTAT
TGCCAATATA GCTGCATTAC AAGCAGTAAT TTACAATCGC GTGGGAGATA TTGGCCTAAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTAGGAGATA TTGGACTAAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGATG TTGGACTAAT
TGCCAACACC GCCGCCTTAC AAGCAGTCAT TTACAACCGA GTAGGGGACA TTGGACTAAT
TGCCAACACC GCCGCCTTAC AAGCAGTCAT TTACAACCGA GTAGGGGACA TTGGACTAAT
TGCCAACACC GCCGCCTTAC AAGCAGTCAT TTACAACCGA GTAGGGGACA TTGGACTAAT
CGCCAACACC GCCGCTTTAC AAGCAGTTAT TTATAACCGA GTGGGGGATA TTGGACTAAT
TGCCAACACC GCCGCCTTAC AAGCAGTCAT TTACAACCGA GTGGGGGATA TTGGGCTAAT
TGCCAACACC GCCGCCTTAC AAGCAGTCAT TTACAACCGA GTGGGGGATA TTGGGCTAAT
CGCCAACACC GCCGCCCTAC AAGCAGTTAT TTACAACCGA GTGGGAGACA TTGGACTAAT
CGCCAACACC GCCGCTCTAC AAGCAGTTAT TTACAACCGA GTGGGAGACA TTGGACTAAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGATA TCGGACTAAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGGGATA TTGGACTAAT
CGCAAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGACA TTGGACTAAT
CGCAAACACC GCTGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGACA TTGGACTAAT
CGCAAACACC GCTGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGACA TTGGACTAAT
TGCCAACACC GCCGCCTTAC AGGCAGTTAT TTACAACCGG GTAGGAGATA TTGGACTAAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGATG TTGGACTAAT
CGCCAACACC GCCGCTTTAC AAGCAGTTAT TTATAACCGG GTAGGAGATA TCGGACTAAT
---------- ---------- --------AT TTACAACCGG GTAGGAGATA TTGGACTGAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTAGGAGATA TTGGACTGAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTAGGAGATA TTGGACTGAT
CGCCAACACC GCCGCCTTAC AAGCAGTTAT TTACAACCGG GTGGGAGATA TTGGACTGAT
CGCCAACACT GCCGCCTTAC AAGCCGTCAT CTACAACCGA GTAGGAGACA TTGGACTAAT
CGCTAATACT GCCGCCCTGC AAGCAGTCAT CTACAACCGA GTAGGAGACA TCGGACTAAT

ACTGTCAATA TCCTGACTTT TAATTAATAC CAACTCATGA GACATCCAAC AGTTATTTAT
AATAACGATA TCTTGACTTC TAGTTAACAC CAACTCATGA GATATTCAAC AATTATTTAT
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACCTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
CTTAAGCATA GCATGATTTG CAATAAACAT AAACACCTGG GAAATTCAAC AAATGTTCGC
CTTAAGCATA GCATGATTTG CAATAAACAT AAACACCTGG GAAATTCAAC AAATGTTCGC
CTTAAGCATA GCATGATTTG CAATAAACAT AAACACCTGG GAAATTCAAC AAATGTTCGC
TTTAAGCATA GCATGATTTG CAATAAGCAT GAACACTTGA GAGATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAGCAT GAACACTTGA GAGATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAGCAT GAACACTTGA GAGATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TCTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TCTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAGATTCAAC AAATATTCGC
```

```
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATACAAC AAATATTCGC
TTTAAGCATA ACATGATTTG CAATAAACAT AAACACTTGG GAAATACAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATACAAC AAATATTCGC
TCTTAGCATA GCATGATTTG CAATAAACAT AAATACCTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCGTGATTTG CAATAAACAT AAACACCTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAGCAT AAACACTTGA GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TTTAAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
TTTGAGCATA GCATGATTTG CAATAAACAT AAACACTTGG GAAATTCAAC AAATATTCGC
CCTAAGCATG GCATGATTTG CAATAAACAT AAACACTTGA GAAATCCAAC AAATATTCGC
CCTAAGCATG GCCTGGTTTG CAATAAACAT AAATACTTGG GAAATTCAAC AAATATTTGC

TCTTACTAAA AACATAGACA TAACCCTGCC CGCCGCGGGA CTATTAATAG CTGCAACAGG
CCTTACTAAA AATATAGATA TAACCCTTCC GGCAGCCGGA CTTCTACTAG CGGCAACAGG
CTCCCCCCAA GACAACCAAG CAACCCTACC ACTCATAGGC CTAATCCTAG CTGCCACAGG
CTCCTCCCAA GACAACCAAG CAACCCTACC ACTCATAGGC TTAATCCTAG CCGCCACAGG
CTCCTCCCAA GACAACCAAA CAACCCTGCC ACTCATGGGC CTAATCCTGG CCGCCACAGG
CTCCTCCCAA GACAACCAAA CAACCCTGCC ACTCATGGGC CTAATCCTGG CCGCCACAGG
CTCCTCCCAA GGCAACCAAA CAACCCTGCC ACTCATGGGC CTAATCCTGG CCGCCACAGG
CTCCTCCCAG GACAACCAAG CGACCCTGCC ACTCATGGGC CTAATTCTAG CCGCCACAGG
CTCCTCCCAA GACAACCAAG CAACCCTGCC ACTCATGGGC CTAATTCTAG CCGCCACAGG
CTCCTCCCAA GACAACCAAG CAACCCTGCC ACTCATGGGC CTAATTCTAG CCGCCACAGG
CACCTCCCAA GGTAGCCAGG CAACCCTACC ACTAATGGGT TTAATCCTAG CTGCCACAGG
CGCCTCCCAA GGTAGCCAGG CAACCCTACC ACTAATGGGT TTGATCCTAG CTGCCACAGG
CTCCTCCCAA GATAACCAGG CAACCCTACC ACTCATAGGT TTAATCCTAG CTGCCACAGG
CTCCTCCCAA GATAACCAGG CAACCCTACC ACTCATAGGT TTAATCCTAG CTGCCACAGG
CTCCTCCCAA GATAACCAGG CAACCCTACC ACTCATAGGT TTAATTCTAG CTGCCACAGG
CTCCCCCCAA GATAACCAGG CAACCTTACC ACTCATAGGT TTAATCCTAG CTGCCACAGG
CTCCCCCCAA GATAACCAGG CAACCTTACC ACTCATAGGT TTAATCCTAG CTGCCACAGG
CTCCGCCCAA GACAACCAAG CAACCCTGCC ACTCATAGGC TTAATCCTAG CCGCCACAGG
CTCCTCCCAA GACAACCAAG CAACCCTACC ACTCATAGGC TTAATCCTAG CTGCCACAGG
CTCCTCCCAA GACAACCAAG CAACCCTACC ACTCATGGGC CTAATCCTAG CCGCCACAGG
CTCCCCCCAA GACAACCAAG CAACCCTACC ACTCATGGGC TTAATCCTAG CTGCCACAGG
CTCCCCCCAA GACAACCAAG CAACCCTACC ACTCATGGGC TTAATCCTAG CTGCCACAGG
CTCCCCCCAA GACAACCAAG CAACCCTACC ACTCATGGGC TTAATCCTAG CTGCCACAGG
CTCCCCCCAA GACAACCAAG CAACCCTACC ACTCATGGGC TTAATCCTAG CTGCCACAGG
CTCCTCCCAA GACAACCCGG CAACCCTGCC ACTCATAGGC CTAATCCTTG CCGCCACAGG
CTCCTTCCAA GACAACCAAG CAACCCTACC ACTCATAGGT CTAATCCTGG CCGCCACAGG

CAAATCCGCC CAATTTGGCC TACATCCATG ACTACCAGCA GCCATAGAAG GCCCTACCCC
CAAATCGGCT CAATTCGGCC TCCACCCATG ACTGCCAGCT GCAATAGAAG GCCCCACTCC
AAAATCAGCC CAATTCGGTC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTTGGCC TTCACCCCTG ACTTCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTTGGCC TTCACCCCTG ACTTCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTTGGCC TTCACCCCTG ACTTCCCTCA GCAATAGAAG GTCCAACACC
AAAGTCAGCT CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCGATAGAAG GTCCAACACC
AAAATCAGCT CAATTCGGCC TTCACCCCTG ACTTCCCTCA GCGATAGAAG GTCCAACACC
AAAATCAGCT CAATTCGGCC TTCACCCCTG ACTTCCCTCA GCGATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGTC TTCACCCCTG ACTTCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGCC TTCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGTC TCCACCCCTG ACTCCCTTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGTC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCC CAATTCGGTC TCCACCCCTG ACTCCCCTCA GCAATAGAAG GTCCAACACC
AAAATCAGCT CAATTCGGCC TCCACCCCTG ACTCCCTTCA GCAATAGAAG GCCCCACACC
AAAATCAGCC CAATTTGGCC TACACCCCTG GCTCCCCTCA GCAATAGAAG GTCCCACGCC
```

```
AGTTTCCGCC TTACTTCACT CTAGCACAAT AGTTGTTGCA GGGATTTTTC TACTAATTCG
AGTATCCGCC CTACTTCATT CTAGCACAAT AGTTGTTGCC GGCATTTTTC TACTAATTCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTCGTGGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT AGTTGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC TTACTACACT CCAGCACCAT GGTTGTAGCT GGCATCTTTC TCCTAATCCG
GGTCTCTGCC TTACTACACT CCAGCACCAT GGTTGTAGCT GGCATCTTTC TCCTAATCCG
GGTCTCTGCC TTACTACACT CCAGCACCAT GGTTGTAGCT GGCATCTTTC TCCTAATCCG
GGTCTCTGCC CTATTACACT CTAGCACCAT GGTCGTAGCC GGCATTTTCC TACTCATCCG
GGTCTCTGCC CTATTACACT CTAGCACCAT GGTTGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTATTACACT CTAGCACCAT GGTTGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTTGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTTC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTCGTAGCC GGCATCTTTC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTTGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT AGTCGTGGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CTAGCACCAT GGTTGTGGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC CTACTACACT CCAGCACCAT GGTCGTAGCC GGCATCTTCC TACTCATCCG
GGTCTCTGCC TTACTACACT CCAGCACTAT GGTTGTAGCC GGCATTTTCC TACTAATCCG
GGTATCTGCC CTACTACATT CTAGCACTAT AGTCGTAGCT GGTATCTTCC TATTAATCCG

CCTCCACCCC CTCATTGAAA ATAACACCAA TATTCTCACC GCCGCACTTT GTCTCGGTGC
CCTACACCCC CTAATCGAAA ACAATAATAA CATCCTCACC GCCACACTTT GTCTTGGCGC
GCTCCACCCC CTAATGGAAC ACAACCAAAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
ACTCCACCCC TTAATGGAAC ATAACCAAAT CGCCCTGACA ACCTGCCTCT GCCTTGGAGC
ACTCCACCCC CTAATGGAAC ACAACCAGGT CGCCTTGACA ACCTGCCTTT GCCTTGGAGC
ACTCCACCCC CTAATGGAAC ACAACCAGGT CGCCTTGACA ACCTGCCTTT GCCTTGGAGC
ACTCCACCCC CTAATGGAAC ACAACCAGGT CGCCTTGACA ACCTGCCTTT GCCTTGGAGC
ACTTCACCCC CTAATAGAAC ACAACCAAGC CGCCCTGACA ACCTGTCTCT GCCTCGGGGC
ACTTCACCCT CTAATAGAAC ACAACCAAGT CGCCCTGACA ACCTGTCTTT GCCTCGGGGC
ACTTCACCCT CTAATAGAAC ACAACCAAGT CGCCCTGACA ACCTGTCTTT GCCTCGGGGC
TCTCCACCCC CTGATAGAAC ACAACCAGGT CGCCCTGACA ACTTGCCTCT GCCTTGGGGC
ACTCCACCCC CTGATAGAAC ACAACCAGGT CGCCCTGACA ACTTGCCTCT GCCTTGGGGC
ACTCCACCCC CTAATAGAAC ACAACCAGGT CGCCCTAACA ACTTGCCTCT GCCTTGGGGC
ACTCCACCCC CTAATAGAAC ACAACCAGGT CGCCCTAACA ACTTGCCTCT GCCTTGGGGC
ACTCCACCCC CTAATAGAAC ATAACCAGAT CGCCCTAACA ACTTGCCTCT GTCTTGGGGC
ACTCCACCCC CTAATAGAAC ACAACCAGGT CGCCCTAACA ACTTGCCTCT GTCTTGGGGC
ACTCCACCCC CTAATAGAAC ACAACCAGGT CGCCCTAACA ACTTGCCTCT GTCTTGGGGC
ACTCCACCCC CTAATGGAAC ATAACCAAAT CGCCCTGACA ACCTGCCTCT GCCTTGGAGC
ACTCCACCCC CTAATAGAAC ATAACCAGAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
ACTCCACCCC CTAATGGAAC ATAACCAAGT CGCCCTAACA ACCTGCCTCT GCCTTGGGGC
GCTCCACCCC CTAATGGAAC ATAACCAAAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
GCTCCACCCC CTAATGGAAC ATAACCAAAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
GCTCCACCCC CTAATGGAAC ATAACCAAAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
GCTCCACCCC CTAATGGAAC ATAACCAAAT CGCCCTAACA ACCTGCCTCT GCCTTGGAGC
ACTTCACCCC TTAATAGAAC ACAACCAACT TGCGCTAACC ACCTGCCTGT GCTTGGGAGC
ACTTCACCCC CTAATAGAAC ACAACCAAAT TGCACTAACT ACCTGCCTAT GCTTAGGAGC

GATCACAACC TTATTTACCG CCACATGTGC TCTTACACAA AACGACATCA AAAAAATTGT
AATCACCACT CTTTTTACTG CCACCTGCGC CCTTACACAA AATGACATCA AAAAAATCAT
TACAACTACC TTATTCACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
TACAACTACC CTATTCACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
CGCAACTACC CTATTCACCG CCACCTGTGC CCTAACGCAA AATGACATCA AAAAAATCGT
CGCAACTACC CTATTCACCG CCACCTGTGC CCTAACGCAA AATGACATCA AAAAAATCGT
CGCAACTACC CTATTCACCG CCACCTGTGC CCTAACGCAA AATGACATCA AAAAAATCGT
CACAACTACC TTATTCACCG CCGCCTGTGC CCTAACACAA AATGATATCA AAAAAATCGT
CACAACTACC TTATTTACCG CGGCCTGTGC CCTAACACAA AATGATATCA AAAAAATCGT
CACAACTACC TTATTTACCG CGGCCTGTGC CCTAACACAA AATGATATCA AAAAAATCGT
CACAACCACC CTATTCACCG CCGCCTGCGC CCTAACACAA AATGACATCA AAAAAATCGT
CACAACCACC CTATTCACCG CCGCCTGCGC CCTAACACAA AATGACATCA AAAAAATCGT
CACAACTACC CTATTCACCG CCGCCTGCGC CCTAACACAA AATGATATCA AAAAAATCGT
```

CACAACTACC CTATTCACCG CCGCCTGTGC CCTAACACAA AATGATATCA AAAAAATCGT
CACAACTACC CTATTCACCG CCGCCTGTGC CCTAACGCAA AATGATATCA AAAAAATCGT
CACAACTACC CTATTCACCG CCGCCTGTGC CCTAACGCAA AATGATATCA AAAAAATCGT
CACAACTACC CTATTCACCG CCGCCTGTGC CCTAACGCAA AATGATATCA AAAAAATCGT
TACAACTACC CTATTCACCG CTGCCTGTGC CCTGACACAA AATGATATCA AAAAAATTGT
CACAACTACC TTATTCACCG CCGCCTGTGC CCTAACACAA AATGATATCA AAAAAATCGT
TACAACTACC TTATTCACCG CTGCCTGTGC CCTGACACAA AATGACATCA AAAAAATCGT
TACAACTACC TTATTTACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
TACAACTACC TTATTTACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
TACAACTACC TTATTTACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
TACAACTACC TTATTTACCG CTGCCTGTGC CCTAACACAA AATGACATCA AAAAAATCGT
CACCACCACC CTATTCACCG CCGCCTGTGC CCTCACTCAA AATGACATCA AAAAAATCGT
CACAACCACC TTATTCACCG CCACCTGTGC CCTCACACAA AATGATATCA AAAAAATCGT

TGCATTTTCC ACATCTAGCC AACTTGGCCT AATGATAGTC GCCATTGGGC TCAACCAGCC
CGCATTCTCC ACATCTAGCC AACTAGGATT AATAATAGTA GCGATCGGAC TTAATCAACC
GGCCTTTTCC ACATCCAGCC AACTAGGCCT AATGATGGTT ACCATCGGCT TAAATCAGCC
AGCATTCTCC ACATCCAGCC AACTAGGCCT AATGACAGGC ACCATCGGCT TAAACCAACC
AGCGTTCTCC ACATCCAGCC AACTAGGCCT AATGATAGTT ACCATCGGCC TAAACCAACC
AGCGTTCTCC ACATCCAGCC AACTAGGCCT AATGATAGTT ACCATCGGCC TAAACCAACC
AGCGTTCTCC ACATCCAGCC AACTAGGCCT AATGATAGTT ACCATCGGCC TAAACCAACC
AGCATTCTCC ACATCCAGCC AACTGGGCCT AATGATAGTT ACCATCGGCT TAAACCAACC
AGCATTCTCC ACATCCAGCC AACTGGGCCT AATGATAGTT ACCATCGGCT TAAACCAACC
AGCATTCTCC ACATCCAGCC AACTGGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
GGCATTCTCC ACATCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
GGCATTCTCC ACATCCAGTC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
GGCATTTTCC ACATCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
GGCATTTTCC ACATCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
AGCATTTTCC ACATCCAGCC AACTAGGCCT AATGATAGTT ACCATCGGCC TAAACCAACC
AGCATTTTCC ACATCCAGCC AACTAGGCCT GATAATAGTT ACAATCGGCT TAAACCAACC
GGCATTTTCC ACATCCAGCC AACTAGGCCT GATAATAGTT ACAATCGGCT TAAACCAACC
AGCCTTTTCC ACATCCAGCC AACTAGGCCT AATGATGGTC ACCATCGGCT TAAACCAACC
AGCATTTTCC ACATCCAGCC AACTAGGCCT AATGATGGTC ACCATCGGCT TAAACCAACC
AGCCTTCTCC ACATCCAGCC AACTGGGCCT AATGATGGTC ACCATCGGCT TAAACCAGCC
AGCCTTTTCC ACGTCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
AGCCTTTTCC ACGTCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
AGCCTTTTCC ACGTCCAGCC AACTAGGCCT AATGATAGTC ACCATCGGCT TAAACCAACC
AGCCTTTTCC ACGTCCAGCC AACTAGGCCT AATGATGGTC ACCATCGGCT TAAACCAACC
AGCATTTTCT ACATCCAGCC AACTAGGCTT AATGATAGTC ACCATCGGCT TAAACCAACC
AGCATTCTCC ACATCTAGCC AGCTAGGCCT AATAATGGTC ACCATCGGCT TAAATCAGCC

ACAACTAGCA TTCCTTCACA TCTGCACTCA TGCTTTCTTT AAAGCAATAC TCTTCTTATG
ACAACTAGCA TTTCTCCACA TCTGCACCCA CGCATTCTTC AAAGCCATAT TATTCCTATG
CCAACTAGCC TTCCTACACA TCTGTACCCA CGCATTCTTC AAAGCAATAC TGTTCCTATG
CCAACAAGCC TTACTACATA TCTGTACCCA CGCATTCTTT AAAGCAATAC TGTTCCTATG
CCAGTTGGCC TTCCTGCACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TATTCCTATG
CCAGTTGGCC TTCCTGCACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TATTCCTATG
CCAGTTGGCC TTCCTGCACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TATTCCTATG
CCAATTGGCC TTCTTACACA TCTGCACCCA CGCATTTTTT AAAGCAATAC TGTTCCTATG
CCAACTGGCC TTCTTACACA TCTGCACCCA CGCATTTTTT AAAGCAATAC TATTCCTATG
CCAACTGGCC TTCTTACACA TCTGCACCCA CGCATTTTTT AAAGCAATAC TATTCCTATG
CCAATTAGCC TTCCTGCACA TCTGTACCCA CGCATTTTTT AAAGCAATAT TATTCCTATG
CCAATTAGCC TTCCTACACA TCTGTACCCA CGCATTTTTT AAAGCAATAT TATTTCTATG
CCAGTTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAT TGTTTCTATG
CCAGTTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TGTTTCTATG
CCAATTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAT TGTTCCTATG
CCAATTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAT TGTTCCTATG
CCAATTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAT TGTTCCTATG
CCAACTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TGTTCCTATG
CCAATTAGCC TTCCTACACA TCTGTACCCA CGCATTCTTT AAAGCAATAC TGTTCCTATG
CCAACTAGCC TTCCTACACA TCTGTACTCA TGCATTCTTT AAAGCAATAC TGTTCCTATG
CCAACTAGCC TTCCTACACA TCTGTACCCA CGCATTCTTT AAAGCAATAC TATTCTTATG
CCAACTAGCC TTCCTACACA TCTGTACCCA CGCATTCTTT AAAGCAATAC TATTCTTATG
CCAACTAGCC TTCCTACACA TCTGTACCCA CGCATTCTTT AAAGCAATAC TATTCTTATG
CCAACTAGCC TTCCTACACA TCTGCACCCA CGCATTCTTT AAAGCAATAC TATTCCTATG
CCAATTAGCT TTCCTACACA TTTGTACCCA CGCATTCTTT AAAGCAATAC TATTCCTATG
CCAACTGGCT TTCCTGCATA TCTGCACCCA CGCTTTCTTC AAAGCAATAC TATTCCTATG

```
CTCCGGATCA ATTATTCACG CCCTTAACGA CGAACAGGAC ATCCGCAAAA TAGGAGGAAT
TTCAGGATCA ATTATTCATT CTTTAAATGA TGAACAAGAC ATTCGTAAAA TGGGAGGTAT
CTCCGGATCT ATTATCCACA GCCTTAACGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGGTCC ATCATCCACA GCCTTAATGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCTGGGTCC ATCATCCACA GCCTCAACGA CGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCTGGGTCC ATCATCCACA GCCTCAACGA CGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCTGGGTCC ATCATCCACA GCCTCAACGA CGAACAAGAC ATCCGAAAAA TAGGAGGCCT
TTCTGGATCC ATTATCCACA GCCTCAACGA CGAACAAGAT ATCCGAAAAA TAGGAGGCCT
CTCTGGATCC ATTATCCACA GCCTCAACGA CGAACAAGAT ATCCGAAAAA TAGGAGGCCT
CTCTGGATCC ATTATCCACA GCCTCAACGA CGAACAAGAT ATCCGAAAAA TAGGAGGCCT
TTCCGGATCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TAGGGGGCCT
TTCCGGATCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TAGGGGGCCT
CTCCGGATCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATCCGAAAAA TAGGGGGCCT
CTCCGGATCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TAGGGGGCCT
CTCCGGATCT ATTATCCATA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TGGGGGGCCT
TTCCGGGTCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TAGGGGGCCT
TTCCGGATCT ATTATCCACA GCCTCAACGA TGAACAAGAC ATTCGAAAAA TAGGGGGCCT
CTCCGGGTCC ATCATCCACA GCCTTAATGA TGAGCAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATCATCCACA GCCTCAACGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATCATCCACA GCCTTAACGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATTATCCACA GCCTTAATGA TGAACAAGAC ATCAGAAAAA GAGGAAGCCG
CTCCGGATCC ATTATCCACA GCCTTAATGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATTATCCACA GCCTTAATGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATTATCCACA GCCTTAACGA TGAACAAGAC ATCCGAAAAA TAGGAGGCCT
CTCCGGATCC ATTATTCACA GCCTAAACGA CGAACAAGAC ATCCGAAAAA TAGGAGGCCT
TTCCGGGTCC ATTATCCACA GCCTAAATGA TGAACAAGAC ATCCGAAAAA TGGGAGGCCT

TCATAAAATA CTTCCCCTAA CCTCCTCATG TCTTACTATT GGCAGCCTCG CACTAATAGG
TCACAAAATA CTTCCACTAA CATCTTCATG TCTTACCATC GGCAGCCTCG CACTAATGGG
TCACACCATG CTTCCGCTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATA CTCCCACTCA CCTCCGCCTG CCTCACCATT GGCAGCCTAG CCCTAACCGG
CCACACCATG CTTCCACTTA CTTCTTCCTG TCTTACCATC GGTAGCCTAG CCCTGACCGG
CCACACCATG CTTCCACTTA CTTCTTCCTG TCTTACCATC GGTAGCCTAG CCCTGACCGG
CCACACCATG CTTCCACTTA CTTCTTCCTG TCTTACCATC GGTAGCCTAG CCCTGACCGG
CCACACTATA CTTCCGCTCA CCTCTACCTG CCTCACCATC GGCAGCCTAG CTCTAACCGG
CCACACCATA CTCCCGCTCA CCTCCATCTG CCTCACCATT GGCAGCCTAG CCCTGACCGG
CCACACCATA CTCCCGCTCA CCTCCACCTG CCTCACCATT GGCAGCCTAG CCCTGACCGG
CCACACCATG CTCCCCCTCA CTTCCACCTG CCTCACCATC GGCAGTCTAG CTTTAACCGG
CCACACCATG CTCCCCCTCA CTTCCACCTG CCTCACCATC GGCAGTCTAG CTTTAACCGG
CCACACCATG CTTCCGCTCA CTTCCACCTG CCTCACCATC GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CTTCCACCTG CCTCACCATC GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CTTCCACCTG CCTCACCATC GGCAGTCTAG CCCTAACCGG
CCACACCATA CTCCCGCTCA CTTCCACCTG CCTCACTATT GGCAGTCTAG CCCTAACCGG
CCACACCATA CTCCCGCTCA CTTCCACCTG CCTCACTATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTCCCACTCA CCTCCACTTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCACTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACAACATG CTCCCGCCCA CCTCCACCCG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATG CTTCCGCTCA CCTCCACCTG CCTCACCATT GGCAGTCTAG CCCTAACCGG
CCACACCATA CTTCCATTCA CCTCAACCTG CCTCACCATT GGCAGCCTAG CTCTCACCGG
CCATACTATA CTTCCACTCA CTTCAACCTG CCTCACCATT GGCAGCCTAG CCCTCACCGG

AACGCCCTTC CTCGCCGGCT TCTTTTCAAA AGATGCAATT ATTGAGGCCA TAAATACCTC
AACTCCATTT CTAGCCGGCT TCTTCTCAAA AGACGCAATT ATTGAAGCCC TCAACACCTC
GATACCATTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GATACCCTTT CTATCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
AATACCATTC CTCTCCGGGT TCTTCTCAAA GGATGCCATC ATTGAAGCCC TAAACACATC
AATACCATTC CTCTCCGGGT TCTTCTCAAA GGATGCCATC ATTGAAGCCC TAAACACATC
AATACCATTC CTCTCCGGGT TCTTCTCAAA GGATGCCATC ATTGAAGCCC TAAACACATC
AATACCTTTC CTCTCCGGGT TCTTCTCAAA AGACGCCATC ATCGAAGCCC TAAACACATC
AATGCCTTTC CTCTCCGGGT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
AATGCCTTTC CTCTCCGGGT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
AATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
AATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
```

```
GATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
GATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
GATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
GATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
GATACCATTC CTCTCCGGAT TCTTCTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
GATACCCTTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GGTGCCATTT CTCTCCGGAT TCTTTTCAAA AGACGCCATC ATTGAAGCCC TAAACACATC
AATGCCATTT CTCTCCGGAT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GATACCATTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GATACCATTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GATACCATTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
GATACCATTT CTCTCCGGGT TTTTCTCAAA AGACGCCATC ATTGAAGCCC TGAACACATC
AATACCATTC CTCTCAGGAT TCTTCTCAAA AGACGCTATT ATTGAAGCTC TTAACACATC
AATGCCATTC CTCTCAGGAT TCTTCTCAAA AGACGCCATC ATCGAAGCCC TAAACACATC

CAACTTAAAC CTGAGCCCTA ACTCTAACAC TAATTGCCAC CTCATTTACC GCCGTATACA
GCACCTAAAC CTGAGCCCTA GTACTTACGT TAATCGCCAC TTCCTTCACC GCCGTATATA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CCTGACCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTAAAC CTGAGCCCTA ACCCTGACTC TCATTGCCAC TTCTTTCACA GCCGTGTACA
ACACCTAAAC CTGAGCCCTA ACCCTGACTC TCATTGCCAC TTCTTTCACA GCCGTGTACA
ACACCTAAAC CTGAGCCCTA ACCCTGACTC TCATTGCCAC TTCTTTCACA GCCGTGTACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TTTTCGCCAC CTCCTTCACA GCCGTGTACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TTATCGCCAC CTCCTTCACA GCCGTGTACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TTATCGCCAC CTCCTTCACA GCCGTGTACA
ACACCTAAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTAAAC CTGAGCCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTG ACCCTAACTC TCATTGCTAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CTGAGCCCTA ACCCTAACCC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATATA
ACACCTGAAC CTGAGCCCTA ACCCTAACTC TCATTGCCAC CTCCTTCACA GCCGTATACA
TCACCTAAAC CTGAGCCCTT ACCCTGACCC TTATTGCCAC CTCCTTCACA GCCGTGTACA
TCACCTGAAC CTGAGCCCTA ACTCTAACCC TCATCGCCAC CTCCTTTACA GCCGTATACA

GCCTACGTAT TATTTATTTT GTATTAATAA ATAACCCACG AACTCTTCCA CTCTCCCCCA
GCCTACGCGT TATCTACTTT GTTCTTATAA ATAATCCGCG AACCTTACCC CTCTCCCCGG
GCTTCCGAGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCCTCCGAAT TATCTTCTTC GCCTCCATGG GCTCCCCCCG GTTCCTCCCC CTATCACCCC
GCCTCCGAAT TATCTTCTTC GCCTCCATGG GCTCCCCCCG GTTCCTCCCC CTATCACCCC
GCCTCCGAAT TATCTTCTTC GCCTCCATGG GCTCCCCCCG GTTCCTCCCC CTATCACCCC
GCTTCCGGGT CATCTTCTTC GCCTCAATGG GCTCCCCCCG ATTCCTCCCC CTGTCACCCC
GCTTCCGGGT TATCTTCTTC GCCTCAATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGGGT TATCTTCTTC GCCTCAATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCCATGG GTGCCCCCCG ATTCCTCCCC TTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCCATGG GTGCCCCCCG ATTCCTCCCC TTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCCATGG GCTCCCCCCG ATTCCTCCCA TTATCCCCCC
GCTTCCGAGT TATCTTCTTC GCCTCCATGG GCTCCCCCCG ATTCCTCCCA TTATCCCCCC
GCTTCCGAGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC TTATCACCCC
GCTTCCGAAT TATCTTCTTC GCCTCTATGG GCTCCCCTCG ATTCCTCCCC TTATTACCCC
GCTTCCGAAT TATCTTCTTC GCCTCTATGG GCTCCCCTCG ATTCCTCCCC TTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGAGT TATCTTCTTC GCCTCCATGG GCTCCCCCCG ATTCCTCCCC TTATTACCCC
GCTTCCGAGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTTCTCCCC CTATCACCCC
GCTTCCGAAT TATCTTCTTC GCCTCTATGG GCTCCCCTCG ATTCCTCCCC TTATCACCCC
GCTTCCGAGT TATTTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGAGT TATTTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGGGT TATCTTCTTC GCCTCTATGG GCTCCCCCCG ATTCCTCCCC CTATCACCCC
GCTTCCGAGT CATCTTCTTC GCCTCCATAG GATCCCCCCG ATTCCTACCC CTATCCCCCC
```

```
GCCTCCGAGT CATTTTCTTT GCCTCCATAG GATCCCCCCG ATTCCTCCCC CTTTCCCCTC

CCAACGAAAA TAACCCATTA ATTATTAAT- -AACCGCACG AAGACAACCA CGAGAATAGC
TAAATGAAAA TAATCCACTA ATTATTAAT- -AACTGCACG CAAACAACCA CGGGAATAAC
TCAACGAAAA CAACCCAACA GTAATCAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA TAACCCAACA GTAATCAACT AAACAACACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA TAACCCAATA GTAATTAACT AAACCGCACG AAGAGCCCCA CGACCAAGCC
TCAATGAAAA TAACCCAATA GTAATTAACT AAACCGCACG AAGAGCCCCC CGACCGAGCC
TCAATGAAAA TAACCCAATA GTAATTAACT AAACCGCACG AAGAGCCCCC CGACCGAGCC
TCAATGAAAA CAACCCAACA GTAATAAACT AAACGGCACG AAGAGCCCCC CGACTCAACC
TCAACGAAAA CAACCCAACA GTAATAAACT AAACGGCACG AAGAGCCCCC CGACTCAACC
TCAACGAAAA CAACCCAACA GTAATAAACT AAACGGCACG AAGAGCCCCC CGACTCAACC
TCAATGAAAA CACCCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTAAGCC
TCAATGAAAA CACCCCGACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTAAGCC
TCAATGAAAA CAACCCGACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCGACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA TAACCCAACA GTAATCAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTAAGCC
TCAACGAAAA CAACCCAACA GTAATTAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAAC- ---------- ---------- ----------
TCAATGAAAA CAACCCAACA GTAATCAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATCAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATCAACT AAACAGCACG AAGAGCCCCC CGACTGAGCC
TCAACGAAAA CAACCCAACA GTAATCAATT AGACAGCCCG AAGAGCCCCC CGACTGAGCC
TCAATGAAAA CAACCCAACA GTAATTAATT AGACAGCACG CAGGGCCCCC CGACCAAGCC

CACGAGTTAA CTCCAGGACA ACTAGTAAAG CCATCAATAA CACCCATCCC GCGATTAATA
CTCGAGTTAA TTCTAAAACA ACTAATAAAG CCATTAATAA TACTCACCCA ACAATCAACA
CCCGAGTAAG CTCCAGCACT ACAAACAAAG TCAACAACAA TACCCACCCC GACACCACTA
CACGAGTAAG CTCCAGTACC ACAAACAAAG TCAACAACAA CACCCACCCC GACACCACTA
CCCGAGTAAG CTCCAACACC ACAAATAAAG TTAACAACAA CACCCACCCA GACACTACTA
CCCGAGTAAG CTCCAACACC ACAAATAAAG TTAACAACAA CACCCACCCA GACACTACTA
CCCGAGTAAG CTCCAACACC ACAAATAAAG TTAACAACAA CACCCACCCA GACACTACTA
CACGAGTAAG CTCCAACACC ACAAATAAAG TTAGCAGCAA CACCCACCCC GACACTACTA
CACGAGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CATCCACCCC GACACTACTA
CACGAGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CATCCACCCC GACACTACTA
CACGTGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CACCCACCCC GACACTACCA
CACGTGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CACCCACCCC GACACTACCA
CACGAGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CACCCACCCC GACACCACTA
CACGAGTAAG CTCCAACACC ACAAATAAAG TCAACAACAA CACCCACCCC GACACCACTA
CACGAGTAAG CTCCAACACC ACAAACAAAG TCAACAACAG CACCCACCCC GACACCACTA
CACGAGTAAG CTCCAACACC ACAAACAGAG TCAACAACAA CACCCACCCC GACACCACTA
CACGAGTAAG CTCCAACACC ACAAACAGAG TCAACAACAA CACCCACCCC GACACCACTA
CCCGAGTAAG CTCCAGTACC ACAAACAAAG TCAACAACAA CACCCACCCC GACACCACTA
CCCGAGTAAG CTCCAACACC ACAAACAAAG TCAACAACAA TACCCACCCC GACACTACTA
CCCGAGTAAG CTCCAACACC ACAAACAAAG TCAACAACAA TACCCACCCC GACACCACTA
---------- ---------- ---------- ---------- ---------- ----------
CCCGAGTAAG CTCCAGCACT ACAAACAAAG TCAACAACAA TACCCACCCC GACACCACTA
CCCGAGTAAG CTCCAACACT ACAAACAAAG TCAACAACAA TACCCACCCC GACACCACTA
CCCGAGTAAG CTCCAGCACT ACAAACAAAG TCAACAACAA TACCCACCCC GACACCACTA
CACGAGTAAG CTCCAACACC ACAAACAACG TTAACAACAA AACCCACCCC GACACTATCA
CACGAGTAAG CTCCAACACC ACTAACAAAG TTAACAACAG CACCCACCCC GATACCACCA

ATATATAACC ACCAGCATTA TACACCTCCG CCACACCAAC -ATATCGCCG CACCGCCCCA
ATATATAACC CCCAGCATTA TACACTTCTG CTACCCCAAC -ATACCGTCG TACAGCATTA
ATATCGCTCC CCCTAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCGCCCC CCCTAAAGAA TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATTACCCC CCCTAAAGAG TACATAAGCG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATTACCCC CCCTAAAGAG TACATAAGCG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATTACCCC CCCTAAAGAG TACATAAGCG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCACTCC CCCCAAAGAG TACATAAATG CCACCCCACT AAAATCCCCC CGAAGCAAGG
GTATCGCTCC CCCCAAAGAG TACATAAATG CCACCCCACT AAAATCTCCC CGAAGCAAGG
GTATCGCTCC CCCCAAAGAG TACATAAATG CCACCCCACT AAAATCTCCC CGAAGCAAGG
ACATCGCTCC CCCCAAAGAG TACATAAATG CCACCCCACT AAAATCCCCC CGAAGCAAGG
```

```
ACATCGCTCC CCCCAAAGAG TACATAAATG CCACCCCACT AAAATCCCCC CGAAGCAAGG
GCATCGCTCC CCCCAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
GCATCGCTCC CCCCAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCACTCC CCCTAAAGAG TACATAAGCG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCACTCC CCCTAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCACTCC CCCTAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCGCCCC CCCTAAAGAA TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
GCATCGCCCC CCCTAAAGAG TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
ACATCGCCCC CCCTAAAGAA TACATAAGTG CCACCCCACT AAAATCCCCC CGAAGCAAGG
---------- ---------- ---------- ---------- ---------- ----------
ATATCGCTCC CCCTAAAGAG TACATAAGTG CCACCCCGCT AAAATCCCCC CGAAGCAAGG
ATATCGCTCC CCCTAAAGAG TACATAAGTG CCACCCCGCT AAAATCCCCC CGAAGCAAGG
ATATCGCTCC CCCTAAAGAA TACATAAGTG CCACCCCGCT AAAATCCCCC CGAAGCAAGG
ACATAAACCC CCCTAAAGAG TACATAAACG CCACCCCACT AAAATCCCCA CGAAGCAACG
ACATAGGCCC CCCTAAAGAA TACATAACCG CCACCCCACT AAAATCCCCA CGAAGCACTG

AAATGCCCCT CAAT--GTCA TCAAGCAACC CAAAATCTCC TACAAATGCA ACCCACCCCA
AAACACACCT CAAT--ACTA TCAAATACCT CAAAATTCCC AACAAACACC ACCCATATTC
ACAGATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAGATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AGTCGCCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AGTCGCCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AGTCGCCCCC
ACAAATCCTT AAACTCATCA ACAATTACTC AAGAAC---- -ACGAATACC AATCGCCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATATC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATATC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AGTCACCCCC
ACAGATCCTT AAATTCATCA ACAACTACCC AAGAAC---- -ACGAATACC AATCACCTCC
ACAGATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AGTCACCCCC
ACAGATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AGTCACCCCC
ACAGATCCTT AAACTCATCG ACAACCACCC AAGAAT---- -ACGAATACC AGTCACCCCC
ACAGATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AGTCACCCCC
---------- ---------- ---------- ---------- ---------- ----------
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAGATCCTT AAATTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCCCC
ACAAATCCTT AAACTCATCA ACAACCACCC AAGAAC---- -ACGAATACC AATCATTACT
ACAAATCCGT GAACTCATCA ACAACCACCC AAGAAT---- -ACGAATACC AATCACCACT

CA-------- ---ACAATTA AA-------- AACATATACC CCAATACATA AGAAAACACC
CA-------- ---ACAATTA A--------C AATACATACC CTAGTACATA AGAAAAAACT
CACCAAACCA CCCACAACTA AAACCCCCGT AATGTAAACC --ACTACATA AAACAACACC
TGCCAAACCA CCCACAACCA AAACCCCCGC AATGTAAACC --ACTACATA AAACAACACT
CACTAAACCA CCCAAAATTA AAACACCCGC AATGTAAGTC --ACCACATA AAGCAGCACA
CACTAAACCA CCCAAAATTA AAACACCCGC AATGTAAGCC --ACCACATA AAGCAGCACA
CACTAAACCA CCCAAAATTA AAACACCCGC AATGTAAGCC --ACCACATA AAGCAGCACA
CACCAAACCT CCCACAACCA AAACTCCCGC AATGTAAACC --ACTACATA AAACAGCACT
CACCAAGCCC CCCACAACCA AAACTCCCGC AATGTAAACC --ACTACATA AAACAACACT
CACCAAGCCC CCCACAACCA AAACTCCCGC AATGTAAACC --ACTACATA AAACAACACT
CGCCAGACCA CCCGCAATCA AGACCCCCAC AATATAAACC --ACTACATA AAACAACACT
CGCCAGACCA CCCGCAATCA AGACCCCCAC AATATAAACC --ACTACGTA AAACAACACT
TACCAGACCA CCCGCAACCA AGACTCCCGC AATATAGACC --ACTACGTA AAACAACACT
CACCAGACCA CCCGCAACCA AGACTCCCGT AATATAGACC --ACTACGTA AAACAACACT
CACCAGACCA CCCGCAACCA AAACCCCCGC AATATAAACC --ACTACATA AAACAACACC
CACCAGACCA CCCGCAACCA AAACTCCCGC AATGTAAACC --ACTACGTA AAACAACACC
CACCAGACCA CCCGCAACCA AAATTCCCGC AATGTAAACC --ACTACGTA AAACAACACC
CACCAGACCA CCCACAACCA AGACTCCCGT AATATAGACC --ACTACGTA AAACAACACT
TACCAAACCA CCCACAACTA AAACCCCCGC AATGTAAACC --ACTACATA AAACAACACT
CACCAAACCA CCCACAACTA AAACCCCCGC AATATAAACC --ACTACATA AAACAACACT
---------- ---------- ---------- ---------- ---------- ----------
CACCAAACCA CCCACAACCA AAACCCCCGT AATGTAAACC --ACTACATA AAACAACACC
CACCAAACCA CCCACAACCA AAACCCCCGT AATGTAAACC --ACTACATA AAACAACACC
CACCAAACCA CCCACAACTA AAACCCCCGT AATGTAAACC --ACTACATA AAACAACACT
```

```
CACCAAAGTG CCCACAATCA AAACACCCAC AACATAAACC --ACAACATA AAACAACACC
CACCAAACCC CCCACGACCA AAACACCCAC AATGTAAACC --ACAACATA AAATAACACC

TCCCAGGAGC CCCACGCAGC CGGGTAAGGC TCCGCTGCTA ATGCAGCAGA ATATGCAACA
TCCCAAGAAC CCCACGCACG AGGGTAAGGC TCCGCCGCTA ATGCAGCCGA ATACGCAATA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCCGCTA ACGCCGCAGA ATAAGCAACA
GACCAGTCCC CCCACGTCTC AGGATAAGGC TCTGCAGCTA ACGCCGCAGA ATAAGCAACA
GACCAGTCCC CCCACGTCTC AGGATAAGGC TCTGCAGCTA ACGCCGCAGA ATAAGCAACA
GACCAGTCCC CCCACGTCTC AGGATAAGGC TCTGCAGCTA ACGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTTTC AGGATAAGGC TCTGCAGCCA GTGCAGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCAGCCA GTGCAGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCAGCCA GTGCAGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAGGGC TCCGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAGGGC TCCGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCCGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCCGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAGTCCC CCCATGTCTC AGGATAAGGC TCTGCGGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCGGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCGGCTA ACGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCCGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTTTC AGGATAGGGC TCTGCAGCTA ACGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCAGCTA ATGCCGCAGA ATAAGCAACA
---------- ---------- ---------- ---------- ---------- ----------
GACCAATCCC CCCATGTCTC AGGATAGGGC TCTGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAGGGC TCTGCAGCTA ATGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAAGGC TCTGCAGCTA ATGCCGCAGA ATAAGCAACA
GATCAATCTC CCCATGTCTC AGGATAAGGC TCCGCAGCTA ACGCCGCAGA ATAAGCAACA
GACCAATCCC CCCATGTCTC AGGATAGGGC TCCGCGGCTA AAGCCGCAGA ATAGGCAACA

CCACCAACAT ACCACCTAAA TAAATTAAAA ACAAAACTAT AGATAAAAAA GTCATACCAC
CTACCAACAT ACCACCTAAA TAAATTAAAA ACAAAATCAT AGACAAAAAA GTTATACCAC
CCACCAACAT CCCACCAAGA TAAATCAAAA ATAAAACTAA AGAAAGAAAA GAACCACCAT
CACCTAGTAT GCCACCGAAA TAAATCAAAG ACCAAACGAA AGAAAGGAAC GACCCACCAT
CCACCAACAT CCCGCCAAGA TAAATCAAAA ACAAAACCAA AGAAAGAAAA GACCCGCCGT
CCACCAACAT CCCGCCAAGA TAAATCAAAA ACAAAACCAA AGAAAGAAAA GACCCGCCGT
CCACCAACAT CCCGCCAAGA TAAATCAAAA ACAAAACCAA AGAAAGAAAA GACCCGCCGT
CTACCAACAT CCCCCCAAGG TAGATCAAAA ACAAAACCAA GGAGAGGAAA GACCCACCAT
CTACCAGCAT CCCCCCAAGA TAGATCAAAA ACAAAACCAA AGAAAGAAAA GACCCACCAT
CTACCAGCAT CCCCCCAAGA TAGATCAAAA ACAAAACCAA AGAAAGAAAA GACCCACCAT
CTACCAGTAT TCCACCGAGA TAAATCAAAA ACAAAACTAA AGAAAGGAAA GACCCACCAT
CTACCAGTAT TCCACCGAGA TAAATCAAAA ACAAAACTAA AGAAAGGAAA GACCCACCAT
CTACCAACAT CCCACCAAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CTACCAGTAT CCCACCAAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CTACCAGCAT CCCACCGAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CTACCAGCAT CCCACCGAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CTACCAGCAT CCCACCGAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CTACCAGTAT CCCACCAAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCACCAT
CCACCAGTAT TCCACCAAGA TAAATCAAAA ACAAAACTAA AGAAAGAAAA GATCCGCCAT
CCACCAACAT CCCACCAAGA TAAATCAAAA ATAAAACTAA AGAAAGAAAA GATCCACCAT
---------- ---------- ---------- ---------- ---------- ----------
CCACCAACAT CCCACCAAGA TAAATCAAAA ATAAAACTAA AGAAAGAAAA GATCCACCAT
CCACCAACAT CCCACCAAGA TAAATCAAAA ATAAAACTAA AGAAAGAAAA GATCCACCAT
CCACCAACAT CCCACCAAGA TAAATCAAAA ATAAAACTAA AGAAAGAAAA GATCCACCAT
CCACCAACAT CCCACCCAGG TAAATTAAAA ACAAAATTAA TGAAAGAAAA GACCCCCCAT
CCACTAACAT CCCACCCAAA TAAATTAAAA ACAAAACTAA TGAAAGAAAA GACCCTCCAT

ACTGCATTAA TATGCCACAT GAAACATAGG CTCACTCCTT GGATTATGCC TAGTTATACA
ACTGCATTAA TATTCCACAT GAAATATAGG CTCACTATTA GGATTATGTT TAATTGTACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTGTCACACA
GCCAAACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTATGACACA
GCCCAACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTATTACACA
GCCCAACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTATTACACA
GCCCAACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTATTACACA
GCCCAACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTGTGACACA
GCCCGACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTGTGACACA
GCCCGACCAA AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTGTGACACA
```

```
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTG GGCCTCTGCC TTGTCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTG GGCCTCTGCC TTGTCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTG GGCCTCTGCC TTATCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTG GGCCTCTGCC TTATCACACA
GCCCAACCAA GATGCCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTATCACACA
GCCCAACCAA GATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTATCACACA
GCCCAACCAA GATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTATCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTATCACACA
GTCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTATTACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCGCTCTTG GGCCTCTGCC TTGTCACACA
---------A AATACCACAT GAAACTTTGG CTCACTCCTA GGCCTCTGCC TTGTGACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTGTCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTGTCACACA
GCCCAACCAA AATACCACAT GAAATTTTGG CTCACTCCTA GGCCTCTGCC TTGTCACACA
GTCCAACCAA GATACCACAT GAAATTCTGG CTCACTCCTT GGCCTCTGCC TAATCACACA
GCCCAACCAG AATACCGCAT GAAACTTTGG CTCACTTCTT GGCCTCTGCT TAATTACACA

AATTATCACA GGCCTATTCC TGGCAATACA CTACGTCTCC GATATTAGCT CCGCCTTCTC
AATCGCTACA GGGCTATTCT TAGCAATACA CTATGTTTCT GATATTAATT CCGCCTTTTC
AATCCTAACG GGACTATTCC TTGCAATACA CTACACAGCT GACATTTCGA CGGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATCTCAA CGGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA TTACACAGCC GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA TTACACAGCC GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA TTACACAGCC GACATTTCAA CAGCCTTCTC
GATCCTAACA GGATTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
GATCCTAACA GGATTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGATTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGATTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGATTATTTC TTGCAATACA CTATACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGATTATTTC TCGCAATACA CTATACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGATTATTTC TCGCAATACA CTATACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATCTCAA CGGCCTTCTC
AATCCTAACA GGACTATTTC TCGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCTTAACA GGACTATTTC TTGCAATACA TTACACAGCC GACATTTCAA CAGCCTTCTC
AATCCTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
AATCCTAACA GGACTATTTC TTGCAATACA CTACACAGCT GACATTTCAA CAGCCTTCTC
GATTCTAACA GGATTATTTC TCGCAATACA CTACACAGCT GACATCTCAA CAGCCTTCTC
AATTCTAACA GGGCTATTTC TCGCAATACA CTACACAGCT GACATCTCAA CAGCCTTCTC

CTCCGTCGCA CATATCTCAT CTTACTACTA ATTGTTGTTT TACTAGCACT TCTTTCACCA
TTCAGTTGCA CATATCTCAT TCTACTATTA GTTATTACAT CACTTGCACT ACTCTCACCC
CTCCGTTGCC CACATCTTCT GATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCT
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTGGCACT GTTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTGGCACT GTTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTGGCACT GTTCTCCCCC
CTCCGTCGCC CACATCTCCT AATACTAATC GGACTCGCCT CTGTGGCACT ATTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGTTAATC GGACTCACCT CCGTAGCACT ATTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGTTAATC GGACTCACCT CCGTAGCACT ATTTTCCCCC
TTCTGTCGCC CACATCTCTT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCC
TTCTGTCGCC CACATCTCTT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCC
CTCTGTCGCC CACATCTCTT AATGCTAGTC GGACTCACCT CTGTGGCACT ATTCTCCCCC
CTCTGTCGCC CACATCTCTT AATGCTAGTC GGACTCACCT CTGTGGCACT ATTCTCCCCC
CTCTGTCGCC CACATCTCTT AATGCTAGTC GGACTCACCT CCGTGGCACT GTTCTCCCCC
CTCTGTCGCT CACATCTCTT AATGCTAGTC GGACTCACCT CCGTAGCACT GTTCTCCCCC
CTCTGTCGCT CACATCTCTT AATGCTAGTC GGACTCACCT CCGTAGCACT GTTCTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCT
CTCCGTCGCC CACATCTCCT AATGTTAGTT GGGCTCACCT CCGTAGCACT ATTTTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTT GGACTCACCT CCGTAGCACT ATTTTCCCCC
CTCCGTCGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ACTCTCCCCC
CTCTGTTGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCC
CTCTGTTGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTCTCCCCC
```

```
CTCTATTGCC CACATCTCCT AATGCTAGTC GGACTCACCT CCGTAGCACT ATTTTCCCCC
CTCCGTCGCC CACATCTCCT GATACTAATC GGACTCACCG CCATTGCACT CTTTTCCCCA
CTCTGTCGCC CACATCTCCT TATGCTAATT GGACTCACCG CCGTAGCACT TTTCTCCCCA

AATTTATTAA ATGACCCTGA AAACTTTACC CCCGCCAACC CACTAATTAC ACCCCCCCAC
AACCTACTAA ACGACCCAGA AAATTTCACT CCGGCCAACC CACTAATTAC CCCTCCCCAT
AACCTCCTCG GCGACCCAGA CAACTTCACA CCCGCTAACC CCCTTGTCAC TCCCCCACAC
AACCTTCTGG GTGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC TCCCCCACAC
AACCTCCTGG GAGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC ACCACCCCAC
AACCTCCTGG GAGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC ACCACCCCAC
AACCTCCTGG GAGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC ACCACCCCAC
AACCTTTTAG GCGACCCAGA CAACTTTACG CCTGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTTTTAG GCGACCCAGA CAACTTTACG CCTGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTTTTAG GCGACCCAGA CAACTTTACG CCTGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTCTTGG GCGACCCAGA CAACTTTACA CCTGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTCTTGG GCGACCCAGA CAACTTTACA CCTGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTCCTGG GTGACCCAGA CAACTTTACA CCCGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTCCTGG GTGACCCAGA CAACTTTACA CCCGCCAACC CCCTTGTCAC ACCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTTACA CCCGCCAACC CCCTTGTCAC ACCCCCGCAC
AACCTCCTGG GCGACCCAGA CAACTTTACA CCAGCCAACC CCCTTGTTAC ACCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTTACA CCAGCCAACC CCCTTGTTAC ACCCCCACAC
AACCTTCTGG GTGACCCAGA CAACTTCACA CCCGCCAATC CCCTTGTCAC TCCCCCACAC
AATCTCCTGG GCGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC TCCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTCACA CCCGCCAATC CCCTCGTCAC TCCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTCACA CC-------- ---------- ----------
AACCTCCTGG GCGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC TCCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC TCCCCCACAC
AACCTCCTGG GCGACCCAGA CAACTTCACA CCCGCCAACC CCCTTGTCAC TCCCCCACAC
AACCTCTTAG GAGACCCAGA CAACTTCACA CCCGCCAACC CCCTCGTCAC ACCCCCACAC
AATCTCCTAG GAGACCCCGA CAACTTCACA CCCGCCAACC CCCTCGTCAC ACCCCCTCAC

ATTAAGCCTG AATGATACTT CCTATTTGCT TACGCAATCC TACGCTCTAT CCCGAACAAA
ATTAAACCTG AATGATACTT CCTATTTGCC TACGCAATCC TACGCTCCAT CCCAAACAAA
ATCAAGCCCG AGTGATATTT TCTCTTCGCC TACGCCATTC TCCGATCCAT CCCAAATAAA
ATCAAACCCG CATCATACTT TCACTTTGCC TACGCCCGTC TACGATCCAT TCAAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCTAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCTAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCTAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATCC TCCGATCTAT TCCGAACAAA
ATCAAACCCG AATGATATTT CCTCTTTGCC TACGCCATCC TCCGATCTAT TCCGAACAAA
ATCAAACCCG AATGATATTT CCTCTTTGCC TACGCCATCC TCCGATCTAT TCCGAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT CCCAAACAAG
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT CCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT TCCAAACAAG
ATCAAACCCG AATGGTACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT TCCAAACAAA
ATCAAACCCG AATGATACTT TCTCTTTGCC TACGCCATTC TACGATCCAT TCCAAACAAA
ATCAAGCCCG AATGATACTT TCTCTTCGCC TACGCCATTC TCCGATCCAT CCCAAACAAA
ATCAAGCCCG AATGATACTT TCTCTTTGCC TACGCCATTC TCCGATCCAT CCCAAATAAA
ATCAAGCCCG AATGATATTT TCTCTTCGCC TACGCCATTC TCCGATCTAT CCCAAACAAA
---------- ---------- ---------- ---------- ---------- ----------
ATCAAGCCCG AATGATACTT TCTCTTCGCC TACGCCATTC TCCGATCCAT CCCAAATAAA
ATCAAGCCCG AATGATACTT TCTCTTCGCC TACGCCATTC TCCGATCCAT CCCAAATAAA
ATCAAGCCCG AATGATACTT TCTCTTCGCC TACGCCATTC TCCGATCCAT CCCAAATAAA
ATCAAGCCAG AGTGATACTT CCTCTTTGCC TACGCCATCC TCCGATCCAT CCCCAACAAA
ATCAAACCCG AGTGATACTT CCTCTTTGCC TACGCCATCC TTCGATCCAT CCCCAACAAA

CTAGGTGGAG TATTAGCCCT ACTCTTCTCG ATCGTTATCT TGATACTTGT ACCATTTCTA
TTAGGCGGAG TGCTTGCCCT ACTTTTCTCA ATCTTAATCC TAATACTCGT ACCTCTACTA
CTAGGCGGGG TTCTAGCCCT TCTATTCTCT ATCCTAGTCC TAATATTAGT GCCAATACTC
CTGGGCGGAG TACTAGCCCT TCTATTCTCT ATCCTAGTAC TAATATTAGT GCCAGTACTC
CTAGGCGGAG TACTGGCCCT CCTATTTTCT ATCCTAGTCC TAATACTGGT CCCAATACTC
CTAGGCGGAG TACTGGCCCT CCTATTTTCT ATCCTAGTCC TAATACTGGT CCCAATACTC
CTAGGCGGAG TACTGGCCCT CCTATTTTCT ATCCTAGTCC TAATACTGGT CCCAATACTC
CTAGGCGGAG TACTGGCCCT TCTATTCTCC ATCCTAGTCC TAATATTGGT ACCAGTCCTC
CTAGGCGGAG TACTGGCCCT TCTATTCTCT ATTCTAGTCC TAATACTGGT ACCAGTCCTC
```

```
CTAGGCGGAG TACTGGCCCT TCTATTCTCT ATTCTAGTCC TAATACTGGT ACCAGTCCTC
CTAGGCGGAG TACTAGCCCT TCTATTTTCC ATTCTAGTCC TAATATTAGT ACCAATACTA
CTAGGCGGAG TACTAGCCCT TCTATTTTCC ATTCTAGTCC TAATATTAGT ACCAATACTA
CTAGGCGGAG TACTAGCCCT TCTATTTTCT ATTCTAGTCC TAATATTAGT ACCAATACTC
CTGGGCGGAG TACTAGCCCT TCTATTTTCC ATTCTAGTCC TAATATTAGT ACCAATACTC
CTAGGCGGAG TACTAGCCCT TCTATTCTCC ATCCTAGTCC TAATATTAGT ACCAATACTC
CTGGGCGGAG TACTAGCCCT TCTATTTTCC ATCCTGGTCC TAGCATTGGT ACCAATACTC
CTGGGCGGAG TACTAGCCCT TCTATTCTCC ATCCTGGTCC TAGCATTGGT ACCAATACTC
CTAGGCGGAG TCCTAGCCCT CCTGTTCTCT ATCCTAGTAC TAATATTAGT GCCAATGCTT
CTAGGCGGGG TACTAGCCCT TCTATTCTCT ATTCTAGTCC TAATATTAGT GCCAGTACTC
TTAGGTGGGG TTTTAGCCCT TCTATTCTCT ATTCTAGTCC TAATATTAGT ACCAATACTC
---------- ---------- ---------- ---------- ---------- ----------
CTAGGTGGAG TTCTAGCCCT TCTATTCTCT ATCCTAGTCC TAATATTAGT ACCAATACTC
CTAGGTGGAG TTCTAGCCCT TCTATTCTCT ATCCTAGTCC TAATATTAGT ACCAATACTC
CTAGGTGGAG TTCTAGCCCT TCTATTCTCT ATCCTAATCC TAATATTAGT GCCAGTACTC
CTAGGTGGGG TACTGGCCCT ACTTTTCTCC ATCCTAGTAC TAATACTAGT ACCAATCCTC
CTAGGCGGAG TACTGGCCCT ACTCTTCTCC ATCCTAGTAC TTATACTAGT ACCAATCCTT

CACACATCTA AAATCCGCAG CACCACATTT CGTCCCCTAT TTAAAATTAC CCTATGAATC
CACACTTCAA AAATCCGCAG CACTACATTC CGCCCCTTAT TCAAAATTAC ACTTTGAATC
CACACCTCTA AACAACGAGG AAACACATTC CGACCCCTCT CCCAAATTCT ATTCTGGGCC
CACACCTCTA AACAACGGGG AAACACATTC CGACCCCTTT CTCAAATTCT ATTCTGGACC
CACACCTCTA AACAACGAGG AAACACATTC CGGCCCCTCT CTCAAATCTT ATTCTGGACC
CACACCTCTA AACAACGAGG AAACACATTC CGGCCCCTCT CTCAAATCTT ATTCTGGACC
CACACCTCTA AACAACGAGG AAACACATTC CGGCCCCTCT CTCAAATCTT ATTCTGGACC
CACACCTCCA AACAACGGGG AAATACATTT CGGCCCCTCT CCCAAATCCT ATTTTGAGCC
CACACCTCCA AACAACGAGG AAATACATTT CGGCCTCTCT CCCAAGTCCT ATTTTGAGCC
CACACCTCCA AACAACGAGG AAATACATTT CGGCCTCTCT CCCAAGTCCT ATTTTGAGCC
CACACCTCTA AACAACGAGG AAACACATTC CGGCCCCCTT CCCAAATCCT ATTCTGAGCC
CACACCTCTA AACAACGAGG AAACACATTC CGACCCCCTT CCCAAATCCT ATTCTGAGCC
CACACCTCTA AACAACGAGG AAACACATTC CGACCCCTTT CTCAAATCCT ATTCTGGGCC
CACACCTCTA AACAACGAGG AAACACATTC CGACCCCTCT CTCAAATCCT ATTCTGGGCC
CACACCTCTA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATCCT ATTCTGGGCC
CACACCTCCA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATCCT ATTCTGGGCC
CACACCTCCA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATCCT ATTCTGGGCC
CACACCTCTA AGCAACGAGG AAACACATTT CGGCCCCTTT CCCAAATTCT ATTCTGGGCC
CACACCTCTA AACAACGGGG AAACACATTC CGACCCCTTT CTCAAATTCT ATTCTGGACC
CACACTTCTA AACAACGAGG AAACACATTC CGGCCCCTCT CCCAAATTCT ATTCTGAACC
---------- ---------- ---------- ---------- ---------- ----------
CACACCTCTA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATTCT ATTCTGAGCC
CACACCTCTA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATTCT ATTCTGAGCC
CACACCTCTA AACAACGAGG AAACACGTTC CGACCCCTTT CTCAAATTCT ATTCTGAGCC
CACACCTCCA AACAACGAGG AAACACATTC CGACCACTCT CCCAAATCCT CTTCTGAACC
CACACCTCTA AACAACGAGG AAACACATTC CGACCACTCA CCCAAATCCT ATTCTGAACC

CTTGCAGCCG ACGTCCTCAT CCTAACCTGA ATCGGAGGAC AACCAGTAGA AGACCCCTAC
CTTGCAGCTG ATGTCCTAAT CCTAACCTGA ATCGGAGGAC AGCCCGTAGA AGACCCATAT
CTAGTGGCCG ACATACTAGT GCTCACATGA ATCGGAGGCC AACCAGTTGA ACACCCGTTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTTGA ACACCCGTTC
CTGGTGGCCG ACATACTGGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTGGTGGCCG ACATACTGGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTGGTGGCCG ACATACTGGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTGGCCG ATATATTAGT ACTCACATGA ATTGGAGGCC AGCCAGTCGA ACATCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATCGGAGGCC AGCCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATCGGAGGCC AGCCAGTCGA ACACCCATTC
CTAGTAGCCG ACATGTTAGT GCTCACATGA ATTGGAGGCC AACCAGTCGA ACATCCATTC
CTAGTAGCCG ACATGTTAGT GCTCACATGA ATTGGAGGCC AACCAGTCGA ACATCCATTC
CTAGTAGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTAGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AGCCAGTCGA ACACCCATTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTTGA ACACCCGTTC
CTAGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AGCCAGTCGA ACACCCGTTC
---------- ---------- ---------- ---------- ---------- ----------
CTGGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTTGA ACACCCGTTC
```

```
CTGGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTTGA ACACCCGTTC
CTGGTGGCCG ACATACTAGT ACTCACATGA ATTGGAGGCC AACCAGTTGA ACACCCGTTC
CTAGTAGCTG ATATGTTAGT ACTCACATGA ATCGGGGGCC AACCAGTCGA ACACCCATTC
CTGGTAGCCG ACATACTAGT ACTCACATGA ATCGGAGGCC AACCAGTCGA ACACCCATTT

ATCACTATTG GTCAAGTAGC CTCAATTCTT TACTTCATAA TCTTCCTGGT ATTCATACCC
ATTGCCATTG GCCAAGCAGC TTCAATTCTC TACTTCCTAA TTTTCCTAGT ATTTATACCC
GTCTTAATTG GACAGGTGGC CTCCACAGTT TATTTTGCCC TATTCCTAAT CGCCCTCCCT
GTCTTAATTG GACAAGTAGC CTCCACAGTC TATTTCACCC TATTTCTAAT CGCCCTCCCT
GTCTTAATCG GACAGGTGGC CTCCACAGCC TATTTCGCCC TGTTTCTAAT CGCCCTCCCT
GTCTTAATCG GACAGGTGGC CTCCACAGCC TATTTCGCCC TGTTTCTAAT CGCCCTCCCT
GTCTTAATCG GACAGGTGGC CTCCACAGCC TATTTCGCCC TGTTTCTAAT CGCCCTCCCT
GTCTTAATTG GACAAGTGGC TTCCACAGTC TATTTCACCC TGTTTTTAGT CGCCCTCCCT
GTCTTAATCG GACAAGTGGC TTCCACAGTC TATTTCACCC TGTTTTTAGT TGCCCTCCCT
GTCTTAATCG GACAAGTGGC TTCCACAGTC TATTTCACCC TGTTTTTAGT TGCCCTCCCT
GTCTTAATCG GACAAGTAGC TTCCACAATC TATTTCGCCC TATTTCTAAT TGCCCTCCCC
GTCTTAATCG GACAAGTAGC TTCCACAATC TATTTCGCCC TGTTTCTAAT TGCCCTCCCC
GTCTTAATCG GACAAGTAGC CTCCACAGTC TATTTCGCCC TATTTCTAAT CGCCCTCCCC
GTCTTAATCG GACAAGTAGC CTCCACAGTC TATTTCGCCC TATTTCTAAT CGCCCTCCCC
GTCTTAATCG GACAAGTAGC TTCCACAGTC TATTTCGCCC TATTTTTAGT CGCCCTCCCT
GTCTTAATCG GACAAGTGGC TTCCACAGTC TATTTCACCC TATTTTTAGT CGCCTTCCCC
GTCTTAATCG GACAAGTGGC TTCCACAGTC TATTTCACCC TATTTTTAGT CGCCTTCCCC
GTCTTAATTG GACAGGTGGC CTCCACAGTC TATTTTGCCC TATTCCTGAT CGCCCTCCCT
GTCTTAATTG GACAGGTAGC CTCCACGGTC TATTTTGCCC TATTTCTAAT CGCCCTCCCT
GTCCTAATTG GACAGGTAGC TTCCACAGTC TATTTTGCCC TATTTTTAAT CGCCCTCCCT
---------- ---------- ---------- ---------- ---------- ----------
GTCTTAATTG GACAGGTAGC TTCCACAGTT TATTTTGCCC TATTCCTAAT CGCCCTCCCT
GTCTTAATTG GACAGGTAGC TTCCACAGTT TATTTTGCCC TATTCCTAAT CGCCCTCCCT
GTCTTAATTG GACAGGTAGC TTCCACAGTT TATTTTGCCC TATTCCTAAT CGCCCTCCCT
GTCCTAATCG GACAAGTTGC CTCCACGATC TATTTCGCCC TATTCCTAGT TGCTCTCCCC
GTCCTAATCG GACAAGTAGC CTCCACAATT TACTTCGCCC TATTCCTAAT TGCCCTCCCC

CTATCCGGCT GAATTGAAAA CAAAATACTT AACCGCAACT CAAAGATTTG GTCCTGGTCT
CTATCAGGCT GAATCGAAAA CAAAATACTT AACCGTAACT CAAAGATTTG GTCCTAGTCT
CTGACAGGCT TACTAGAAAA TAAAGCCTTA AACTGAAA-- ---------- ----------
ATAACGGGAT GACTAGAAAA TACAGACCTA AACTGAACCT CAAAGGATTG GTCCTGGCCG
CTGGCCGGCC TACTGGAGAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGACCT
CTGGCCGGCC TACTGGAGAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGACCT
CTGGCCGGCC TACTGGAGAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGACCT
ATAACGGGCT GACTGGAAAA TAAAACCCTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
GTATCTGGCT GACTGGAAAA TAAAACCCTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
GTATCTGGCT GACTGGAAAA TAAAGCCCTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAACCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACCGGCT GACTAGAAAA TAAAACCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACTGGCT TACTAGAAAA CAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACTGGCT TACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACTGGCT TACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
---------- ---------- ---------- ---------- ---------- ----------
CTGACTGGCT TACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACTGGCT TACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTGACTGGCT TACTAGAAAA TAAAGCCTTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
CTAACCGGTT GACTAGAAAA CAAAATCCTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT
TTAATCGGCT GACTAGAAAA CAAAATCCTA AACTGAAACT CAAAGGCTTG GTCCTGGCCT

TACTATCATT TTTCAACTAA ACTTACACAT GCAAGTATCC GCACTCCGGT GAAAATGCCC
TACTATCATT TTT-AACTAA ACTTACACAT GCAAGTATCC GCGCTCCGGT GAAAATGCCC
---------- ---------- ---------- ---------- ---------- ----------
CACTATCAAT TTT-AACCCA AGTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTATACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTATACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTATACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCCGT GAGAATGCCC
```

```
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCATCCCTGT GAGAATGCCC
---------- ---------- ---------- ---------- ---------- ----------
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCCGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCTGT GAGAATGCCC
TACTATCAAT TTT-AACCCA ATTTACACAT GCAAGTCTCC GCACCCCCGT GAGAATGCCC

TCAATCTTCC AGAAGACGAG GAGCCGGCAT CAGGCTCGTG TCCACGGCCC AAGACGCCTT
TCAATCTTCC AGAAGATAAG GAGCCGGCAT CAGGCTCGTG TCCACGGCCC AAGACGCCTT
---------- ---------- ---------- ---------- ---------- -------CTT
TTAATCACCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACCCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCGGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCGGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCGGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
---------- ---------- ---------- ---------- ---------- ----------
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAGACGCCTT
TTAATCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA ACCGCAGCCC AAGACGCCTT
TTAGTCCCCC AGGGGAAAAG GAGCAGGTAT CAGGCACGCA CCCGCAGCCC AAAACGCCTT

GCTTTGCCAC ACCCCCACGG GATTTCAGCA GTGATAAACA TTAGGCAATT AGCGAAAGCT
GCTTAGCCAC ACCCCCACGG GATTTCAGCA GTGATAAATA TTAGGCAATC AGCGAAAGCT
TTCAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCC GCGATAAACA TTGAGACATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAGACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCCATG AGCGCAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCCATG AGCGCAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGCAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGCAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCCATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
---------- ---------- ---------- ---------- ---------- ----------
```

```
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCTATG AGCGTAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCCATG AGCGCAAGCT
GCTAAGCCAC ACCCCCAAGG GAACTCAGCA GTGATAAACA TTGAGCCATG AGCGCAAGCT

AGACCTAGTT ATGGTTAAAT AGAGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
AGACCTAGTT ATAGTTAAAT AGAGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGAATCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCAGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
---------- ---------- ---------- ---------- ---------- ----------
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA
CGACTCAGCC AGAGTTAAG- AGGGCCGGTA AAACTCGTGC CAGCCACCGC GGTTATACGA

GAGGCTCAAA ATGATAGTCA TTCGGCGTAT GGTGTGGTTA GCAATTAAGT TATATACTGA
GAGGCTCAAA ATGATGGTCA TTCGGCGTAT GGCGTGGTTA ACAATAAACT A-TAAACTGA
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATACCT AGTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ATTACACTAG
GAGGCCCCAA CTGATAGTT- CACGGCGTAA AGCGTGATTA AAGGATACCC ACTGCACTAG
GAGGCCCCAA CTGATAGTT- CACGGCGTAA AGCGTGATTA AAGGATACCC ACTGCACTAG
GAGGCCCCAA CTGATAGTT- CACGGCGTAA AGCGTGATTA AAGGATACCC ACTGCACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGTCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATATCT ATTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATATCT ATTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGACACCT ACTACACTAG
GAGACCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGACACCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCT ACTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATACCC ATTATACTAG
GAGGCCCCAA CTGATGGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCC ATTATACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATGCCC ATTATACTAG
---------- ---------- ---------- ---------- ---------- ----------
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATACCT ATTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATACCT ATTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGATACCT ATTACACTAG
GAGGCCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGACGCCC ACTACACTAG
GAGACCCCAA CTGATAGTC- CACGGCGTAA AGCGTGATTA AAGGACTCCC GCCACACTAG

GAACAAAATA CCTTTAAGCT GTCATACGCT AATAAGTATA TGAAAATCAT TAACGAAAGT
GAACAAAATA CCTTTAAGCT GTCATACGCT AACAAGTATA TGAAAATCAT AAACGAAAGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CTTGAGAGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CTCCTAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCCAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCCAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCCAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
```

```
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAAGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAAGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAAGT
AGCCAAAAGC CTCCTAAGCC GTCATACGCA CCTGAAGGCC AGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CTCCTAAGCC GTCATACGCA CCTGAAGGCC AGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CTCCTAAGCC GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CTCCTAAGCC GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CTCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAGGGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAAGGCC CGAAGCCCAA CCACGAAGGT
---------- ---------- ---------- ---------- ---------- ----------
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAGAGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAGAGCC CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCA CCTGAGAGCC CGAAGCCCAA CCACGAAGGT
AGTCAAAACC CCCCTAAGCT GTCATACGCA TTTGGAGGCT CGAAGCCCAA CCACGAAGGT
AGCCAAAAGC CCCCTAAGCT GTCATACGCT CCTGGAAGCC CGAAGCCCAA CCACGAAGGT

TATCTCACTA AATG------ ---TGAACCC ACGAAAACCA AGGCACAAAC TGGGATTAGA
TATCTCATCA CATG------ ---TGAACCC ACGAAAACCA AGGCACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC GACAAGGATC CCTTGAACCC ACGACAACTG AGGCACAAAC TGGGATTAGA
AGCTCTACCC GACAAGGATC CCTTGAACCC ACGACAACTG AGGCACAAAC TGGGATTAGA
AGCTCTACCC GACAAGGATC CCTTGAACCC ACGACAACTG AGGCACAAAC TGGGATTAGA
AGCTCTACCT AACGAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGAACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGAACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGGACC CCTTGAACCC ACGGCAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
---------- ---------- ---------- ---------- ---------- ----------
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAAGGACC CCTTGAACCC ACGACAACTG AGACACAAAC TGGGATTAGA
AGCTCTACCT -ACAAGGACC CCTTGAACCC ACGACAACCG AGACACAAAC TGGGATTAGA
AGCTCTACCC AACAGGGACC CCTTGAACCC ACGACAACCG AGACACAAAC TGGGATTAGA

TACCCCACTA TGCTTGGAAC TAAACCAAGG CGGCACAAAT AC-TAAGCCG CCCGCCAGGT
TACCCCACTA TGCTTGGAAC TAAACTAAGG CGGCATAAAT AC-TAAGCCG CTCGCCAGGT
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
GACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG CGATA-AATC ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG CGATA-AATC ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG CGATA-AATC ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTA CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTTAGTCA TAAACTTTGG TAATA-AATT ACATATATTA CCCGCCAGGG
TACCCCACTA TGCTTAGTCA TAAACTTTGG TAATA-AATT ACATATATCA CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACCTTGG TAATA-AATT ACACATATTA CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACCTTGG TAATA-AATT ACACATATTA CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TAATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TAATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TAATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TAATA-AATT ACACATATTG CCCGCCAGGG
```

```
---------- ---------- ---------- ---------- ---------- ----------
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATTG CCCGCCAGGG
TACCCCACTA TGCTCAGTCA TAAACTTTGG TGATA-AATT ACACATATCA CCCGCCAGGG
TACCCCACTA TGCTCAGTCG TAAACTTTGG TGATA-AATT ACACATATCG CCCGCCAGGG

TACTACGAGC GCAAGCTTAA AACCCAAAGG ACTTGGCGGT GCTTCAGAAC CACCTAGAGG
TACTACGAGC GCAAGCTTAA AACCCAAAGG ACTTGGCGGT GCTTCAGAAC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGC
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
---------- ---------- ---------- ---------- ---------- ----------
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG
TACTACGAGC GCTAGCTTAA AACCCAAAGG ACTTGGCGGT GCCCCAGACC CACCTAGAGG

AGCCTGTTCT ATAATCGATA ATCCTCGTTC AACCTCACCG CATCTTGCA- -TCCAGCCTA
AGCCTGTTCT ATAATCGATA ATCCTCGTTA AACCTCACCA CATCTTGCA- -TCCAGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
ACCCTGATCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGGCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
---------- ---------- ---------- ---------- -----TGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTACCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA
AGCCTGTTCT AGAACCGATA ATCCCCGTTA AACCTCACCA CTTCTTGTCA TTTCCGCCTA

TATACCGCCG TCGCCAGCCT ACCTTTTGAA AGACCCTAAG TAGGCACAAC AAGCATAGCT
TATACCGCCG TCGCCAGCCT ACCTTCTGAG AGACCATAAG TAGGCAAAAT AAGTAGAACT
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCA TCGTAAGCTA ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
```

```
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACTAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAATAG TAAGCAAAAA TGGCACACCC
TATACCGCCG TCGTCAGCTT ACCCTGTGAA AGACCAACAG TAAGCAAAAA TGGCACACCC

TAACACGTCA GGTCAAGGTG TAGCTTATGA TGCGGGAAGA AATGGGCTAC ATTTTCTAAA
TAATACGTCA GGTCAAGGTG TAGCTTATGA TGTGGGAAGA AATGGGCTAC ATTTTCTAAA
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATAGGCTAC ATTTTCTGAT
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAT
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC
AAAAACGTCA GGTCGAGGTG TAGCGAATGA AGTGGAAAGA AATGGGCTAC ATTTTCTGAC

TTAG--AATA TACGGACGAC ACTATGAAAC CTGTGTCTAA AGGAGGATTT AGCAGTAAGC
TTAG--AACA CACGGACGAC ACTATGAAAC CTGTGTCTGA AGGAGGATTT AGCAGTAAGC
ACAGAAAACA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAACA CACGAATAAT ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTATGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAG
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAG
```

```
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGGTTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAACA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA
ACAGAAAATA CACGAATAAC ACTGTGAAAC CAGTGATTGA AGGTGGATTT AGCAGTAAAA

GGGGAGTAGA GTGCCCCACT GAAGCCTGGC GCTGAAGCGC GCACACACCG CCCGTCACTC
GGGGAATAGA GTGCCCCACT GAAACT-GGC ACTGAAGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCT-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA GAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA GAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAGTTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGAGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGAGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGAGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGAGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA GAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC
AGAAAATAGA AAATTCTTTT GAAGCC-GGC TATGGGGCGC GCACACACCG CCCGTCACTC

TCCTCGAAAC AAT-CTACAC TTTTAATAAA ATACTACTAA C--AAGAAGA GGAGGCAAGT
TCCTCGAAAT ATCCAAAAAA ATTTAATAAA ACACCATTAT C--AAGAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCAA AT---ATATA AATCTATAAC CCAAGCAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA AT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG GACGCCCTAA GT---ACATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG GACGCCCTAA GT---ACATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG GACGCCCTAA GT---ACATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACACCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACACCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACACCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ACATA AGTCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ACATA AGTCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACCCCAA GT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCGA AT---ATATA AACCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCAA AT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCAA AT---ATATA AACCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AACACACCAA GT---ATATA AATCTACAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCAA AT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCAA AT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATAACCCAA AT---ATATA AATCTATAAC CCAAACAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCCCA GT---ATATA AAAATATAAC CCAAGCAAGA GGAGGCAAGT
TCCTCAAAGG AATACCCTCA AT---ATATA AAAACATAAC CCAAACAAGA GGAGGCAAGT

CGTAACATGG TAAGCGTACC GGAAGGTGTG CTTGGAATAA TGCATTATAG CTAGCCTGAC
CGTAACATGG TAAGCGTACC GGAAGGTGTG CTTGGAACAA TGCATTACAG CTAGCCTAAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTGAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
```

```
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC
CGTAACATGG TAAGTGTACC GGAAGGTGCA CTTGGAACAA CGCTAAATAG CTAGCCTCAC

CATGCAATCA AATAAATATT AAAATAAACA C--------- -ACAAAAAAT AAAAACAAAA
CATACAACCA AACAATTATT AAAATATATA C--------- -TCAAATAAT AAAAACAAAA
CACACACATA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACATACACA AATGAATATT TATATACAAC CCCCCATAAG ATAAAAGAAA ATAAACAAAC
CACACACACA AATGGATACT TATACATAAA CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGGATACT TATACATAAA CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGGATACT TATACATAAA CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACATACACA AATGAATATT TATATACAAC CCCCCATAAG ATAAAAGAAA ACAAACAAAC
TACATACACA AATGAATATT TATATATAAC CCCCCATAAA ATAAAAGAAA ACAAACAAAC
TACATACACA AATGAATATT TATATATAAC CCCCCATAAA ATAAAAGAAA ACAAACAAAC
CACATACACA AATGAATATT TATATATAAA CCCCAATAAG ATCAAAAAAA ATAAACAAAC
CACATACACA AATGAATATT TATATATAAA CCCCAATAAG ATCAAAAAAA ATAAACAAAC
CACATACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAC ATCAAAAAAA ACAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ACAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ACAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATACT TATACATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CTCCCATAAG ATCAAAAAAA ATAAACAAAC
CACACACACA AATGAATATT TATATATAAC CTCCCATAAG ATCAAAAAAA ATAAATAAAC
CACACACACA AATGAATATT TATATATAAC CCCCCATAAG ATCAAAAAAA ATAAACAAAC
CACAAACACA AATAAATATT TATATATAAC CCCCCATAAA AACAAAAAAA ATAAACAAAC
CACAAACACA AATGAATATT TATATATAAC CCCCCATAAG AATAAAAAAA ATAAACAAAC

CATTTGTCAA CTTTAGTATA GGCGATAGAA AAAGACCCA- -CAGAGCTAT AGCAATAGTA
CATTTGTTAA CTCCAGTATA GGTGATAGAA AGAGAACAA- -CAGAGCTAT AGCAATAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT ACCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCACAAT AGAGAAAGTA
CATTTAATCC CCCTAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAAAAAGTA
CATTTAATCC CCCTAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAAAAAGTA
CATTTAATCC CCCTAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAAAAAGTA
CATTTAATCC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAAAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAAAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATAC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCC CCCCAGTATA GGCGATAGAA AAGGACAAAG -CAGCGCAAT AGAGAAAGTA
```

```
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCC CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATCT CCCCAGTATA GGCGATAGAA AAGGATAAAG -CAGCGCAAT AGAGAAAGTA
CATTTAATTT CCCCAGTATA GGCGATAGAA AAGGACAAAA -CAGCGCAAT AGAGAAAGTA
CATTTAATCC CCCCAGTATA GGCGATAGAA AAGGACAAAG TCAGCGCAAT AGAGATAGTA

CCGCAAGGGA AAGCTGAAAG AGAAATGAAA TAAATCGTTA AAGCACAACA AAGCAGAGAT
CCGCAAGGGA AAGCTGAAAA AGAAATGAAA TAAATCGTCA AAGCACAACA TAGCAGAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAACAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTCGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAATAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAATGAAA CAACTTGTTA AAGCAACAAA AAGCAAAGAT
CCGCAAGGGA AAGCTGAAAG AGAAGTGAAA CAACTTGTTA AAGCAACAAA AAGCAAAGAT

AAAGCCTCGT ACCTTTTGCA TCATGGTCTA GCAAGTCAGA TCCAAGCAAA ATGATTTATA
TAAATCTCGT ACCTTTTGCA TCATGATCTA GTAAGTAAGC -CCAAGCAAA ATGATTTATA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTCGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTCGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTCGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAGAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTCA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTCA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTCA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT ATCAGGCGAA GAGAACTTTA
TAAATCTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT GTCAGGCAAA GAGAACTTTA
TAAAACTTGT ACCTTTTGCA TCATGATTTA GCCAGTTCTT CTCAGGCAAA GAGAACTTTA

GTTTGACCCC CCGAAACTAG ACGAGCTACT TCGAAGCAGT TGAA-AGGAC CAACCCGTCT
GTTTGACCCC CCGAAACTAG ACGAGCTACT TCGAAGCAGT TGAA-AGGAC CAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAACTACT CCAAGACAGC CTATTAGGGC ATACCCGTCT
```

```
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAACAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAACAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAACAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
TTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAATAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAACAGGGC AAACCCGTCT
GTCTGACCCC CCGAAACTAG ACGAGCTACT CCGAGACAGC CTAACAGGGC AAACCCGTCT

CTGTGGCAAA AGAGTGGGAA GACTTCCAAG TAGAGGTGAC AAGCCTAACG AGCCTAGTGA
CTGTGGCAAA AGAGTGGGAA GACTTCCAAG TAGAGGTGAC AAGCCTAACG AGCCTAGTGA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA AATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA
CTGTGGCAAA AGAGTGGGAA GATCTCCGAG TAGAGGCGAC AAACCTAACG AGCCTAGTAA

TAGCTGGTTA CTTGAGAAAT GAATAAAAGT TCAGCCTCAA AAATTTT--C TAAAAAATAT
TAGCTGGTTA CTTGAGAAAT GAATAAAAGT TCAGCCTCAA ATATTCTAAA AAATAAACAA
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
GAGCTGGTTG CGCAGGAAAC GAACACTCCT TCAGCCACAA GGCTTCT-AC TGTCACTTAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAA
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAA
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAA
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTTAA GACTTCT-AC TGCCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTTAA GGCTTCT-AC TGCCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTTAA GGCTTCT-AC TGCCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
```

```
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACTCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACTCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC CGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAGGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCACCCAG
TAGCTGGTTG CTCAAGAAAT GAATATTAGT TCAGCCTCAA GGCTTCT-AC TGTCAACCAG
TAGCTGGTTG CTCAAGAAAT GAATATTAGT TCAGCCCTAA GGCTCCT-AC TGTCAACCAA

GTAAGTTCTA ACTAAAAAAT TTTAAGAACA TTTGAGAGTT ATTCACAGGA GGTACAGCTC
GTTCTTACTA A--AAAATTT CAAGGATA-- TTTGAGAGTT ATTCACAGGA GGTACAGCTC
GTCATCACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACCAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACCAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACCAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCACTACCA ACAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCACTACCA ACAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATCACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATCACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATCACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCACTACCA ACAAAGACAT CAAGAAACC- CTTAAGAGTT ATTCAACACK CCTCCCCCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAT CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATTACCA ACAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTCATCACCA ACAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC
GTTAACACCA ATAAAGACAC CAAGAAAAC- CTTAAGAGTT ATTCAAGAGA GGTACAGCTC

CTATGAACCG GGAAACAACC CAATAAGGAG GAAAAAGATC A-----TAAT CCAAAGGACA
CTATGAACTG GGAAACAACC CAATAAGGAG GAAAAAGATC A-----TAAT TTAAAGGATA
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
CCTTGAAAAA GAACACAACC TTAACAGGTG GATAAAGATC ATATT-AAAT CAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGTG GATAAAGATC ATATT-AAAT CAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGTG GATAAAGATC ATATT-AAAT CAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGTG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGTG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGAACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGACCT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGACCT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGACCT
TCTTCCCCYT TTTTTCAAKA WWAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT TAAAGGACCT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGAACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGAACT
TCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACATT-AAAT CAAAGGGACT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ATATTCAAAC CAAAGGAAAT
CCTTGAAAAA GAACACAACC TTAACAGGCG GATAAAGATC ACACT-AAAC CAAAGGAAAT

AAATCCAAGT GGGCCTGAAA GCAGCCACCT TTAAAGAAAG CGTTATAGCT TAA-----AT
AGATCCAAGT GGGCCTGAAA GCAGCCACCT TTTAAGAAAG CGTTAAAGCT TAAATATTAC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
```

```
TTGTTTCAGT GGGCCTAAAA GCAGCCAACT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTCCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC
TTGTTTCAGT GGGCCTAAAA GCAGCCACCT GCACAGAAAG CGTTAAAGCT CAG-----AC

ATAA---ACC CTATATCCGG ATAAAGTACT CTAAATCCCC TATAAATATC AAGTTAATCT
AAAAAATACC CAATATCCAG ATATAAC-CT CTAAATCCCC TATAAATATC AAGTTATTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCTCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCTCATC CCATATCCCG ATAAAACGAT CACAATCCCC TAAACCTACA GAGCCGCTCT
AAACCTCATC CCATATCCCG ATAAAACGAT CACAATCCCC TAAACCTACA GAGCCGCTCT
AAACCTCATC CCATATCCCG ATAAAACGAT CACAATCCCC TAAACCTACA GAGCCGCTCT
AAACCCTACC CCATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCTACC CCATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCTACC CCATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCTCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCCCTCT
AAACCTCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCCCTCT
AAACCCAACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCCCTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCCCTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCTCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAACCCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCATC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACCCCACC CTATATCCCG ATAAAACAAT CACAATCCCC TAAACCTACA GAGCCACTCT
AAACTCCACC CTATATCCCG ATAAAACAAT CACAATCCCY TA-ACTTACT GAGCTACTCT
AAACCCCGTC CCGCATCCCG ATAAAATAAT CACAATCCCC TAGACCTACA GAGCTACTCT

ATTAAA--AT AGAAGAAACA ATGCTAGAAC TAGTAATAAG AAAA-ATGAT TTTC-TCCTA
ATCCCA--AT AGAAGAAATT ATGCTAGAAT TAGTAATAAG AAAA-ATGAT TTTC-TCCTA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAGCTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGTACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGTACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGTACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCGA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCGA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGTACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGTACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
```

```
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATATAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCACGAC CTTC-TCCAA
ATACAACTAT AGAAGCAATA ATGCTAAA-- TAGTAACANG AAGGCNCGAC CTTCTTCCCA
ATACAACTAT AGAAGCAATA ATGCTAAAAT TAGTAACAAG AAGGCAAGAC CTTC-TCCCA

GCACAAGTGT AAGTT-AGAA CGGACAAACC ACTAACAGTT AAACGAACCC ACCGGAGGAG
GCACAAGTGT AAGTT-AGAA CGGACAAACC ACTAACATTT AAACGAACCC TCCAGAGGAT
GCACACGTGT AAGTT-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGCC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACATGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTT-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTT-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAC CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAC CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAC CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAC CGGACCCTCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAGT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAGT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGCC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGCC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGGCCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCACC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACACGTGT AAGTC-AGAT CGGACCCGCC ACTGACAAAT AACGGACCCA ACCAAAGAGG
GCACATGTGT AAGTCAAGAT CGGACTCACC GCTGNCAAAT AACGGACCC- ACCAAAGAGG
GCACATGTGT AAGTC-AGAT CGGACTCACC ACTGACACAT AACGGACCCA ACCAAAGAGG

CAAATAGTGC ATA-TAATAA AAA-CAAGAA AACCCTAT-T AATTCTATCG TAAATCTTAC
CAAATAGCGT ATA-TAATAA AAA-CAAGAA AATCCTAT-T AACCTTATCG TTAATCTTAC
GAAGTACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGT ATAATAATAG AAGTCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGT ATAATAATAG AAGTCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGT ATAATAATAG AAGTCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACGACCG TTAACCCAAC
GAAATACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACGACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AATCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AATCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAGTACAGA ATAGTAATAG AAATCAAGAA AACCCTGTAA AATACTACCG TTAACCCAAC
GAAATACAGA ATAATAACAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAGTACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAGTACAGA ATAATAATAG AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAGTACAGA ATAATAATAA AAATCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC
GAAATNCRGA ATACCAATGA AACTCAAGAA AA-CCTGTAA AACACAACCG TTAATCCCAC
GAAATACAAA ATAATAATGA AACTCAAGAA AACCCTGTAA AACACAACCG TTAACCCAAC

ACAAGAGTGC CCA--AAGGA AAGACTAAAA GAGAAAAAAG GAACTCGGCA AACCCGAGCC
ACAAGAGTGC CCA--AAGGA AAGACTAAAA GAGAAAAAAG GAACTCGGCA AACTCGAACC
```

```
ACAGGAGTGC GCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACTCCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC GCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC GCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC GCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACGCCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACGCCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACGCCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACGCCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACATCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACATCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACATCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACATCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC ACATCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC TCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC
ACAGGAGTGC TCACCAAGGA AAGACTAAAA GAAAAAGAAG GAACTCGGCA AACACGAGCC

CCGCCTGTTT ACCAAAAACA TCGCCTTCAG CTTTCC-ATG TATTGAAGGT CCTGCCTGCC
CCGCCTGTTT ACCAAAAACA TCGCCTTCAG CTTTTC-ATG TATTGAAGGT CCTGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAACCAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAA-CAATG TATTAGAGGT CCCGCCTGCC
TCGCCTGTTT ACCAAAAACA TCGCCTCTTG CAAA-CAATG TATTAGAGGT CCCGCCTGCC

CAGTGA-CAT GAGTTTAACG GCCGCGGTAT CCTGACCGTG CAAAGGTAGC GTAATCACTT
CAGTGA-CAC AAGTTTAACG GCCGCGGTAA CCTGACCGTG CAAAGGTAGC ATAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
```

```
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT
CTGTGACCAA AAGTTTAACG GCCGCGGTAT TTTGACCGTG CGAAGGTAGC GTAATCACTT

GTCCTTTAAA TGGGGACTGG TATGAATGGC CCCACGAGGG CTCAACTGTC TCTTTTCTCC
GTCCTTTAAA TGGGGACTGG TATGAATGGC CCCACGAGGG TTCAACTGTC TCTTTTCTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATTACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATTACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATTACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG TATGAATGGC ATAACGAGGG CTCGACTGTC TCCTTTTTCC
GTCTTTTAAA TGAAGACCTG T-TGAATGGC ATAACGAGGG CTCAACTGTC TCCTTTTTCC

CGTCAATTAA ACTGATCTAC CCGTGCAGAA GCGGGTATAA AAACATAAGA CGAGAAGACC
AGTCAATTAA ATTGATCTAC CCGTGCAGAA GCGGGCATAA AACCATAAGA CGAGAAGACC
AGTCAGNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA TCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA ACCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA ACCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA ACCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA TCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA TCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA TCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA TCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCACATAAGA CGAGAAGACC
AGTCAGTGAA ATTGACCTGC TCGTGCAGAG GCGAGCATAA CCCCATAAGA CGAGAAGACC

CTGTGGAGCT TTAGACTAAA TCCAAATGA- -CTCCTAACT ATACCATGCC GTATAGATAA
```

```
CTGTGGAGCT TTAGACTAAA TCCAAACGA- -CTCCCCACT ATACCCTACC ATATAAATAA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNN-NNNNN NNNNNNNNNN
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCT ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCT ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCT ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CTAT-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CTAT-CGCCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CTAT-CGCCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCCATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCCATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-TACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-TACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAGATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC GCGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CACC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAAC-CACCC ACGGAAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAACACACCA ACAGGAATAA
CTATGGAGCT TAAAACACAA GATCAACTAT GCTATCAAGC CAACACGCCA ACAGGAATAA

---------A CACAGCATTA -----TGGCC ATAAAGTCTT AGGTTGGGGC GACCACTGAG
--ACT----- -----TAGCG --TTATGGCC ATACAGTCTT AGGTTGGGGC GACCACTGAG
NNNNNNNNNN NNNNNNNNNN --NNNNGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CACAATAGTA --CTATGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CACAATAGTA --CTATGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CACAATAGTA --CTATGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGCA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CCAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CGTAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTAAAAG CATAATAGTA --CCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
CAGCTATAAG CATAATAGTA -CCTCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG
TAGCTAAAAG CATAATAGTC ACCCCTGATC CTAATGTTTT CGGTTGGGGC GACCACGGAG

AACAAGCAAT CCTCAGCGAT GATTGAAGCA CAGCTTTATA AACTAAGAAT GACAATTCGA
AACAAAATAT CCTCAGCGAC GATTGAAGCA CAGCTTTATA AGCTAAGAAT GACAATTCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
```

```
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAATAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AACCTAGGGC GACAGCCCAA
GACAAAAAAG CCTCCATGTC GACGGGGGCA CTGCCCCTAA AGCCTAGGGC GACAGCCCAA

CGCATCAGGA CATCTGAC-- ATTAAGATCC AGACTAA--T CTGATCAACG AACCAAGTTA
AGCATCAGGA CATCTGAC-- ATTAAGATCC AGACTAA--T CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCCAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCCAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCCAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ATAATGACCC AGGCTAAAAC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAGCAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCTAAAGC CTGATCAACG AACCAAGTTA
AGCAACAGAA TATCTGACGA ACAATGATCC AGGCCCA-GC CTGATCAACG AACCAAGTTA
AGCAACAGAA CATCTGACGA ACAATGACCC AGGCCTA-GC CTGATCAACG AACCAAGTTA

CCCCAGGGAT AACAGCGCAA TCTTTTCCAA GAGCCCAAAT CGACGAAAAG GTTTACGACC
CCCCAGGGAT AACAGCGCAA TCTTTTCCAA GAGCCCAAAT CGACGAAAAG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTGA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTGA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTGA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAAGG GTTTACGACC
CCCTAGGGAT AACAGCGCAA TCCTTTCTAA GAGTCCATAT CGACGAAGGG GTTTACGACC
```

```
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG TAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAT CCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC
TCGATGTTGG ATCAGGACAA TCTAATGGTG CAGCCGCTAT TAAGGGTTCG TTTGTTCAAC

GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCGTACTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCGTACTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCGTACTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCGTACTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCGTACTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT
GATTAAAGTC CTACGTGATC TGAGTTCAGA CCGGAGTAAT CCAGGTCAGT TTCTATCTAT

GAAGTTATTT ATCCTAGTAC GAAAGGATTG GATAAATGAG GCCTATATTT AAGACACGCC
GAAGTTATTT ATCCTAGTAC GAAAGGATTG GATAAATGAG GCCTATATTA AAAACATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTACGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG AAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG AAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG AAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA TAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA TAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA TAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA CAAGTATGCC
```

```
GCAGTGACCC TTCCTAGTAC GAAAGGGCCG AAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG AAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGACTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGACTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGACTGAG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTATGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATACTA CAAGCACGCC
GCAGTGACCC TTCCTAGTAC GAAAGGACCG GAAGGCTGGG GCCAATGCTA CAAGTACGCC

TCTCTTTAAC CTACTGAAGC CAAATCAAGT AGATAATAAA GAACGCACAC CACACCCTAG
TCCCCTTAAC CTACTGAAAC CAAATCAAGT AGATAATAAT GAACAAAATT TTCACCCAAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
ACACCCCAAC CTAATGAAGA CAACTAAAAT AGGTAAAGGG GCACAAATCC TCCCCCCTAG
ACACCCCAAC CTAATGAAGA CAACTAAAAT AGGTAAAGGG GCACAAATCC TCCCCCCTAG
ACACCCCAAC CTAATGAAGA CAACTAAAAT AGGTAAAGGG GCACAAATCC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAC AGGTAAAGGG GCACGAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAC AGGTAAAGGG GCACGAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAA
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAA
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAGTGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
CTACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TTCCCCCCAG
TCACCCCGAC CTAATGAAAA CAACTAAAAC AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAG
TCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG ACACAAA-CC TCCCCCCTAG
CCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TTTCCCCTAA
CCACCCCAAC CTAATGAAAA CAACTAAAAT AGGTAAAGGG GCACAAA-CC TCCCCCCTAA

AACAGGGCTA GCTAGTGTAG CTTAACCTCA AAGCATAACA CTGAAAATGT TAAGATGGAC
ATCAGGGTCA GCTAGTGTAG CTTAACCTCA AAGCATAACA CTGAAAATGT TAAGATGGAC
AATAGGGCA- ---------- ---------- ---------- ---------- ----------
AATAGGGCA- ---------- ---------- ---------- ---------- ----------
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGGGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGACGGGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGGGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AACAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGAGC
AATAGGGCA- ---------- ---------- ---------- ---------- ----------
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGGGT
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAGGATGGGT
AATAGGGCAA GCTAGCGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGGGT
AACAGGGCAA GCTAGTGTAG CTTAACT--A AAGCATAACA CTGAAGATGT TAAGATGGAC
AACAGGACAA GCTAGCGTAG CTTAACTT-A AAGCATAACA CTGAAGATGT TAAGATAGAC
```

```
CCTAGAAAAG TCCCACAAGC ACAAAATGTA GCTTAATAGA AAAGTATCTC CCTTACACTG
CCTAGAAAAG TCCCACAAGC ACAGAATGTA GCTTAATAGA AAAGTATCTC CCTTACACTG
---------- ---------- ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
---------- ---------- ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAGATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAGATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CTTAG-GCAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CTTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAA-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAA-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAA-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG -TCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAG CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
---------- ---------- ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAA CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAA CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-ACAA CTCCGCAGGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-AAAG TCCCACAAGC ACAAAATGTG GCTCAATAGA AAAGCACCTC CCTTACACCG
CCTAG-AAAG TCCCGCAAGC ACAAAATGTA GCTCAATAGA AAAGCACCTC CCTTACACTG

AGATGAAATC TGCGCAATTC AGATCATTTT GAGTTAGAGT GGCAGAGCCT GGTAATTGCA
AGATGAAGTC TGCGCAATTC AGACCATTTT GAGTTAGAGT GGCAGAGCCC GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTGAGAT GGCAGAGCTT GGTAATCGAA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGCAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGCAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGCAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATCTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTGAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTGAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTGAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGGGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTGAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGGGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGAGACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAATTGCA
AGGAAACATC TGTGCAAATC AGATCATTTT GAGCTAAGAT GGCAGAGCTT GGTAACTGCA

TAAGGCCTAA GCCCTTACCC CCAGGGGTTC AACTCCCCTC TCTAACT--- -CTATGCCCG
TAAGGCCTAA GCCCTTACTC CCAGGGGTTC GACTCCCCTT TCTAACT--- -CTATGCCCG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA AAAGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ACCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ACCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ACCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
```

```
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTC-A ACAGAGGTTC AAGTCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTC-A ACAGAGGTTC AAGTCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTCAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAGTCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ATCGTGCCTG
AAAGGCCTAA GCCCTTTCCA ACAGAGGTTC AAATCCTCTT CTTAGCTGGA ACCGTGCCTG
AAAGGCCTAA GCCCTTTCCT ACAGAGGTTC AAATCCCCTT CTTAGCTGGA ACCGTGCCTG

AAAGTTAGGG GCTACTTTGA TAGAGTAGAA AATAGGGGTT CAAACCCCCT TAGTGCCT-A
AAAGCTAAGG ACTACTTTGA TAGAGTAGAA AATAGGGGTT CAAACCCCCT TAGTGCCTAG
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAACG GCCACTTTGA TAGAGTGGAT AATGGGGGTT CAAGTCCCCT CGCTTCCTTA
AAAGTCAAGG GCCACTTTGA TAGAGTGAAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAAGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAAGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAGGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATGGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGATCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGCCAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA
AAGGTTAAGG GCCACTTTGA TAGAGTGGAT AATAGGGGTT CAAGTCCCCT CGCTTCCTTA

GAAAAGAAGG ACTCGAACCT CCACCAGAAA GATCAAAACT TTCAGTGCTC CCATTACACC
GAAAGAA-GG ATTCGAACCT CCACTAAAAA GATCAAAACT TTTAGTGCTC CCATTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GACCAAAACT CTTGGTGCTT CAACTACACC
GAAAGAAGGG ACTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG ACTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG ACTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CCTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAGAAGGG GTTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
GAAAAAAGGG ATTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC
```

```
GAAAGAAGGG GCTCGAACCC ATCCTCAAGA GATCAAAACT CTTGGTGCTT CCACTACACC

ACTTCCTAGT AAGGTCAGAT GATTCAAGCT TTTGGGCCCA TACCCCAAAT ATGTTGGTTA
ACTTCCTAGT AAGGTCAGCT AATTTAAGCT TTTGGGCCCA GACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACATCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGGTA
ACCTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACCTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACCTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA
ACTTCCTAGT AAAGTCAGCT AATTAAAGCT TTTGGGCCCA TACCCCAAAC ATGTTGGTTA

AAATCCTTCC CTTACTAAGG AGTTTAGATT AACT---TAA ACCAAGAGCC TTCAAAGCTT
AAATCCTTCC CATACTAAGG AGTTTAGATT AACT---TAA ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAAGCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AACAA--CAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AACAA--CAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AACAA--CAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AACAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GACTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GACTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTACTAAGG GGCTTAGGAT AGCAATCTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTATTAAGG GGCTTAGGAT AA-GCCTTAG ACCAAAAGCC TTCAAAGCTT
AAATCCTTCC TTTATTAAGG GGCTTAGGAT A--AACCTAG ACCAAAAGCC TTCAAAGCTT

TAGACAGAGG TTAAAATCCT CTAACTCCTG -TATTAGGAT TTGCAGGACT TTATCCCACA
TAGACAGAGG TTAAAATCCT CTAACTCCTG CTTATAGGAT TTGCAGGACT TTATCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTACAGGATA TTACCCCACA
TAAGCAGGAG TGAAAACCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA CTACCCCACA
TAAGCAGGAG TGAAAACCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA CTACCCCACA
TAAGCAGGAG TGAAAACCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA CTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC CTGCAGGATA CTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -C--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -C--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGTCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
```

```
TAAGCAGGAG TGAGAATCTC CTAGTCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -T--TAAGAC TTGCGGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAGAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAT TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGACA TTACCCCACA
TAAGCAGGAG TGAAAATCTC CTAGCCCCTG -T--TAAGAC TTGCAGGATA CTACCCCACA

---------- ---------- ---------- ---------- ---------- ----------
TCTCCTGAAT GGAAATCAAA CACTTTAATT AAGCTAAAAC C-CTACTAGG AGGGTAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTGCTGGGT CCAACACTGA TACTTTAATC AATCTACAGC C-TCACTAGA TGAGACGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTGCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTGCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTGCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTGCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTTCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC C-TTACTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC CTTTCCTAGA TGAGAAGGCC
TCTTCTGAAT GCAACCCAGA TACTTTAATT AAGCTAAAGC CTTTTCTAGA TGAGAAGGCC

-------TAC A----TTTAG TTAACAGCTA AAAGCTTTAA CCGGCAAGCT TCCACCTAGA
TCAATCCTAC ATACATTTAG TTAACAGCTA AAAGCTTTAA CCGGCAAGCT TCCACCTAGA
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGCGCCCCAT CCAGCGAGCA TTCATCT--A
TCAATCCTTC AAAATCTTAG TTAACAGCTA AGCGCCCCAA CCACCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAACGGGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAACGGGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAACGGGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
```

```
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCT--A
TCGATCCTAC AAAATCTTAG TTAACAGCTA AGTGCCCTAT CCAGCGAGCA TTCATCTA-A

GCCCCGGCAG GCG-TTAACC TACGTCTCAG GGTTTGCAAT CCTGCATGAA ACTACACTAC
GCCCCGGCAG GCG-TTAACC TACGTCTCAG GATTTGCAAT CCTGCATGTA GCTACACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCTCGGCAG GCGGCGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G TCTTCACCAC
GTCTCGGCAG GCGGCGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G TCTTCACCAC
GTCTCGGCAG GCGGCGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G TCTTCACCAC
GTCCCGGCAG GCGGCGAGCC TGCGTCTTCA GGTTTGCAAT CTGACGTG-A TCTTCACCAC
GCCCCGGCAG GCGGCGAGCC TGCGTCTTCA GATTTGCAAT CTGATGTG-A TCTTCACCAC
GCCCCGGCAG GCGGCGAGCC TGCGTCTTCA GATTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGATGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGATGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGGCGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCAACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGACGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-A TCTTCACCAC
GTCCCGGCAG GCGGCGAGCC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G CCTCCACCAC
GTCCCGGCAG GCGACGAACC CGCGTCTTCA GGTTTGCAAT CTGATGTG-G TCTTCACCAC

AAGGCTTGGT AAAAGAAGGA ATTAAACCTC CCTTTACAGG GCTACAACCT GCCGCTTAAA
AAGGCTTGAT AAGAGAAGGA ATTAAACCTT CATTTACAGG GCTACAGCCT GCCGCTTAAA
GGGACTTGAT AAGAAGGGGA TTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GAGACTTGAT AAGAAGGGGA CTTTAACCTC TGTACATGGG GCTACAACCC ACCGCTTAAA
GAGACTTGAT AAGAAGGGGA CTTTAACCTC TGTACATGGG GCTACAACCC ACCGCTTAAA
GAGACTTGAT AAGAAGGGGA CTTTAACCTC TGTACATGGG GCTACAACCC ACCGCTTAAA
GAGACTTGAT AAGAAGGGGA CTTTAACCTC TGTACATGGG GCTACAACCC ACCGCTTAAA
GGAACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGAACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGAGCTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
AGGACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
AGGACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
AGGACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
AGGACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
AGGACTTGAT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGAT AAGAAGGGGA CTCTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGAT AAGAAGGGGA CTCTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GAGACTTGAT AAGAAGGGGA CTTTAACCTC TGTACATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGGACTTGGT AAGAAGGGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTAAA
GGAACTTGGT AAGAAGAGGA CTTTAACCTC TGTGCATGGG GCTACAATCC ACCGCTTAAA
GGGACTTGAT AAGAAGAGGA CTTTAACCTC TGTGCATGGG GCTACAACCC ACCGCTTGAA

CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATTGAACCC CCTTAAATCG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATTGAACCC CCTTAAACCG ATTTCAAGCC
CACTCAGCCA TCTTACCTNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCCAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
```

```
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCGGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
CGCTCAGCCA TCTTACCTGA GAAAGGAAGG AATCGAACCC CCATTTGCTG GTTTCAAGCC
TACTCAGCCA TCTTACCTGA GAAAGGAAGG AATTGAACCC CCATTTACTG GTTTCAAGCC
CACTCAGCCA TCTTACCTGA GAAAGGAAGG AATTGAACCC CCATTTGCTG GTTTCAAGCC

GACTACATAA CCACTCTGCC ACTTTCTT-- ----ATAAGA TTTTAGTTAA A---ATATAA
GGCTGCATAA CCACTCTGCC ACTTTCTT-- ----ATAAGA TATTAGTTAA A---CTATAA
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCACATAA CCGCTCTGCC ACTTTCTTAT CCCCATGAGA CACTAGTAAA ATCACTATGA
AGCCACATAA CCGCTCTGCC ACTTTCTTAT CCCCATGAGA CACTAGTAAA ATCACTATGA
AGCCACATAA CCGCTCTGCC ACTTTCTTAT CCCCATGAGA CACTAGTAAA ATCACTATGA
AGCCGCATAA CCGCTCTGCC ACTTTCTTTC CCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTTT CCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTTT CCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGG CACTAGTAAA ATTGATATAA
AACCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AACCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGATATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT CCCCATGAGG CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCGCTCTGCC ACTTTCTTAT TCCCATGAGA CACTAGTAAA ATTGCTATAA
AGCCGCATAA CCACTCTGCC ACTTTCTTAT CCCAATGAGA CACTAGTAAA ATTGCTATTA
AACCGCATAA CCGCTCTGCC ACTTTCTTAT CCCCATGAGA CACTAGTAAA ACTGTTATTA

CACTGCCTTG TCAAGGCAAA ATTACGGGTT AAACCCCCGT ATATCTTACC T---------
CTCTGCCTTG TCAAGGCAAA ATTATGGGTT AAATTCCCAT ATATCTTACC T---------
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTG GAAGGCCTGC GTGTCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTG GAAGGCCTGC GTGTCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTG GAAGGCCTGC GTGTCTTAAG TCCAAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCGAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCGAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCGAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCAAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCTAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTT AAAGACCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA GTTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
CACTGCCTTG TCAAGGCAGA ATTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
```

```
CACTGCCTTG TCAAGGCAGA ATTGCAGGTT AAAGGCCTGC GTGCCTTAAG TCCAAGGACT
CATTGCCTTG TCGAGGCAAA ATTGCAGGTT AAAGCCCTGC GTGCCTTAAG TCCATGGACT
CATTGCCTTG TCAAGGCAAA ATTGCAGGTT GAAACCCTGC GTGCCTTAAG TCCGAGGACT

--CACTAAGA AGCT-AATAG GATATAGCAT T-AGCCTTTT AAGCTAAACA CAGGTGACTG
--CATTAAGA AGCT-AATAG GGTATAGCAT T-AGCCTTTT AAGCTAAAAA TAGGTGACTC
NNCACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCTTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACTAAGA AGCTAAATAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC
AACACCAAGA AGCTAAATAG GGACTAGCGT TAAGCCTTTT AAGCTAAAGA TTGGTGGCCC
AACACTAAGA AGCTAAACAG GGAATAGCGT T-AGCCTTTT AAGCTAAAGA CTGGTGACCC

CCGACCACCC TTAGTGAATC TTTCTAGTAT TAAAA---AG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAATGAATC TTTCTAGTAT TAAAA---AG TACAAATGAC TTCCAATTAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC ATTCTAGCAT CAACATACAG AACAAATAAC CTCCAATCAA
CCAACCACCC TTGGTGAACC TTTCTAGTAT TAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTGGTGAACC TTTCTAGTAT TAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTGGTGAACC TTTCTAGTAT TAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT TAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT TAACATTCAG TACAAATGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT TAACATTCAG TACAAATGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAATGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAATGAC TTCCAATCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT CAACATTCAG TACAAGTGAC TTCCATTCAT
CCAACCACCC TTAGTGAACC TTTCTAGTAT TAGCACTCAG TACAAGTGAC TTCCAATCAT
CCAACCACCC TTAGTGAGCC TTTCTAGTAT TAACATTCAG TACAAATGAC TTCCAATCAT

TTAGTCTTGG TTAAATTCCA AGGAAAGATA AAGTCCCTAG TCCAACGTAA GATTATTGAT
TTAGTCTTGG TTAAAACCCA AGGAAAGATA AAGTCCCTAG TCCAACGTAA GATTATTGAT
TTAGCCTTGG TTAGAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG CTAAAATCCA AGGAAAGGTA AGGTGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGACGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGACGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
```

```
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGATGACTAG TCCAAAGTAA GACCGCTGAT
TAAGCCTTGG TTAAAATCCA AGGAAAGGTA AGACGACTAG TCCAAAGTAA GACCGCTGAT
TTAGCCTTGG TTAAAATCCA AGGAAAGGTA AGACGACTAG TCCAAAGTAA GACCGCTGAT

TTCGGCTCAA TAGATTGTGG CTCAAATCCA CAGAGACTTA GTTAATATAG TCTAACTAAG
TTCGGCTCAA TAGATTGTGG CTCAAATCCA CAGAGACTTA GTTAATATAG TCTAAACAAG
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG AGTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TCGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAGCATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAGCATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAGCATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAGCATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAGTCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA
TTCGGCTCAA CTAATTATGG TGCAAATCCA TAGTCGTCTT GTAAACATAG TTTAAACAAA

ACATTAGATT GTGATTCTAA TAATAGAAGT TAAATTCTTC TTATTGACCG AGAGATGCC-
ACATTAGATT GTGATTCTAA TAATAGAAGT TAAACTCTTC TTATTGACCG AGAGATGCC-
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
AGATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGT TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGT TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGT TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
```

ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCG
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAAACCCTCC TTGTTCACCG AGAGAAGCCA
ATATTAGATT GTGGTTCTAA AGATGGGAGC TAGACCCTCC TTGTTCACCG AGAGAAGCCG

TGCGCACCAA GAACTGCTAA TTCTT-ACTA CTTTGGTTAA ACTCCAAAGA TCACTCAATG
TGCGCACTAA GAACTGCTAA TTCTT-ATTA CCTTGGTTAA ACTCCAAGGA TCACTCACCG
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATAGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGAA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGAA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTAAGCA CCATGGTTCA AGTCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCGTGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G
GGGGCACTAA GAACTGCTAA TTCTTCAGCA CCATGGTTCA AATCCATGGG TCACTCG--G

CTTTCAAAGG ATAATAGTT- ATCCATTGGT CTTAGGAACC AAAAACTCTT GGTGCAAATC
CTTTCAAAGG ATAATAGTT- ATCCATTGGT CTTAGGAACC AAAAACTCTT GGTGCAAATC
CTTTTAAAGG ATAATAGTTC ATCCATTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTCC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTCC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCATTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGTTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTCTAAAGG ATAATAGCTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC
CTTTTAAAGG ATAATAGCTC ATCCGTTGGT CTTAGGAACC AAAAACTCTT GGTGCAACTC

CAAGTGAAAG CTCATAATTC TTGCCAGGAC TCTAACCAGG GCTAATGACT TGAAAAATCA
CAAGTGAAAG CTCATAATTC TTGCCAGGAT TTTAACCAGG ACTAATGACT TGAAAAATCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA

```
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAGTGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAGTGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG ACCATAAATC TTACTCGGAC TCAAACCAAG ACCAGTGACA TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA
CAAGTAGAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACTAATGACT TGAAAAACCA
CAAGTAAAAG CTCATAATTT TTACTCGGAC TCTAACCAAG ACCAATGACT TGAAAAACCA

CCGTTGTT-A CTCAACTACA AAAATCAAAT GGCAATCATA CGTAAAACCC ACCCATTAGC
CCGTTG-TAA CTCAACTACA AAAACTGCAT GGCAATCTTA CGCAAAACCC ACCCACTAAC
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCGCTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCGCTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACTA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACTA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGAAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AC GGCAAACATC CGAAAAACAC ACCCGCTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGGAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGGAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCAAACATC CGGAAAACAC ACCCACTACT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCTAACATC CGAAAAACAC ACCCCCTTCT
CCGTTGTTAA TTCAACTATA AAAACCA-AT GGCTAACATC CGAAAAACGC ACCCACTTCT

AAAAATCATC AACAGTGCAT TCATCGACCT GCCAGCCCCA TCAAACATCT CATCATGGCC
AAAAATTATC AACAGCGCAT TCATCGACCT GCCTGCCCCA TCCAATATTT CTTCATGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CACCACACCC TCCAACATCT CCGTGTGTCC
TAAAATTATT AATGGAGCAC TAATCGACCT CCCCACACCC TCTAATATCT CAGTGTGGCC
TAAAATTATT AATGGAGCAC TAATCGACCT CCCCACACCC TCTAATATCT CAGTGTGGCC
TAAAATTATT AATGGAGCAC TAATCGACCT CCCCACACCC TCTAATATCT CAGTGTGGCC
TAAAATTATT AATGGAGCAT TCATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATC AATGGAGCAT TCATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATC AATGGAGCAT TCATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATCATT AATGGAGCAT TTATCGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAC TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAC TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATC AATGGAGCAT TCATTGACCT CCCCACACCC TCCAACATCT CCGTGTG---
```

```
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGACCT CCCCACACCC TCCAACATCT CCGTGTGGCC
TAAAATTATT AATGGAGCAT TTATTGATCT TCCAGCACCC TCTAACATCT CCGTATGGCC
TAAAATTATC AATGGAGCAT TCATTGATCT CCCAGCACCC TCCAACATCT CCGTGTGGCC

CCAGTAACTT AAA-GCTAAA GTATCGGTCT TGTAAGCCGA ACAT-GACGG TTAAATTCCA
TAAGTAACTT AAA-GTTAAA GTGTCGGTCT TGTAAACCGA ATAT-GACGG CTAAATTCCA
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTC AGACATAAAA GCCCCGATCT TGGAAACCGA AGATCATAGG TTAAAACCCT
CTAGTAGCTT AGACATTAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATTAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATTAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTC AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTC AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAGAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
---------- ---------- ---------- ---------- ---------- ----------
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATCAAA GCACCGGTCT TGTAAACCGA AGACCGAAGG TTAAAATCCT
CTAGTAGCTT AGACATTAAA GCACCGGTCT TGTAAACCGA AGATCGAAGG TTAAAATCCT

TCCTTGGGCT CAAAGAGAAG AGACTCTAAC TCCTGCCTTT AACTCCCAAA GCTAAAATTC
TCCTTAGGCT CAAAGAAAAG AGACTTTAAC TCCTACCTTT AACTCCCAAA GCTAAAATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCCAARATTC
TCGTACCAAC CAGAGAAAAG AGAATGCAAC CCTAAACCTT AACTCCCAAA GCGAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AGCTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AGCTCCCAAA GCTAAGATTC
TCCTAGCGCC CAAAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAGTTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAGTTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
---------- ---------- ---------- ---------- ---------- ----------
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC
TCCTAGCGCC CAGAGAAAAG AGAATTCAAC TCTCACCCTT AACTCCCAAA GCTAAGATTC

TAC-TTAAAC TACTCTTTGA TATATTCCTC ATTAATTT-- ATCCATAACA TAAAAGTATT
TAC-TTAAAC TACTCTTTGG TATATTTTAC ATTAAATT-- ATCTACAATC CAACAGAATT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCC AGTCACTACA CCGTAGTATT
TACATTAAAC TATCCTCTGA TTTTCTCCTC ATTAACCTCT AACCACTATA ACGTAGTATT
TACATTAAAC TATCCTCTGA TTTAATCCAC ATTAATTTCT AGTCAACATA TCGTAGTACT
TACATTAAAC TATCCTCTGA TTTAATCCAC ATTAATTTCT AGTCAACATA TCGTAGTACT
TACATTAAAC TATCCTCTGA TTTAATCCAC ATTAATTTCT AGTCAACATA TCGTAGTACT
```

```
TACATTAAAC TATTCTCTGA TTTAATCCCC ATTAATTTCT AGTCAACATA TCGTAGCATT
TACATTAAAC TATTCTCTGG TTTTATCCTC ATTAACTTCT AGTCACTATA ACGTAGTATT
TACATTAAAC TATTCTCTGG TTTTATCCCC ATTAACTTCT AGCCACTATA ACGTAGTATT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGCCACCATA TTACAGAAAC
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGCCACCATA ACGTAGTATT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGTCACCATA TCGTAGAGAT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGTCACTATA TCGTAGAGAT
TACATTAAAC TATTCTCTGA TTTAATCCCC ATTAATTTCT AGTCACCATA TCGTAGAGAT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATCTCT AGTCACCATA CCGTAGAGGT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATCTCT AGTCACCATA CCGTAGAGGT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGCCACCATA ATGTAGTTGT
TACATTAAAC TATTCTCTGA TATAATCCAC ATTAATTTCT AGCCACCATA CCGTAGTTGT
TACATTAAAC TATTCTCTGG TTTAATCCAC ATTAACTTCT AGCCACCATA TCGTAGTAAT
---------- ---------- TTTTATCCCC ATTAACTTCT AGCCACTATA ACGTAG----
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGCCACCATA CCGTAGTTGT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGTCATCATA CCGTAGTTGT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTCT AGTCACCATA CCGTAGTTGT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTAATTTTT ACACAACATA --AAAGATAT
TACATTAAAC TATTCTCTGA TTTAATCCAC ATTACTT--- -------GTA CAAAAGTAAA

CAGCTAGAAT ATATATGACT TGAACAATCA TTAACATTTG GTTCCTGTTT CAGGTACATA
AAGCCAGAAT ATATATGACT TGAACAATCA TTAGCATTTG GTTCCTAACT CAAGTACACA
CTGTCCAAAT ATAGGGTAAC TGAGCTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATA
CTTTTTAAAC ATAAAATAAG TGAACTATTA TTGGCATCTG GTTTATATCT CAGGCACATA
TTACCCAAAC ATGACACACC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATA
TTACCCAAAC ATGACACACC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATA
TTACCCAAAC ATGACACACC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATA
TGTCTTAAAC ATAAGGTAAC TGAACTATTA CTAACATCTG GTTCCTATTT CAGGTACATA
CTTTTTAAAC ATAAACTAAT TGAACTATTA CTGGCATCTG GTTTGTATCT CAGGCACATG
CTTTTTAAAC ATAAACTAAT TGAACTATTA CTGGCATCTG GTTTATATCT CAGGCACATA
ATATCAAAAC ACAAGATAAC TGAACTATTA CTGGCATCTG GCTTCTATCT CAGGTCCATA
CTTTTTAAAC ACAAGATAAT TGAACTATTA CTGGCATTTG GCTTCTATCT CAGGCCCATA
ATGTCTAAAC ATAAGATAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATG
ATGTCTAAAC ATAAAGTAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATA
CTGTCTGAAC ATAAGGTAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATA
CTGTCTAAAC ATTAAGTAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATG
CTGTCTAAAC ATTAGGTAAC TGAACTATTA CTGGCATCTG GCTCCTATTT CAGGTCCATG
CTGTCTAAAC ATGGGATAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATG
CTGTCTAAAC ATGAAGTAAC TGAATTATTA CTAACATCTG GCTCCTATTT CAAGTCCATA
ATGTCTAAAC ATAAAATAAC TGAATTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATA
---------- ---------- ---------- ---------- ---------- ----------
CTGTCTGAAC ATAGAATAAC TGAATTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATA
CTGTCTAAAC ATAGAATAAC TGAATTATTA CTGGCATCTG GTTCCTATTT CAGGCCCATA
CTGTCTAAAC ATAAGGTAAC TGAACTATTA CTGGCATCTG GTTCCTATTT CAGGTCCATA
TTCGTTAAAT ATAAAGTGAA TGAACTATTA CTGGC-TCTG GGT-CTATTT CA-GTCCATA
TTTTAGTAAC ATAAAATAAA TGAACTATTA CTGGCATCTG GTTCCTATTT CAAGTCCATA

ATAGATTATA CCCCATAACT GA
ATAGAATATT CCCCATAACT GA
GTGGCTACA- CCCCATAACT AG
ACAGTTCTGT CCCCAGAATT GG
ACAGTCTCCT CCCCATGACC AG
ACAGTCTCCT CCCCATGACC AG
ACAGTCTCCT CCCCATGACC AG
ACAGTCTTTT CCCCATAACT GA
ACAGCTCTCT CCCCATAACT GA
ACAGTTCTGT CCCCATAATT GG
ACAGCTAATT CCCCATAACT GA
ATAGTTAACT CCCCATAACT GA
ACGGCTATTT CCCCATAACT GA
ACAGTTATTT CCCCATAACT GA
ATAGCTACCT CCCCATAACT GA
ATAGCTATTT CCCCATAACT GA
ATAGTTATTT CCCCATAACT GA
ATAGTTATCT CCCCATAATC AG
ACAGTTATTT CCCCATAATA AG
ATGATTA-TT CCCCATAACC AG
```

```
---------- ---------- --
ATAGTTGTAT CCCCATAACT AG
ACAGTTGTAT CCCCATAACT AG
ACAGTTGTAT CCCCATAATC AG
ATGGTTATTT CCACATAACT GA
ATAGTTTCTT CCCCATAACT GA
```