University of Massachusetts Medical School
eScholarship@UMMS

GSBS Dissertations and Theses Graduate School of Biomedical Sciences

2009-08-03

Regulation of DNA Replication Origins in Fission Yeast: A Dissertation

Naveen Kommajosyula
University of Massachusetts Medical School

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss

Part of the Amino Acids, Peptides, and Proteins Commons, Enzymes and Coenzymes Commons, Fungi Commons, and the Genetic Phenomena Commons

Repository Citation
Kommajosyula N. (2009). Regulation of DNA Replication Origins in Fission Yeast: A Dissertation. GSBS Dissertations and Theses. https://doi.org/10.13028/pkad-8f31. Retrieved from https://escholarship.umassmed.edu/gsbs_diss/436

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

REGULATION OF DNA REPLICATION ORIGINS IN FISSION YEAST

A Dissertation Presented

By

NAVEEN KOMMAJOSYULA

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

IN

BIOMEDICAL SCIENCES

AUGUST 03, 2009

REGULATION OF DNA REPLICATION ORIGINS IN FISSION YEAST

A Dissertation Presented By

Naveen Kommajosyula

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

____
Nicholas Rhind, Ph.D., Thesis Advisor

____
Michael Brodsky, Ph.D., Member of Committee

____
Craig Peterson, Ph.D., Member of Committee

____
Oliver Rando, Ph.D., Member of Committee

____
Lee Zou, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

____
Dannel McCollum, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the School

____
Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program
August 3, 2009

COPYRIGHT NOTICE

Portions of this dissertation are represented in the following publications.

Kommajosyula N, and Rhind N. "Cdc2 tyrosine phosphorylation is not required for the S-phase DNA damage checkpoint in fission yeast," Cell Cycle. 2006 Nov 1;5(21):2495-500.

Patel PK, **Kommajosyula N**, Rosebrock A, Bensimon A, Leatherwood J, Bechhoefer J, Rhind N. "The Hsk1(Cdc7) replication kinase regulates origin efficiency," Mol Biol Cell. 2008 Dec;19(12):5550-8.

ACKNOWLEDGEMENTS

Nick Rhind has been a great mentor to me. I came to Umass to join his lab and I have not regretted my decision a single day in the last 6 years. It is great to see his single minded enthusiasm for science. I have learnt a lot from him. Nick has always been supportive and enthusiastic towards each experiment. The greatest thing about Nick is the time he devotes to lab and how available he is for any help I need. I have truly enjoyed working under a very cool, kind and understanding mentor. Thank you for everything Nick.

Rhind lab has been the place for insane fun, inane chatter, random conversations and loud singing. To say that I have enjoyed my time in the lab is to put it very mildly. I have shared with Nick the DUDE, Chaitali and Mary numerous stresses, joys and laughter's. We have been there for each other and I could not ask for better labmates. I will miss being around them. Thanks for the great memories. Prasanta and Shankar, the post-docs in our lab have been great supporters of me. Dan and Jill have been loads of fun to work with. BMP is the best department to have been a part of. I have been extremely close to Munson lab especially Mary Munson, Melonnie, Dan and Jen. I am glad to have known Jen and I will always miss her. Rando lab has been responsible for adding more craziness to our side of the department and I am thankful for that. Many thanks are due to the labs whose help I took at various times during my Ph.D.

Mary and Mel have especially been great friends and they have taken care of me over the years. Luca, Moses, Sagar and Madhavi have been other members in

the department who have been great to me. Thanks to all the secretaries in the GSBS and BMP for being so wonderful and taking care of our needs. Thanks to Red Sox, Bad news bears and Worcester Speedsters for many memories while I was here.

Life in Worcester would not be the same without Samriddha, Srivatsan, Pranav, Nitya and Amit. They have been my friends and family here. I have counted on Sam to be there for me everyday and no one has been more important to me all these years then her. She has supported, encouraged, listened to me whenever I needed her. I am very thankful for having her as my friend. Srivatsan has been my other support system, one whom I can count upon to be on my side come what may. You guys will always be very special to me.

Sankar, Sridevi, Rahul, Rujuta and Deb have been my other friends who have always supported me and been part of many crazy parties. My friends outside Umass have always loved me and supported me no matter what. Polly, Amrita, Suman, Amar, Payal, Bhumi Shruti and other ACBR nuts. Thank you.

My parents and sister have been through a lot over the years that I have been here. It has been very hard to be here and not being able to provide them with the emotional support by being there. They have made sure that I am not worried about anything that happens and have given me their complete support, love and encouragement. I love you guys and this is for your dream of me becoming a doctor.

To Dad, Mom and Vandana

ABSTRACT

Cells need to complete DNA replication in a timely and error-free manner. To ensure that replication is completed efficiently and in a finite amount of time, cells regulate origin firing. To prevent any errors from being transmitted to the next generation, cells have the checkpoint mechanism.

The S-phase DNA damage slows replication to allow the cell to repair the damage. The mechanism of replication slowing by the checkpoint was not clear in fission yeast, *Schizosaccharomyces pombe,* at the start of my thesis. The downstream targets of the DNA damage checkpoint in fission yeast were also unclear. I worked on identifying the downstream targets for the checkpoint by studying if Cdc25, a phosphatase, is a target of the checkpoint.

Work from our lab has shown that origin firing is stochastic in fission yeast. Origins are also known to be inefficient. Inefficient origins firing stochastically would lead to large stretches of chromosome where no origins may fire randomly leading to long replication times, an issue called the random gap problem. However, cells do not take a long time to complete replication and the process of replication itself is efficient. I focused on understanding the mechanism by which cells complete replication and avoid the random gap problem by attempting to measure the firing efficiency of late origins.

Genome-wide origin studies in fission yeast have identified several hundred origins. However, the resolution of these studies can be improved upon.

I began a genome-wide origin mapping study using deep sequencing to identify origins at a greater resolution compared to the previous studies. We have extended our origin search to two other *Schizosaccharomyces* species- *S. octosporus* and *S. japonicus*. There have been no origin mapping studies on these fission yeasts and identifying origins in these species will advance the field of replication.

My thesis research shows that Cdc25 is not a target of the S-phase DNA damage checkpoint. I showed that DNA damage checkpoint does not target Cdc2-Y15 to slow replication. Based on my preliminary observation, origin firing might be inhibited by the DNA damage checkpoint as a way to slow replication. My efforts to measure the firing efficiency of a late replicating sequence were hindered by potentially unidentified inefficient origins firing at a low rate and replicating the region being studied. Studying the origin efficiency was maybe further complicated by neighboring origins being able to passively replicate the region. To identify origins in recently sequenced *Schizosaccharomyces* species, we initiated the genome-wide origin mapping. The mapping was also done on *S. pombe* to identify inefficient origins not mapped by other mapping studies. My work shows that deep sequencing can be used to map origins in other species and provides a powerful tool for origin studies.

TABLE OF CONTENTS

REGULATION OF DNA REPLICATION ORIGINS IN FISSION YEAST I

COPYRIGHT NOTICE III
ACKNOWLEDGEMENTS IV
ABSTRACT VII

TABLE OF TABLES XII

TABLE OF FIGURES XII

CHAPTER I INTRODUCTION- DNA REPLICATION AND ORIGIN EFFICIENCY 1

The process of DNA Replication 2
Formation of Pre-RC complex 3
Initiation of Replication 9
How is re-replication prevented? 12
Early studies on origin sites 19
Origins in fission yeast 21
Origin location influences timing of firing 25
What is Random gap problem? 26
DNA damage checkpoints and origins 27
How does S-phase DNA damage checkpoint work? 29
Origin regulation by checkpoints 35

CHAPTER II CDC2 TYROSINE PHOSPHORYLATION IS NOT REQUIRED FOR THE S-PHASE DNA DAMAGE CHECKPOINT IN FISSION YEAST 37

ABSTRACT 38
INTRODUCTION 39
MATERIALS AND METHODS 44
Yeast methods 44
Flow cytometry methods 44
Asynchronous Experiments 46
Synchronous Experiments 46
RESULTS 47
Over-expressing Cdc25 fails to override the S-phase DNA damage checkpoint 56
Inhibitory phosphorylation of Cdc2 is not required for the S-phase DNA damage checkpoint 57

Cdc25 is not required for the checkpoint 63
Cdc25Δ nmt1:Pyp3 cells have a partial slowing in the presence of damage 64
DISCUSSION...69

CHAPTER III MEASURE IF ORIGIN EFFICIENCY INCREASES DURING S PHASE....................................77

INTRODUCTION..78
Random gap problem 79
MATERIALS AND METHODS..83
Strain construction...84
Synchronization experiments..85
Two-dimensional gel electrophoresis 86
Micro-array design..86
DNA preparation and microarray experiment 86
Deep sequencing experiments 87
Flow cytometry: 87
RESULTS AND DISCUSSION 88
Origin efficiency at *AT2062*..88
Oligonucleotide arrays at *AT2062* 92
Using deep sequencing to look at AT2062 region 97
Origin efficiency at *AT3003*...105
Fork bypass in a recombination dependent manner..105
CONCLUSION 108

CHAPTER IV GENOME-WIDE ANALYSIS OF ORIGINS IN *SCHIZOSACCHAROMYCES* GROUP 109

INTRODUCTION..110
Different origin identification methods...110
Sequencing to identify origins..113
MATERIALS AND METHODS 115
Strain maintenance..115
G2 synchronization..115
Time-course experiment..116
Flow cytometry: 116
Deep sequencing experiments 116
Alignments..117
RESULTS AND DISCUSSION 118
Identifying *Schizosaccharomyces pombe* origins 118
Identifying *Schizosaccharomyces octosporus* origins..137

Identifying *Schizosaccharomyces japonicus* origins 138
CONCLUSIONS 144

CHAPTER V UNPUBLISHED DATA **146**

APPENDIX V.1. ORIGIN INHIBITION BY DNA DAMAGE CHECKPOINT 147
INTRODUCTION 147
MATERIALS AND METHODS 148
Two-Dimensional gel electrophoresis 148
RESULTS 149
CONCLUSIONS 149
APPENDIX V.2 USING MICROARRAYS TO MEASURE REPLICATION ORIGIN FIRING EFFICIENCY 151
INTRODUCTION 151
MATERIALS AND METHODS 152
Synchronization experiments 152
Micro-array design 152
DNA preparation and microarray experiment 153
RESULTS AND DISCUSSION 154
CONCLUSION 165
APPENDIX V.3. EFFECTS OF MODULATING DFP1 LEVELS 166
INTRODUCTION 166
RESULTS AND DISCUSSION 167
CONCLUSION 172

CHAPTER VI DISCUSSION AND FUTURE DIRECTIONS **175**

Mechanism of slowing and role of Cdc25 in damage checkpoint 176
Hsk1 as an alternate target 178
Origin efficiency studies 179
Ways of using RTS1 sites to study origin efficiency 182
Identifying replication origins 184
Potential uses of using sequencing to identify origins 186
SUMMARY 187

REFERENCES **188**

Table of Tables

Table I.1 Mechanisms for preventing pre-RC formation....................................17
Table I.2 Checkpoint components across the species 31
Table II.1 - Strain list 45
Table III.1 - Strain list 83
Table IV.1 - Strain list...115
Table IV.2 – List of experiments performed for deep sequencing 145
Appendix Table List of oligonucleotides for *AT2062* microarray........................203
Appendix Table List of oligonucleotides for *ura4* microarray............................233

Table of Figures

Figure I.1 Formation of protein complexes leading to origin firing.........................5
Figure I.2 Hsk1 and Cdc2 regulate initiation of replication 14
Figure I.3 Steps leading to the formation of Initiation complex............................16
Figure I.4 Origins in fission yeast and stochasticity of origin selection................24
Figure I.5 Model for the S phase DNA damage checkpoint in Mammals 33
Figure II.1A S-phase DNA damage checkpoint analysis in asynchronous wild type cells...49
Figure II.1B S-phase DNA damage checkpoint analysis in asynchronous mutant cells...50
Figure II.2A S-phase DNA damage checkpoint analysis in G2 synchronized wild type cells. 53
Figure II.2B S-phase DNA damage checkpoint analysis in G2 synchronized mutant cells. 54
Figure II.2C S-phase DNA damage checkpoint analysis in G2 synchronized *cdc25Δ cdc2-Y15F* cells...55
Figure II.3A S-phase DNA damage checkpoint analysis in G1 synchronized wild type cells 60
Figure II.3B S-phase DNA damage checkpoint analysis in G1 synchronized mutant cells...61
Figure II.4 S-phase DNA damage checkpoint analysis in asynchronous Pyp3 cells...66
Figure II.5 S-phase DNA damage checkpoint analysis in G2 synchronized Pyp3 cells...68
Figure II.6 S-phase DNA damage checkpoint analysis in asynchronous cells using low MMS...74
Figure III.1 Schematic representation of RTS1 integration.................................90
Figure III.2 A general description of 2-Dimensional gel electrophoresis..............90
Figure III.3 *AT2062* replicates passively with or without RTS1 94
Figure III.4 Forks are getting blocked at RTS1 sites 96

Figure III.5 *AT2062* does not fire in early S phase 99
Figure III.6 Replication profile during S phase at *AT2062* in wild type cells 102
Figure III.7 Replication profile during S phase at *AT2062* in RTS1 strain 104
Figure III.8 Replication profile at *AT2062* in HU arrest using deep sequencing 107
Figure IV.1 Replication arrest in the presence of HU for *S.pombe*....................120
Figure IV.2 Process of identifying origins on Chromosome I............................123
Figure IV.3 Replication profiles of *S. pombe* chromosomes.............................125
Figure IV.4 comparison of *ura4* locus between array and sequencing data......127
Figure IV.5 Comparison of origins identified with previous studies 129
Figure IV.6 Replication arrest in the presence of HU in wild type *S. pombe* 132
Figure IV.7 Forks progress slowly during HU arrest in *S. pombe*......................134
Figure IV.8 Comparison between two independent *S. pombe* datasets............136
Figure IV.9A HU arrest of *S. octosporus*..140
Figure IV.9B Replication profiles of *S. octosporus* chromosomes.....................141
Figure IV.10 Replication profiles of *S. japonicus* scaffolds................................143
Figure V.1 Origin inhibition in the presence of damage.....................................150
Figure V.2 Timing of HU arrest to look at replication profile of *ura4* locus.........157
Figure V.3 Control for HU arrest profiles using Self hybridizations...................159
Figure V.4 No change in replication profile when dfp1 is overexpressed..........162
Figure V.5 Replication kinetics at ura locus 164
Figure V.6 Expression levels of *dfp1* alleles..169
Figure V.7 Over-expression of Dfp1 does not activate or inhibit the replication checkpoint..171
Figure V.8 Genome-wide transcript levels of in cells with Gal4-Dfp1 tethered at *AT3003*. 174

List of files to be included later

1. List of 448 oligonucleotides for microarray studies at *AT2062* as described in Chapter III will be included in the final E-copy of the thesis.

2. List of 768 oligos for microarray studies at the *ura4* locus as described in Appendix III.2 will be included in the final E-copy of the thesis.

Chapter I

Introduction- DNA replication and origin efficiency

The process of DNA Replication

DNA replication is a crucial step in the cell cycle of all organisms (Bell and Dutta 2002). High fidelity must be maintained during this process in order to maintain genomic integrity. In prokaryotes, replication initiates at a single locus or origin, while in eukaryotes, replication is initiated from multiple points along the genome (Dutta and Bell 1997; Gilbert 2001; Bell and Dutta 2002). Therefore, eukaryotic replication is highly complex process. Activation of origins and initiation of replication must be tightly regulated.

The process of initiation requires the sequential and cell cycle dependent binding of proteins (Fangman and Brewer 1992; Kelly and Brown 2000). Several protein complexes are assembled at sites of replication initiation allowing these origins to fire.

In the fission yeast *Schizosaccharomyces pombe*, the first complex formed is the pre-replicative complex (pre-RC) at the origin. This multiprotein complex consists of the origin recognition complex (ORC), Cdt1, Cdc6 and a hexameric complex of the minichromosomal maintenance proteins (Mcm 2-7) (Aparicio *et al.* 1997; Donovan *et al.* 1997; Tanaka *et al.* 1997; Nishitani *et al.* 2000). The fission yeast homolog of Cdc6 is Cdc18. Formation of the pre-RC on the origin is referred to as origin licensing. Two protein kinases, the cyclin-dependent kinase (CDK) Cdc2 and the Cdc7 homologue, Hsk1, phosphorylate pre-RC components, which lead to the activation of the pre-RC and loading of

replication protein Cdc45, a replication initiation protein essential for replication. Loading of Cdc45 is followed by the attachment of Replication protein A (RPA), polymerases, establishment of replication forks and the initiation of replication {Figure I.1}(Zou and Stillman 1998; Zou and Stillman 2000; Masai and Arai 2002).

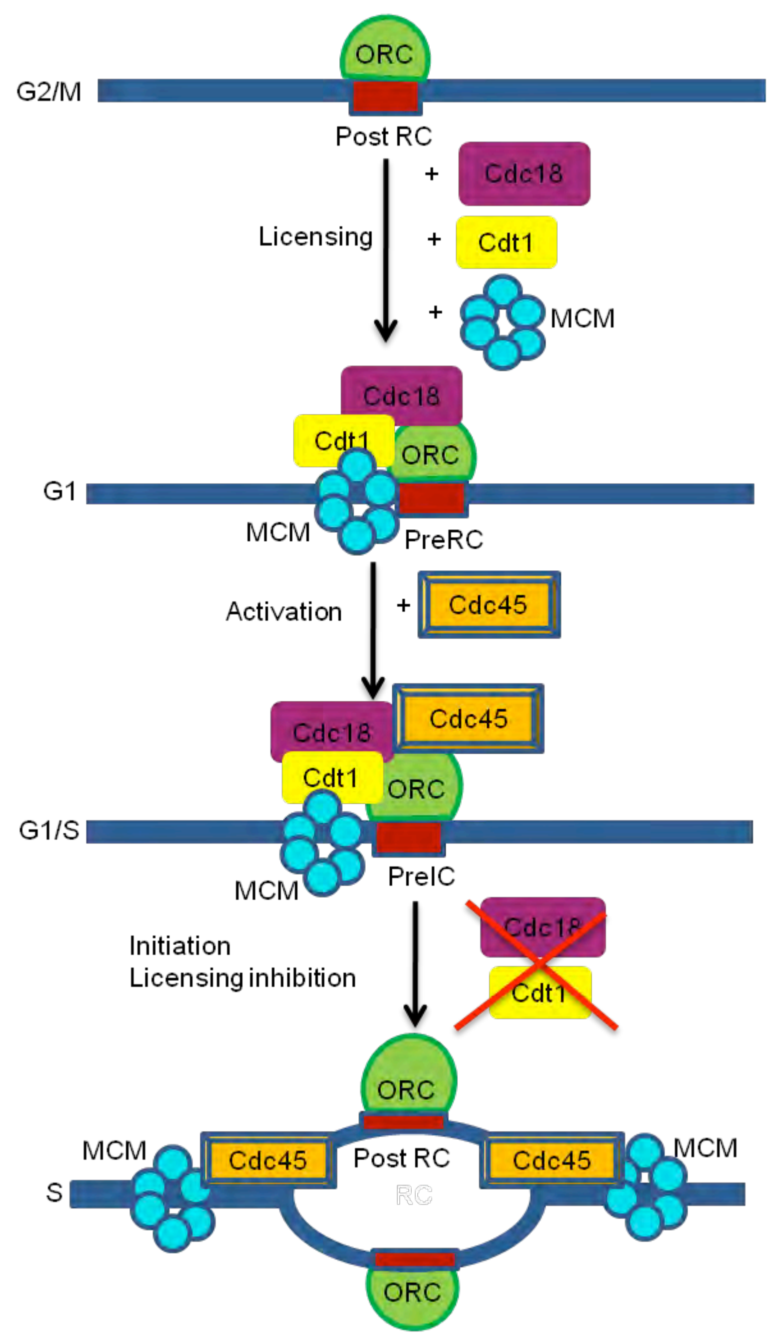
Formation of Pre-RC complex

The pre-RC is assembled during the G1 phase of the cell cycle. The origin recognition complex (ORC) is composed of six proteins Orc1-6. The ORC proteins are conserved from yeast to metazoans. ORC components display DNA binding activity and were originally identified in budding yeast (Bell and Stillman 1992). The timing of ORC recruitment does not appear to be conserved between organisms (Gilbert 2001). In fission yeast, ORC associates with replication origins both *in vivo* and *in vitro* and is formed during the late M and G1 phase (Chuang *et al.* 2002; Kong and DePamphilis 2002; Takahashi *et al.* 2003). Human ORC is capable of binding to any primary DNA sequence. ORC binds to the DNA in an ATP dependent manner. Fission yeast Orc4p has multiple AT hook motifs present at N terminus, which are absent in other organisms (Chuang and Kelly 1999). This hook is essential for viability in fission yeast and is shown to prefer AT rich tracts of DNA. ORC remains bound to the DNA throughout the cell cycle in the case of budding and fission yeast (Santocanale and Diffley 1996; Aparicio *et al.* 1997; Tanaka *et al.* 1997; Ogawa *et al.* 1999). However, in

Figure I.1 Formation of protein complexes leading to origin firing

Origin recognition complex marks the origin sites. At the beginning of G1, MCM complex is recruited by Cdc18 and Cdt1 loading onto the chromatin and this complex is known as pre-replicative complex (pre-RC). The origin is now licensed to fire. G1-S transition marks the recruitment of more factors, shown in the figure, to the pre-RC forming the pre-initiation complex (pre-IC) and leading to the initiation of replication. Once the origins are fired, further licensing is prevented and the complexes are converted to post-replicative complex (post-RC).

Figure I.1 Formation of protein complexes leading to origin firing



mammals, only the Orc2 subunit is continuously associated with the chromatin whereas Orc1 disassociates with the chromatin at the end of S phase and attaches again only at the beginning of G1 (Natale *et al.* 2000; Tatsumi *et al.* 2000). This data is consistent with the observation that Orc1 chromatin association is diminished in mitosis. Similarly, in *Xenopus* ORC is cleared from chromatin during metaphase (Carpenter *et al.* 1996; Romanowski *et al.* 1996).

ORC binding to the origin recruits the initiation proteins Cdc6 and Cdt1. ORC binding to the chromatin is essential for Cdc6 and Cdc6 in turn is required for Mcm2-7 association (Coleman *et al.* 1996; Aparicio *et al.* 1997). A direct association of the budding yeast protein with the origins has also been shown by ChIP (Tanaka *et al.* 1997). Cdc6 is an AAA+ ATPase and has been shown to increase the stability of ORC on chromatin while simultaneously inhibiting nonspecific ORC binding (Mizushima *et al.* 2000; Harvey and Newport 2003). Cdc6 is a cycling protein. In budding and fission yeast, after cells have entered S phase, Cdc6 is targeted for degradation by SCF^{CDC4} dependent ubiquitination and undergoes proteosome mediated degradation (Jallepalli *et al.* 1998; Wolf *et al.* 1999; Perkins *et al.* 2001). Degradation occurs after CDK dependent phosphorylation of the N terminus as shown by mutation studies (Jallepalli *et al.* 1998; Perkins *et al.* 2001). In mammals, CDK dependent phosphorylation promotes the export of Cdc6 from the nucleus. Cdc6 is subsequently degraded by the anaphase promoting complex (APC) during metaphase (Petersen *et al.* 2000).

In fission yeast, Cdt1 associates with the C terminus of Cdc6 and leads to the recruitment of MCM proteins to growing origin complex (Nishitani *et al.* 2000). Much like Cdc6, Cdt1 levels peak in G1 and as the cell progresses through S phase it declines. Crystal structure of the Cdc6 homolog *Pyrobaculum aerophilium* has shown that it is an AAA+ ATPase and is suspected to be a prime candidate to act as the clamp loader as a part of ORC (Perkins and Diffley 1998; Liu *et al.* 2000). Some of the proteins of the ORC- Orc1, 4 and 5- also belong to AAA+ ATPase family (Lee *et al.* 2000). It is highly likely that these proteins along with Cdc6 act as a clamp loader to load the replicative helicase MCM complex on the origins.

After Cdt1 and Cdc6 loading, origin DNA must be unwound. This process requires the Mcm2-7 complex (Walter and Newport 2000). The MCM complex is a heterohexamer formed by six different Mcm proteins. Mcm2-7 are essential in both budding and fission yeast (Kelly and Brown 2000). ORC, Cdc6 and Cdt1 are all required for MCM origin recruitment. The recruitment of MCM requires ATP hydrolysis by both Cdc18 and ORC subunits (Randell *et al.* 2006). After the MCM complex has been loaded, ORC and Cdc6 are no longer required for origin firing indicating that these two proteins act primarily to load the MCM helicase (Rowles *et al.* 1999; Walter and Newport 2000). MCM proteins are the only components of the pre-RC known to associate with replication forks (Labib *et al.* 2000). Previous works using various techniques have suggested the Mcm complex to act as the replicative helicase during S phase (Labib and Diffley 2001; Forsburg

2004). MCM complex has only recently been shown to actually have the *in vitro* helicase activity (Bochman and Schwacha 2008). The subcomplex of Mcm2-7- Mcm4, 6 and 7- display limited *in vitro* helicase activity (Lee *et al.* 2001; You *et al.* 2002) suggesting that Mcm4/6/7 acts as the core helicase with Mcm2/3/5 working as the regulatory subunits. All six proteins are required for ATPase activity (Schwacha and Bell 2001). The structure of the MCM complex as seen by EM studies shows the fission yeast MCM proteins forming a doughnut-like structure with a central cavity (Adachi *et al.* 1997). A similar toroidal structure with six lobes surrounding a central cavity has been observed by EM for human Mcm4/6/7 complex (Sato *et al.* 2000). Presumably DNA strand(s) occupy this central cavity.

MCM proteins are present in the nucleus only during G1 and S phase in budding yeast and are actively exported to the cytoplasm during G2. Only an intact six-subunit complex is able to re-enter nucleus in both budding and fission yeast (Labib *et al.* 1999; Nguyen *et al.* 2000). Export of the unbound MCM proteins has been shown to be mediated by the Crm1 nuclear export factor in fission yeast (Pasion and Forsburg 1999). In fission yeast, the bulk of the MCM proteins are constitutively nuclear (Bell and Dutta 2002). Likewise, in metazoans, MCM proteins are present in the nucleus constitutively and their chromatin association weakens through the S phase (Lei and Tye 2001).

MCM recruitment completes the formation of the pre-RC. Origins bound by the pre-RC are licensed to fire or are capable of firing. However, only a fraction of licensed origins actually fire (Santocanale and Diffley 1996; Walter and Newport 1997; Okuno *et al.* 2001). If, for some reason, the origins that are licensed, fail to fire, then neighboring licensed dormant origins fire (Santocanale *et al.* 1999).

In order for the origins to fire, the pre-RC must be converted to pre-initiation complex (pre-IC) during the G1 to S transition. Not all pre-RC's are converted to pre-IC. Cdc2 and Hsk1-dependent phosphorylation is required for initiating replication at the pre-RCs and ultimately, activation of the MCM helicase.

Initiation of Replication

Mcm10 is the earliest initiation factor that binds to the pre-RC. It is essential for the subsequent steps of the complex formation to take place. A role has been suggested for Mcm10 in the fork elongation and presumably travels along with the replication fork (Gregan *et al.* 2003). Mcm10 seems to have several critical functions while the replication fork is traveling and is needed for pre-RC formation only in budding yeast and not in fission yeast or *Xenopus* (Homesley *et al.* 2000). Mcm10 binds to the chromatin independently of Mcm2-7 except in *Xenopus* where the MCM complex needs to be loaded first

(Wohlschlegel *et al.* 2002; Gregan *et al.* 2003). In mammals, MCM10 displays a cyclic chromatin association, which is highest during S phase, unlike budding yeast, which displays constitutive chromatin binding (Izumi *et al.* 2000; Izumi *et al.* 2001).

Replication is triggered at the origins when the S phase cell cycle regulated kinases, Hsk1 (Cdc7 in budding yeast) and cyclin dependent kinase (CDK) Cdc2, are activated (Kelly and Brown 2000; Bell and Dutta 2002; Masai and Arai 2002; Kim *et al.* 2003). CDK and Dbf4-dependent kinase (DDK) modify the pre-RC and facilitate the loading of additional factors, which are required for the initiation of DNA synthesis (Jares and Blow 2000).

Cyclin dependent kinases ensure the progression through various phases of the cell cycle including initiation of replication. In fission yeast, the CDK activity is low at the beginning of the S phase but the levels gradually increase through the S phase. It is presumed that CDK's intermediate levels are sufficient to initiate replication but are below the threshold for its mitotic functions. However, CDK does play a major role during replication, phosphorylating many replication factors (Kelly and Brown 2000). Mammals have different CDKs for different stages of the cell cycle. Fission yeast however expresses a single CDK, Cdc2. The major cyclin Cdc2 associates with in fission yeast is Cdc13. There are three other cyclins Cig1, Cig2 and Puc1 which are also present and have overlapping functions (Fisher and Nurse 1996).

In addition to CDK Cdc2, the Cdc7 kinase is required. Cdc7 has a catalytic partner, Dbf4 and the heterodimer is known as Dbf4-dependent kinase (DDK). The fission yeast analog of budding yeast Dbf4 is Dfp1. Dfp1 expression is periodic with levels peaking at the beginning of S phase. DDK kinase activity is required for Cdc45 recruitment to the pre-RC (Walter and Newport 2000; Zou and Stillman 2000). DDK activates origins throughout S phase (Bousset and Diffley 1998; Patel *et al.* 2008). DDK phosphorylates MCM and Mcm10 is required for the interaction between DDK and MCM complex (Lee *et al.* 2003). The phosphorylation has been shown both *in vivo* and *in vitro* (Lei *et al.* 1997; Jares and Blow 2000; Jares *et al.* 2000). DDK preferentially phosphorylates chromatin bound MCM (Sheu and Stillman 2006).

Recently the essential targets for phosphorylation by CDK have been identified as Sld2 and Sld3 (Tanaka *et al.* 2007; Zegerman and Diffley 2007). Phosphorylation allows them to bind to Dpb11. The fission yeast homolog of Dpb11 is Rad4. This in turn recruits Cdc45, Go, Ichi, Nii and San (GINS) and DNA polymerases to the origin DNA (Labib and Gambus 2007).

Cdc45 is the initiation protein, which is required not only for initiating replication but also for maintaining replication. It has been suggested that Cdc45 along with GINS are responsible for stimulating the helicase activity of MCM by forming the Cdc45/Mcm2-7/GINS (CMG) complex (Moyer *et al.* 2006).

It has also been suggested that CDK and DDK function in parallel and lead to the loading of Cdc45 onto the chromatin {Figure I.2} (Dolan *et al.* 2004). Cdc45 recruitment initiates replication, presumably by activating the MCM helicase unwinding origin DNA. DNA unwinding leads to the recruitment of the DNA polymerases- DNA pol α, δ and ε. The processivity factor, Proliferating Cell Nuclear Antigen (PCNA) then encircles the DNA and topologically links the polymerase to DNA (Jonsson and Hubscher 1997). This loading of PCNA is done by the clamp loader, replication factor C (RFC) (Ellison and Stillman 2001). Replication forks then travel in a bi-directional manner with MCM and Cdc4 traveling with the forks {Figure I.3}.

How is re-replication prevented?

Replication must be a highly coordinated and controlled process. Cells must ensure replication occurs only once during each round of the cell cycle. Re-replication would lead to chromosomal breakage and genomic instability. To prevent re-replication, cells must ensure any origin will fire only once during S phase. Several studies have identified different mechanisms utilized to ensure fired origins cannot fire again. In budding yeast, origin firing leads to disassembly of the pre-RC, hence ensuring that origins fire only once during each cell cycle {Table I.1} (Diffley 1996).

Figure I.2 Hsk1 and Cdc2 regulate initiation of replication

Hsk1-dfp1 kinase regulates the assembly of the replication complex on the origins. It does so by phosphorylating MCM proteins to allow for the binding of Sld3. Rad4 and Drc1 on the other hand mediate the signal from CDK Cdc2 to initiate replication. Cdc2 phosphorylates Drc1. The regulation by both DDK and CDK is required to activate the preRC and initiate replication and disrupting either signals leads to the disruption of the origin firing. Adapted from Dolan and forsburg 2004

Figure I.2 Hsk1 and Cdc2 regulate initiation of replication

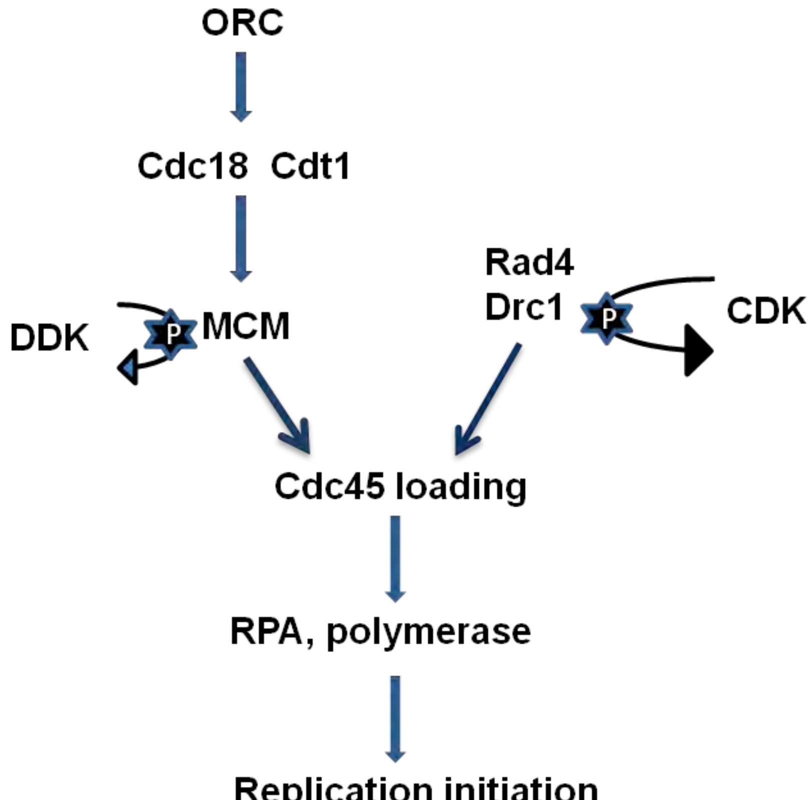


Figure I.3 Steps leading to the formation of Initiation complex

The model is based on studies of the various replication factors. At the beginning of S phase, after the formation of pre-RC, Sld3 binds to the origins where the pre-RC is bound in a DDK dependent manner. CDK is then needed for the binding of GINS and Cut5. Cut5 and GINS are mutually dependent for binding. Cdc45 then binds to the origins only if these three factors have bound to the origin. Binding of cdc45 leads to the initiation of replication. Adapted from Yabuuchi and Yamada 2006

Figure I.3 Steps leading to the formation of Initiation complex

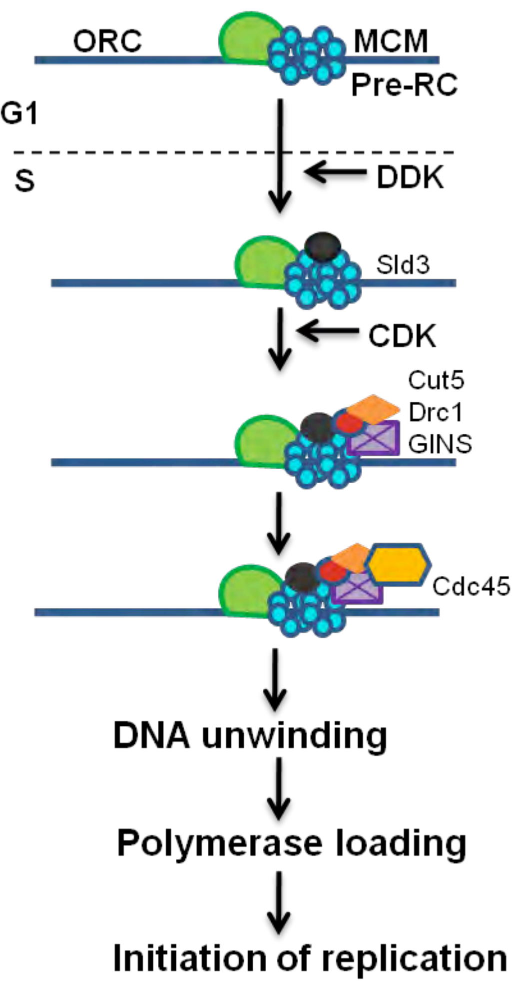


Table I.1 Mechanisms for preventing pre-RC formation

	S. cerevisiae	*S. pombe*	*Xenopus*	*Mammals*
ORC	CDK phosphorylation on chromatin	?	Destabilization of ORC on chromatin	Orc1 dissociation/ degradation
Cdc6	Degradation	Degradation	Nuclear Export	Nuclear Export CDK phosphorylation on chromatin
Cdt1	Nuclear export	Degradation	Degradation Geminin	Degradation Geminin
MCM	Nuclear export	?	Reduced affinity for chromatin	Reduced affinity for chromatin CDK inhibition of helicase activity

Cdc6, ORC and Cdt1 are the licensing factors that need to be regulated or prevented from reattaching to origin sites and licensing them. CDK is the primary regulator for preventing licensing once the S phase has begun, which it does by phosphorylating its targets. Intermediate levels of CDK are required for initiating S phase but also lead to prevention of pre-RC formation. Therefore pre-RC may only be formed when CDK levels are low at the M/G1 transition. Beginning with S phase initiation, CDK levels rise and thereby prevent pre-RC formation.

CDK interacts and phosphorylates subunits of ORC, which is necessary to prevent further pre-RC formation (Nguyen *et al.* 2001; Vas *et al.* 2001). Studies have shown that replication may be initiated in G2 phase if CDK activity is inhibited (Itzhaki *et al.* 1997). In mammals, ORC affinity for chromatin decreases after origins fire. In fission yeast, CDK is recruited by ORC and disruption of this interaction allows re-replication (Wuarin *et al.* 2002). In eukaryotes, CDK activity results in reduction of Cdc6 activity. In budding yeast, beginning in S phase, CDK protein levels increase, phosphorylate Cdc6 and target it for SCF mediated ubiquitination and proteolytic degradation (Drury *et al.* 1997; Jallepalli *et al.* 1997). In vertebrates, Cdc6 activity is prevented by its export from the nucleus upon CDK phosphorylation (Saha *et al.* 1998; Delmolino *et al.* 2001). Cdt1 on the other hand becomes ubiquitinated only after PCNA loading at fired origins in fission yeast (Arias and Walter 2006). In mammals, Cdt1 is phosphorylated by CDK and targeted for degradation (Liu *et al.* 2004). In budding yeast Cdt1 is exported away from the nucleus (Tanaka and Diffley 2002).

MCMs on the other hand, travel with the forks and hence are no longer present at the origins (Aparicio *et al.* 1997). The MCMs that are nuclear but not bound to the chromatin have a decreased affinity for chromatin due to phosphorylation by CDK (Coue *et al.* 1996; Fujita *et al.* 1998). In budding yeast, MCMs released from the forks are exported from the nucleus (Labib *et al.* 1999; Nguyen *et al.* 2000).

Another mechanism for preventing re replication is seen in metazoans. Geminin forms a dimer with and prevents Cdt1 dependent origin licensing (McGarry and Kirschner 1998; Maiorano *et al.* 2004). Geminin prevents Cdt1 dependent Mcm recruitment through steric hindrance of the Mcm-Cdt1 interaction (Cook *et al.* 2004).

Early studies on origin sites

Replication is initiated at defined regions of the genome called origins. Origins are best characterized in budding yeast. Origins are defined sequences capable of initiating replication or autonomously replicating sequences (ARS). Origins were originally identified using plasmid stability studies in proliferating yeast. Two-dimensional gel electrophoresis has also been used to study origins in both plasmid and native chromosomal context (Brewer and Fangman 1987; Brewer and Fangman 1991). Budding yeast contains a 10-12bp ARS conserved signature sequence called the ARS consensus sequence (ACS). Budding yeast

origins are about 100-150 bp containing one ACS element and 2-3 additional origin B elements (Newlon and Theis 1993; Bell *et al.* 1995). Only 1 of these ACS elements is actually conserved between origins (Rao et al, 1994). Budding yeast origins fire with high efficiency at defined periods of the S phase (Kelly and Brown 2000; Gilbert 2001). Origins are defined as either early or late firing (Raghuraman *et al.* 2001; Yabuki *et al.* 2002). However, there is no clear demarcation between early and late firing origins. Rather, origins fire throughout S phase (Raghuraman *et al.* 2001).

Due to a lack of well-defined difference between early and late firing origins it is very difficult to identify which origins fire early or which fire late. This problem can be circumvented by using the drug hydroxyurea. Origins firing in the presence of hydroxyurea are defined as early origins (Santocanale and Diffley 1998; Kim and Huberman 2001; Lopes *et al.* 2001). Hydroxyurea is a ribonucleotide reductase inhibitor, which prevents deoxyribonucleotide synthesis. The inhibition leads to nucleotide depletion and because of this, replicating forks cannot incorporate nucleotides at the regions where replication is occurring and hence the forks stall. Hydroxyurea triggers the replication checkpoint that prevents origins from firing. Early origins are able to fire before the pools of nucleotides are depleted and before the checkpoint activity prevents firing, hence they fire during early S phase (Yabuki *et al.* 2002). Late origins do not fire in HU and are prevented from doing so by the replication checkpoint (Shirahige *et al.* 1998). Genome-wide studies in yeast have identified and mapped many origins

(Raghuraman *et al.* 2001; Yabuki *et al.* 2002; Segurado *et al.* 2003; Feng *et al.* 2006; Heichinger *et al.* 2006; Eshaghi *et al.* 2007; Hayashi *et al.* 2007).

Unlike budding yeast origins, metazoan origins are not well defined and average origin firing efficiency is low. Several studies indicate that in *Xenopus* and *Drosophila* embryos any region is capable of acting as an origin and replication therefore may be initiated from anywhere in the genome (Hyrien and Mechali 1993; Shinomiya and Ina 1994). In mammals origins have been difficult to identify. Regions where ORC binds have not been identified and the origins that are known are highly inefficient with the best studied example, hamster DHFR locus firing only 20% of the time (Burhans and Huberman 1994; Gilbert 2001; Dijkwel *et al.* 2002).

Origins in fission yeast

Fission yeast serves as an excellent model organism for origin studies related to higher eukaryotes. Unlike budding yeast origins, but like metazoans, fission yeast origins contain no consensus sequence. However, replication does initiate at defined regions in the genome in fission yeast (Dubey *et al.* 1994; Gomez and Antequera 1999). Fission yeast origins are large AT rich regions (Clyne and Kelly 1995; Dubey *et al.* 1996; Segurado *et al.* 2003; Dai *et al.* 2005). Similar to metazoans, fission yeast origins are inefficient with the average efficiency ranging between 25-40% (Dubey *et al.* 1994; Gomez and Antequera

1999; Kim and Huberman 2001; Segurado *et al.* 2002; Segurado *et al.* 2003; Patel *et al.* 2006). Lastly, fission yeast origins are not always interchangeable with the budding yeast origins (Clyne and Kelly 1995). Hence the mechanisms of origin regulation in fission yeast maybe more similar to metazoans {Figure I.4}.

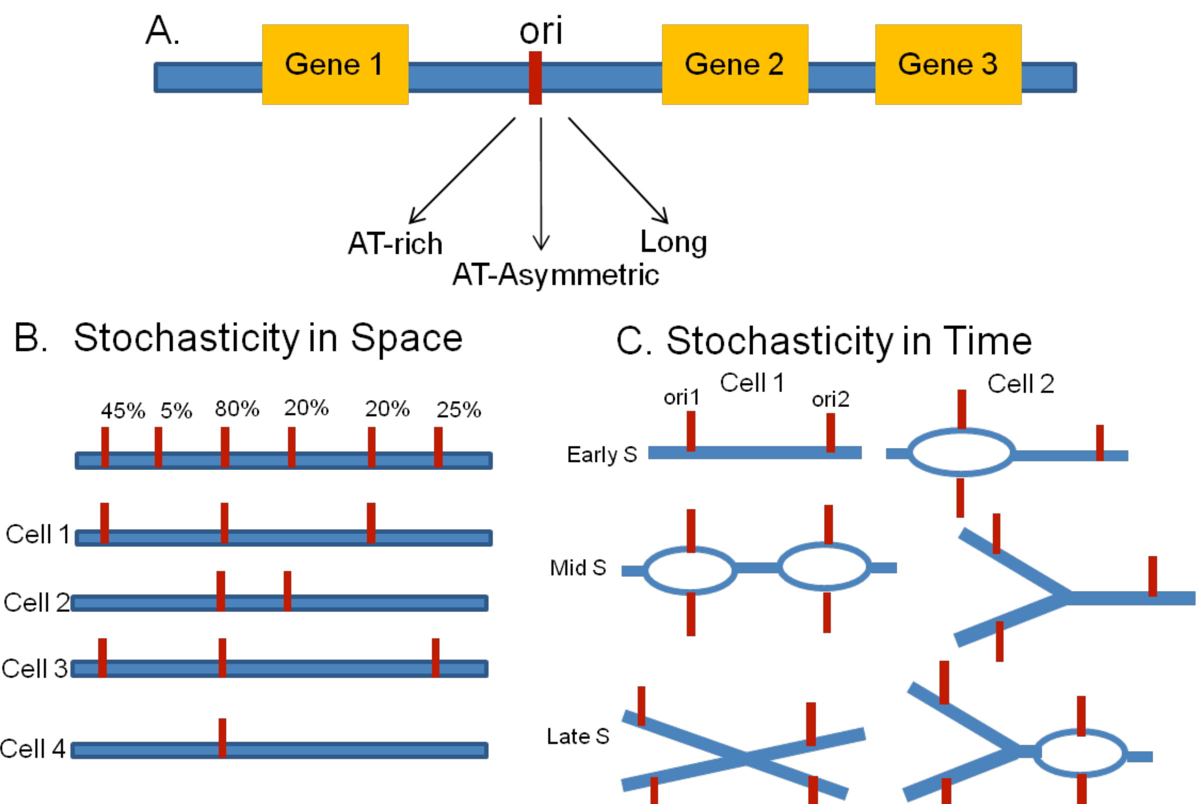
Genome-wide studies have been conducted recently in an effort to identify origins across the genome. One of the earliest efforts was done by bioinformatic analysis where AT rich sequences were the criteria used to select origins (Segurado *et al.* 2003). They looked for regions greater than 72% AT rich, and these AT rich segments should be present in 0.5-1 Kb windows. This method identified 384 AT rich islands, which could serve as origins and 20 of them tested for origin activity by 2-D gel electrophoresis. It was shown recently that fission yeast origins had properties similar to the inter-genic regions (Dai *et al.* 2005).

Another genome-wide analysis is based upon mapping single stranded DNA on ORF microarrays in the presence of HU. The analysis in an S-phase checkpoint deficient strain identified 321 origins in fission yeast (Feng *et al.* 2006). 61% of these origins function during a regular S-phase which suggests that S-phase checkpoint functions in suppressing many origins which will be discussed later.

Figure I.4 Origins in fission yeast and stochasticity of origin selection

(A) Fission yeast origins are characterized by AT rich islands. These are regions of asymmetric stretches of adenine or thiamine. The origins are located in the inter-genic regions as identified by various studies. (B) Origins in fission yeast display stochasticity in firing. Different origins fire in each cell during cell cycle. Each origin fires only in a fraction of cells. Origins are marked by a red line and the corresponding efficiency is given above the origin location. (C) Temporal stochasticity is marked by origins firing throughout S phase. There are no clear demarcations as to when an origin fires early and when it fires late. Adapted from Legouras and Lygerou, 2006

Figure I.4 Origins in fission yeast and stochasticity of origin selection



Recently, Heichinger *et al* have identified more origins in fission yeast based on an increase in copy number with a resolution of about 6.5 Kb (Heichinger *et al.* 2006). DNA content was measured in the G2 phase and S phase of the cell cycle and the regions that had doubled their amount of DNA were the regions that had replicated. They identified 401 strong and 503 putative weak origins which seemed to be spaced on average every 14 kb throughout the genome (Heichinger *et al.* 2006).

As part of this thesis, I have participated in a genome-wide study to find origins using a similar copy change number described in Heichinger *et al*, 2006. However, this study was done using deep sequencing the details of which will be described in Chapter IV. We have also done similar studies on two other *Schizosaccharomyces* species; *S. octosporus* and *S. japonicus*.

Origin location influences timing of firing

Origin location is important for its efficiency (Friedman *et al.* 1995). Inefficient origins can be made to fire if their passive replication is prevented by neighboring origins (Santocanale *et al.* 1999). Late firing origins located in a heterochromatic region may fire early if transferred to euchromatin (Stevenson and Gottschling 1999; Vogelauer *et al.* 2002). Also, early firing origins may be forced to fire late by placing them in heterochromatin (Friedman *et al.* 1996;

Zappulla *et al.* 2002). This data shows that origin location dictates whether that origin will fire early or late.

What is Random gap problem?

In contrast to budding yeast where efficient origins are spaced relatively evenly across the genome, only a few licensed origins actually fire in humans and fission yeast. Origin firing in fission yeast is random in nature compounding potential problems completing replication. Random firing was determined by measuring the distance between origins that had fired. Patel *et al* found an exponential distribution of inter-origin distances which was interpreted as stochasticity of origin firing {Figure I.4} (Patel *et al.* 2006). Random and inefficient origin firing may lead to disastrous consequences. Large regions of the genome may have no origin firing due to the stochastic firing of origins, thereby leading to cells taking a long time to complete replication, a problem known as 'random gap problem' (Lucas *et al.* 2000; Herrick *et al.* 2002; Jun *et al.* 2004). In a recent paper on budding yeast, it was shown that although budding yeast seems to have a highly regulated temporal program, when looked at globally, at the individual cell level replication seems to be stochastic. This data argues that even in budding yeast there is a randomness at a local level instead of a regulated temporal program as shown previously (Czajkowsky *et al.* 2008). However, cells are able to finish replication in an efficient manner (Hyrien *et al.* 2003). Hence

there is no random gap problem and cells must employ a mechanism to regulate origin firing. Several models have been proposed to explain this discrepancy.

How fission yeast reconciles random origin firing and inefficient origins with efficient replication is the focus of my studies. My hypothesis is that the origin efficiency increases as the cells progress through S phase. In order to test this, my thesis has focused on measuring the efficiency of an origin during the later part of S phase. Chapter III discusses the model that I believe explains this conundrum and my efforts to show this.

In order to understand why origins are inefficient biochemically, we hypothesized that there is a rate-determining factor that is responsible for activating all the origins. This factor would have to be present in limited quantities and be physically present at each origin. Any of the factors that are responsible for activating the pre-RC would be good candidates. Studies in our lab have shown that this factor is Dfp1 which is the activating co-factor of Hsk1 Kinase (Patel *et al.* 2008). This work is presented in Appendix III.1.

DNA damage checkpoints and origins

Replication is far from perfect and its progress is hampered by damage both endogenous and spontaneous, to DNA. In order to ensure an error free transmission of genetic material, cells have devised elaborate mechanisms such

as various repair systems and cell cycle checkpoints to detect unreplicated DNA, DNA damage and repair aberrant DNA structures (Zhou and Elledge 2000).

Checkpoints are molecular signaling cascades that delay or arrest the cell-cycle in response to DNA damage, thereby providing sufficient time for repair. Checkpoint signaling consists of damage sensors, which sense the damage, transducers which relay these signals, and effectors which regulate the various targets of the checkpoint (Elledge 1996). The phase of the cell cycle where the damage occurs determines the specific response. Checkpoints ensure the accurate segregation of genetic material and repair of damage and ensure that cells meet the specific cell size, mass and nutrition requirements. The absence of checkpoints can be lethal to cells. DNA damage results in mutations, chromosomal rearrangements and aneuploidy which can lead to cancer (Hartwell *et al.* 1994). There are four different DNA damage checkpoints. The G1-S, S-M and G2-M checkpoints are responsible for arresting the cell cycle until the damage is repaired. The S-phase damage checkpoint slows replication till the damage is repaired.

The G1-S checkpoint ensures that cells have reached a sufficient size before entering the S phase of the cell cycle and repair any damage during G1. G2-M checkpoint prevents mitosis in the presence of damage to ensure that damaged chromosomes do not undergo chromosomal segregation. The S-M checkpoint prevents mitosis till the entire DNA is replicated to ensure that cell

division does not take place with incomplete copies of the genome. The S-phase DNA damage checkpoint slows replication in the presence of damage to allow cells to repair the damage before completing replication. My work focuses on the S-phase DNA damage checkpoint.

How does S-phase DNA damage checkpoint work?

S-phase DNA damage checkpoint is activated and responds to DNA damage occurring in S phase. This checkpoint slows replication in contrast to the other checkpoints, which induce a complete cell cycle arrest (Painter and Young 1980; Rowley *et al.* 1999). This checkpoint is conserved in eukaryotic organisms and requires the Ataxia-Telangectasia Mutated (ATM) family of protein kinases (Kastan and Lim 2000). There are two members of this family, Ataxia-Telangectasia Related (ATR) and ATM in metazoans (Savitsky *et al.* 1995; Bentley *et al.* 1996), Mec1 & Tel1 in budding yeast, and Rad3 & Tel1 in fission yeast (Rhind and Russell 1998). In fission yeast Rad3 responds to all forms of DNA damage (Bentley *et al.* 1996). Effector proteins are also conserved and consist of Chk1 & Chk2 in vertebrates, Chk1 and Rad53 in budding yeast, and Chk1 & Cds1 in fission yeast {Table I.2} (Rhind and Russell 2000). DNA damage induced by methyl methane sulfonate (MMS) also slows fork progression (Tercero and Diffley 2001). Delay in S phase progression can be induced by the checkpoint either by inhibition of origin firing or by slowing fork progression. The presence of well defined origins in budding yeast has shown that late origins are

inhibited from firing and are Mec1 and Rad53 dependent (Shirahige *et al.* 1998; Tercero and Diffley 2001). However, the downstream targets of Rad3 have not been identified for slowing of replication.

A well characterized target of the checkpoint effectors in metazoans is cyclin dependent Kinase, Cdk2 {Figure I.5} (Falck *et al.* 2001). The inhibition of origin firing is mediated mainly by the effector Chk2 targeting Cdk2 via Cdc25. Cdk2 is inactivated when phosphorylation at Tyrosine-15 (Tyr-15) occurs and this inhibitory phosphate must be removed to activate Cdk2. The Cdk2-cyclin E complex, which facilitates loading of Cdc45 onto chromatin, is activated by Cdc25 phosphatase that removes the inhibitory phosphate from Cdk2. Chk2 phosphorylates Cdc25 thereby targeting it for degradation (Mailand *et al.* 2000; Falck *et al.* 2001; Sorensen *et al.* 2003; Xiao *et al.* 2003). However, in yeasts Cdc25 inactivation by the checkpoint has not been shown and the regulation of Cdc2 via this pathway has not been demonstrated.

A parallel pathway in the S-phase DNA damage checkpoint has been shown to exist in vertebrates and fission yeast which appears to act through a heterotrimeric complex consisting of Mre11, Rad50 and Nbs1 (MRN) (Costanzo *et al.* 2001; Falck *et al.* 2002). Mutations in one of the pathways show only a partial loss of the checkpoint and a loss of both the pathways is required for a complete loss of slowing {Figure I.5} (Falck *et al.* 2002).

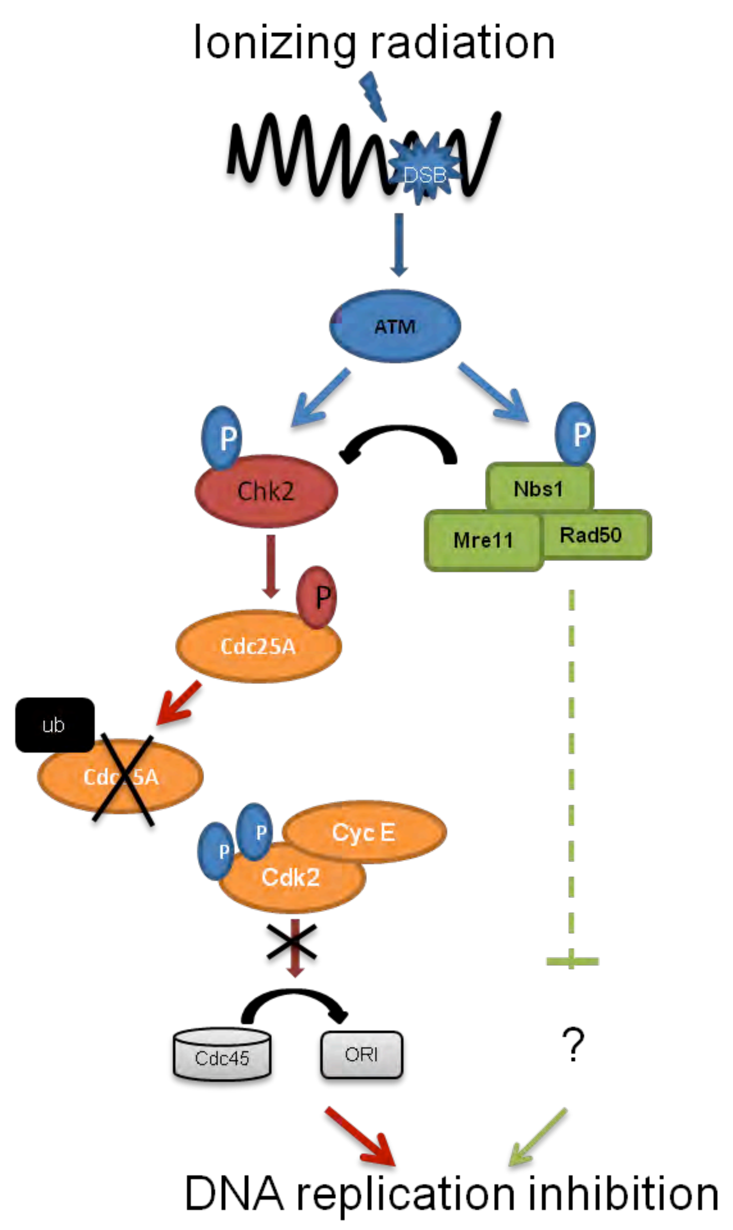
Table I.2 Checkpoint components across the species

Protein functions	S. cerevisiae	*S. pombe*	Mammals
ATM/ATR-like kinases	Mec1p Tel1p	Rad3 Tel1	ATR ATM
ATR-interacting proteins	Ddc2p	Rad26	ATRIP
RFC-like proteins	Rad24p Rfc2-5p	Rad17 Rfc2-5	Rad17 Rfc2-5
Mediators	Rad9p Mrc1p	Crb2 Mrc1	BRCA1 Claspin
Replication fork stabilizers	Tof1p	Swi1	Timeless
DSB recognition processing	Mre11p Rad50p Xrs2p	Rad32 Rad50 Nbs1	Mre11 Rad50 Nbs1
Effector kinases	Rad53p Chk1p	Cds1 Chk1	Chk2 Chk1

Figure I.5 Model for the S phase DNA damage checkpoint in Mammals

Ionizing radiation induces double stranded breaks (DSB). IR activates the checkpoint where ATM triggers the two parallel pathways, which work together to inhibit DNA replication. ATM phosphorylates Chk2, which in turn induces the destruction of Cdc25A phosphatase. The destruction of Cdc25A prevents the activation of the S phase cyclin E/Cdk2 complex by dephosphorylation, and does not allow the binding Cdc45 onto the origins. This inhibits the firing of origins and slows replication. ATM also initiates the second pathway by phosphorylating Nbs1, required for activating Nbs10-Mre11-Rad50 complex. The mechanism of replication slowing by this pathway is unknown. Adapted from Falck and Petrini, 2002

Figure I.5 Model for the S phase DNA damage checkpoint in Mammals



Similar to vertebrates, MRN mutants in fission yeast display a partial defect in S phase slowing implying the existence of a parallel pathway similar to mammals (Willis and Rhind 2009). The MRN independent pathway of the checkpoint has been shown to be dependent upon Rad3 and Cds1 (Marchetti *et al.* 2002). Although the downstream targets of this pathway are not known, recent studies have identified Cdc25 a possible candidate (Kumar and Huberman 2004). However, this seems unlikely in the case of fission yeast as the dephosphorylation of Cdc2 (fission yeast Cdk2) would lead to cells undergoing premature mitosis (Moser *et al.* 2000). It has always been assumed that Cdc2 remains phosphorylated during S phase in fission yeast. Therefore the role of

Cdc25 in the checkpoint needs to be studied to remove the ambiguity in the field. I have shown that Cdc25 is not the target of the intra S-DNA damage checkpoint in chapter II of my thesis (Kommajosyula and Rhind 2006). This shows that although the inhibition of origin firing is conserved between vertebrates and yeast; the mechanism is different (Shirahige *et al.* 1998; Kommajosyula and Rhind 2006; Kumar and Huberman 2009).

A key regulator required for origins to fire in S phase is the Hsk1/Dfp1 kinase. Hsk1 is a serine/threonine kinase that becomes activated after binding to its regulatory subunit Dfp1 (Jackson *et al.* 1993; Johnston *et al.* 1999). The activated Hsk1 then phosphorylates Mcm proteins at the origins (Lei *et al.* 1997). Budding yeast homologue Cdc7 has been shown to be necessary for initiation of

early and late firing origins (Bousset and Diffley 1998; Donaldson *et al.* 1998). Cdc7 has been implicated in the checkpoint as a potential downstream target (Jares and Blow 2000). Hsk1 has been shown to be phosphorylated by Cds1 (Chk2 homolog) upon treatment with Hydroxyurea (HU) making Hsk1 a potential target for the checkpoint (Snaith *et al.* 2000). Thus, the DNA damage checkpoint can target Hsk1 through Cds1 and inhibit origin firing and slowing replication.

Origin regulation by checkpoints

A recent genome-wide study in fission yeast to identify origins was conducted in a checkpoint deficient strain. In this study the origins that were identified had a 61% overlap with origins firing in a regular S-phase. This showed that checkpoints also play a role in regular S phase (Feng *et al.* 2006). Various studies have shown that damage during S phase activates the S-phase DNA damage checkpoint, which inhibits origin firing (Shirahige *et al.* 1998; Kelly and Brown 2000). In mammals the lack of well defined origins has hampered the study of the checkpoint mechanisms. A few studies have shown that upon damage origin firing is inhibited by the checkpoint (Larner *et al.* 1999). Inhibition of origin firing has been supported by 2D gel analysis on replication of rDNA locus. rDNA is one of the few loci in mammals showing well defined early or late origins (Larner *et al.* 1999). The presence of well defined origins in mammals being an exception rather than a rule makes it harder to extrapolate the results from these studies to the whole genome. These studies have all been carried out

on a population level and not on individual origins. A recent study using DNA fiber labeling technique has shown that different DNA damaging agents slow replication by different mechanisms including inhibition of origin firing and slowing of fork progression (Merrick *et al.* 2004). It is however not clear as to how fork progression is slowed and what molecules are playing a role in it. Since fission yeast have a similar origin setup, identification of the targets and mechanism of the checkpoint in them will help in understanding the human checkpoint due to the conserved nature of the checkpoints.

Work in my thesis has shown that Cdc25 is not a target of the intra-S DNA damage checkpoint. I have attempted to measure the efficiency of a late replicating sequence to fire. We have also shown that deep sequencing is a powerful tool for the identification of new origins in the *Schizosaccharomyces* species and further analysis will continue on this project.

Chapter II

Cdc2 tyrosine phosphorylation is not required for the S-phase DNA damage checkpoint in fission yeast

ABSTRACT

The S-phase DNA damage checkpoint slows replication when damage occurs during S phase. Cdc25, which activates Cdc2 by dephosphorylating tyrosine-15, has been shown to be a downstream target of the checkpoint in metazoans, but its role is not clear in fission yeast. The dephosphorylation of Cdc2 has been assumed not to play a role in S-phase regulation because cells replicate in the absence of Cdc25, demonstrating that tyrosine-15 phosphorylated Cdc2 is sufficient for S phase. However, it has been reported recently that Cdc25 is required for the slowing of S phase in response to damage in fission yeast, suggesting a modulatory role for Cdc2 dephosphorylation in S phase. We have investigated the role of Cdc25 and the tyrosine phosphorylation of Cdc2 in the S-phase damage checkpoint, and our results show that Cdc2 phosphorylation is not a target of the checkpoint. The checkpoint was not compromised in a Cdc25 overexpressing strain, a strain carrying non-phosphorylatable form of Cdc2, or in a strain lacking Cdc25. Our results are consistent with a strictly Cdc2-Y15 phosphorylation-independent mechanism of the fission yeast S-phase DNA damage checkpoint.

INTRODUCTION

Cells slow replication in response to DNA damage during S phase (Bartek *et al.* 2004). This S-phase DNA damage checkpoint, also know as the intra-S checkpoint, does not completely block replication. Instead, it reduces the rate of bulk replication, about 50% in human cells, presumably allowing cells to coordinate replication with repair or bypass of the damage (Painter and Young 1980; Bartek *et al.* 2004). Although this checkpoint has been proposed to allow for the repair of damage during S phase, there is not a strong correlation between checkpoint proficiency and damage tolerance. Furthermore, DNA damage induced before or during S phase can persist through the checkpoint and be repaired in G2 (Orren *et al.* 1997; Rhind and Russell 1998). Nonetheless, loss of the checkpoint leads to increased chromosomal rearrangements and profound cancer predisposition in humans (Petrini 2000; Myung *et al.* 2001).

The checkpoint pathway regulating replication in response to DNA damage is conserved amongst eukaryotes (Bartek *et al.* 2004). Members of the ATM-family of protein kinases form the center of the checkpoint pathway, serving to recognize DNA damage and initiate checkpoint signaling. ATM itself appears to be the major kinase in the vertebrate S-phase DNA damage checkpoint; the related kinases, Mec1 and Rad3, are required for the checkpoint in budding and fission yeast, respectively. When activated, these kinases phosphorylate a number of downstream effectors including the FHA-containing effector kinases –

Rad53 in budding yeast, Cds1 in fission yeast and Chk2/Cds1 in vertebrates. In addition to these checkpoint kinases, an array of accessory damage recognition and checkpoint mediator proteins are also conserved. Although this signaling pathway is well conserved, it is less clear if its targets are also conserved.

A priori, there are two ways that the checkpoint could slow replication. It could reduce the number of replication forks by inhibiting origin firing or arresting a subset of active forks, or it could slow the rate of progression of a majority of forks. Origin firing is inhibited by the checkpoint in vertebrates and in budding yeast (Santocanale and Diffley 1998; Shirahige *et al.* 1998; Larner *et al.* 1999; Costanzo *et al.* 2000). In vertebrates, Chk2 regulates origin firing by targeting Cdc25A for proteolysis, thus preventing the dephosphorylation and activation of S-phase cyclin-dependent kinases such as Cdk2/Cyclin E, which are required for origin firing throughout S phase (Costanzo *et al.* 2000; Falck *et al.* 2001).

In addition to the Cdc25-dependent regulation of origin firing, there is a parallel, Cdc25-independent checkpoint mechanism in mammals (Falck *et al.* 2002; Henry-Mowatt *et al.* 2003). Although the mechanism of this branch of the checkpoint is not well understood, it is known to require ATM phosphorylation of MRN, a heterotrimeric recombinational repair complex consisting of Mre11, Rad50 and Nbs1. MRN is involved in homologous and non-homologous recombinational repair, as well as meiotic recombination, DNA damage signaling and telomere maintenance. The fact that MRN is required only for the Cdc25-

independent branch of the checkpoint suggests that it acts downstream in the checkpoint pathway, rather than as an upstream signaling factor (Falck *et al.* 2002). The regulation of fork progression has also been shown to require the XRCC3 recombination protein (Henry-Mowatt *et al.* 2003). The role of MRN in the regulation of recombination, and the role of XRCC3 in regulating fork progression, has lead to the speculation that the checkpoint may slow replication fork progression through induction of replication-coupled recombinational repair (Rhind and Russell 2000; Henry-Mowatt *et al.* 2003).

The targets of the S-phase DNA damage checkpoint in fission yeast are less well defined. The role of the tyrosine-15 phosphorylation of Cdc2 (the only cyclin-dependent kinase in fission yeast) as a checkpoint target has been well established. In response to DNA damage in G2 or replication blocks during S phase, Cdc25 is inhibited, preventing the dephosphorylation of Cdc2 tyrosine-15 and arresting cells before mitosis (Rhind *et al.* 1997; Rhind and Russell 1998). It has also been reported that inhibition of Cdc25 and phosphorylation of Cdc2 tyrosine-15 are required to slow replication in response to DNA damage (Kumar and Huberman 2004).

The published work notwithstanding, there is reason to suspect that Cdc2 tyrosine phosphorylation is not the target of the S-phase DNA damage checkpoint. Since Cdc2 is the only cyclin-dependent kinase in fission yeast, and since it is required for both replication and mitosis, it has been assumed that

there must be different mechanisms of Cdc2 regulation that independently regulate these two events. The model with the most experimental support proposes that different levels of Cdc2 activity trigger the different events: replication is triggered by moderate level of Cdc2 activity, comprised of tyrosine phosphorylated Cdc2/cyclin complexes, and mitosis is triggered by the high level Cdc2 activity achieved when Cdc2/cyclin complexes are dephosphorylated (Stern and Nurse 1996). Consistent with this model, tyrosine-15 kinase activity of the Mik1 tyrosine kinase is high in S-phase, while Cdc25 levels are low, favoring Cdc2 tyrosine phosphorylation during S-phase (Moreno *et al.* 1990; Christensen *et al.* 2000). Consequently, Cdc2 remains largely phosphorylated during S-phase (Gould and Nurse 1989). Furthermore, it is clear that the bulk of Cdc2 cannot be dephosphorylated during S-phase, because such premature dephosphorylation leads to immediate and catastrophic mitosis (Lundgren *et al.* 1991). These observations are inconsistent with general activation of Cdc25 during S-phase. Yet, for inhibition of Cdc25 to be an important target of the S-phase DNA damage target, Cdc25 would have to be active during S-phase, and required for timely replication. Therefore, its activity would have to be limited, either in extent or location, to prevent premature mitosis. Such a subtle regulatory role for Cdc25 seems unlikely, because Cdc25 can be replaced by unrelated tyrosine phosphatases, either human T-cell protein tyrosine phosphatase or over-expression of fission yeast Pyp3 (Gould *et al.* 1990; Millar *et al.* 1992). In both cases, replication appears normal (our unpublished result). This line of reasoning

argues against a role for Cdc25 in the S-phase DNA damage checkpoint. Therefore we have revisited the question of whether Cdc25 or the tyrosine phosphorylation of Cdc2 is required for the S-phase DNA damage checkpoint in fission yeast.

MATERIALS AND METHODS

Yeast methods

Yeast were grown in YES at 30°C and manipulated by standard methods (Forsburg and Rhind 2006). Temperature-sensitive (ts) cells were grown at 25°C unless otherwise stated. Strains used for this study are listed in Table 1.

Flow cytometry methods

Isolated nuclei were prepared for flow cytometry by an adaptation of the protocol of Carlson el al (Carlson et al. 1997; Forsburg and Rhind 2006). 1.0 OD of cells was fixed in 70% EtOH, washed in 1 ml 0.6M KCl, resuspended in 1 ml 0.6 M KCl, 1 mg/ml Novozym 234 (Sigma L1412), 0.3 mg/ml Zymolyase 20T and incubated for 30 min at 37°C. The cells were pelleted, resuspended in 1 ml 0.1 M KCl 0.1% triton-X100 and incubated for 5 minutes at room temperature. The cells were washed and resuspended in 1 ml 20 mM Tris-HCl, 5 mM EDTA pH 8.0. 10 µl 20 mg/ml RNase A was added and the cells were incubated overnight at 37°C. The spheroplasted cells were disrupted, and isolated nuclei released, by sonication with a Branson Sonifier using a microtip at 0.7 power for 5 seconds. 300 µl of disrupted cells were added to 300 µl of 2 mM Sytox Green (Molecular Probes) in PBS and analyzed on a Becton-Dickinson FACScan flow cytometer. G1 synchronized experiments were quantitated in CellQuest (Becton-

Table II.1 - Strain list

Strain	Genotype	Source
yFS104	*h+ leu1-32 ura4-D18*	Lab Stock
yFS189	*h- leu1-32 ura4-D18 ade6-704 rad3::ura4*	Lab Stock
yFS260	*h- leu1-32 ura4-D18 cdc10-M17 rad3::ura4*	Lab Stock
yFS280	*h+ leu1-32 ura4-D18 ade6-210 cdc10-M17*	Lab Stock
yFS357	*h+ leu1-32 ura4-D18 his3-237 ura4 adh1:cdc25*	Russell Lab (Russell and Nurse 1986)
KGY14	*h- leu1-32 ura4-D18 cdc2::ura4 cdc2-Y15F LEU2*	Gould Lab (Gould and Nurse 1989)
yFS430	*h- leu1-32 ura4-D18 ade6-210 his3-237 cdc10-M17 ura4 adh1:cdc25*	This study
yFS445	*h- leu1-32 ura4-D18 cdc2::ura4 cdc2-Y15F LEU2 cdc25::ura4*	This study
yFS437	*h+ leu1-32 ura4-D18 cdc2::ura4 cdc2-Y15F LEU2 cdc10-M17*	This study

Dickinson) by measuring the mean of the S-phase peak as a percentage of the position of between the means of the 1C and 2C controls.

Asynchronous Experiments

Asynchronous experiments were carried out as described (Kumar and Huberman 2004), except that flow cytometry was carried out using the isolated nuclei protocol described above. Briefly, cells were grown to an O.D. of 1.0, diluted to an O.D. 0.1 and allowed to recover for 1 hour. At this time, the culture was divided and treated as described. Samples were collected after every hour, fixed by 70% ethanol and processed for flow cytometry.

Synchronous Experiments

We used centrifugal elutriation to synchronize cells either in G1 or G2. Since fission yeast spends a short time in G1, experiments were conducted in *cdc10-M17* background to synchronize cells in G1. Cultures were grown to O.D. 0.5 arrested at 35°C for 1.5 hours and then synchronized by elutriation. The culture was divided and treated with 0.03% MMS, 10 mM hydroxyurea (HU) or mock treated. The cells were kept at 25°C and samples collected after every 20 minutes for 3 hours.

For G2 synchronization, cultures were grown to O.D. 1.0 and elutriated. The synchronized samples were divided and treated with 0.015% MMS, 10 mM hydroxyurea (HU) or mock treated. Cells were collected after every 20 minutes and processed for flow cytometry.

RESULTS

We employed flow cytometry to assay the S-phase checkpoint response of fission yeast to DNA damage. To reduce cytoplasmic background and thus increase sensitivity, we performed our analyses on isolated nuclei (Carlson *et al.* 1997). Initially, we examined the response of asynchronous cultures (Kumar and Huberman 2004). Fission yeast spend most of their cell cycle in G2, therefore the cytometry profile of an asynchronous culture is largely 2C, with a small 1C and S-phase population (Figure II.1A). The alkylating agent methyl methane sulfonate (MMS), which produces DNA damage in the form of base adducts, was used to induce DNA damage and activate the checkpoint. MMS damage is most efficiently recognized during replication, and therefore preferentially activates the S-phase checkpoint, rather than the G2 checkpoint. However, at 0.03% MMS, the standard concentration used in previous synchronous checkpoint experiments (Lindsay *et al.* 1998), a significant fraction of cells in an asynchronous culture arrest in G2 (our unpublished observation). Therefore, for these experiments we used 0.015% MMS, a concentration used in previous asynchronous experiments (Kumar and Huberman 2004). Hydroxyurea (HU), a ribonucleotide reductase inhibitor which arrests cells in the early S-phase by depleting deoxynucleotides, was used as a control for cells containing close to 1C DNA content. As previously reported, wild type cells respond to MMS treatment by accumulating as sub-2C cells, presumably due to slowing of bulk replication (Figure II.1A) (Lindsay *et al.* 1998; Kumar and Huberman 2004). In

Figure II.1 S-phase DNA damage checkpoint analysis in asynchronous cells

(A). A mid-log, asynchronous cultures wild type culture (yFS104) was split three ways and incubated in the presence or absence of 0.015% MMS or 10 mM HU; samples were taken for flow cytometry every hour.

(B) Asynchronous cultures of wild type (yFS104), *rad3*Δ (yFS189), *adh1:cdc25* (yFS357) and *cdc2-Y15F* (KGY14) were treated and collected for flow cytometry as in panel A; for clarity, only the MMS treated samples are shown.

Figure II.1A S-phase DNA damage checkpoint analysis in asynchronous wild type cells

1A

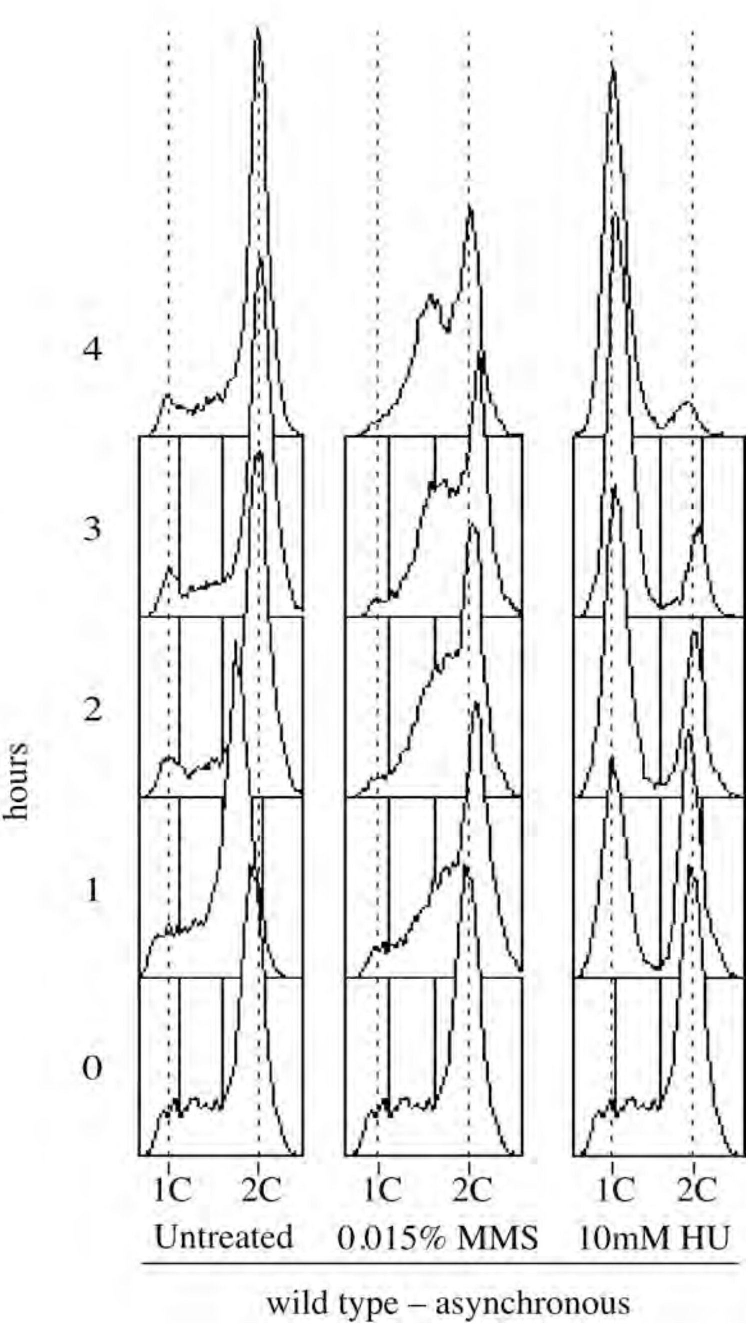
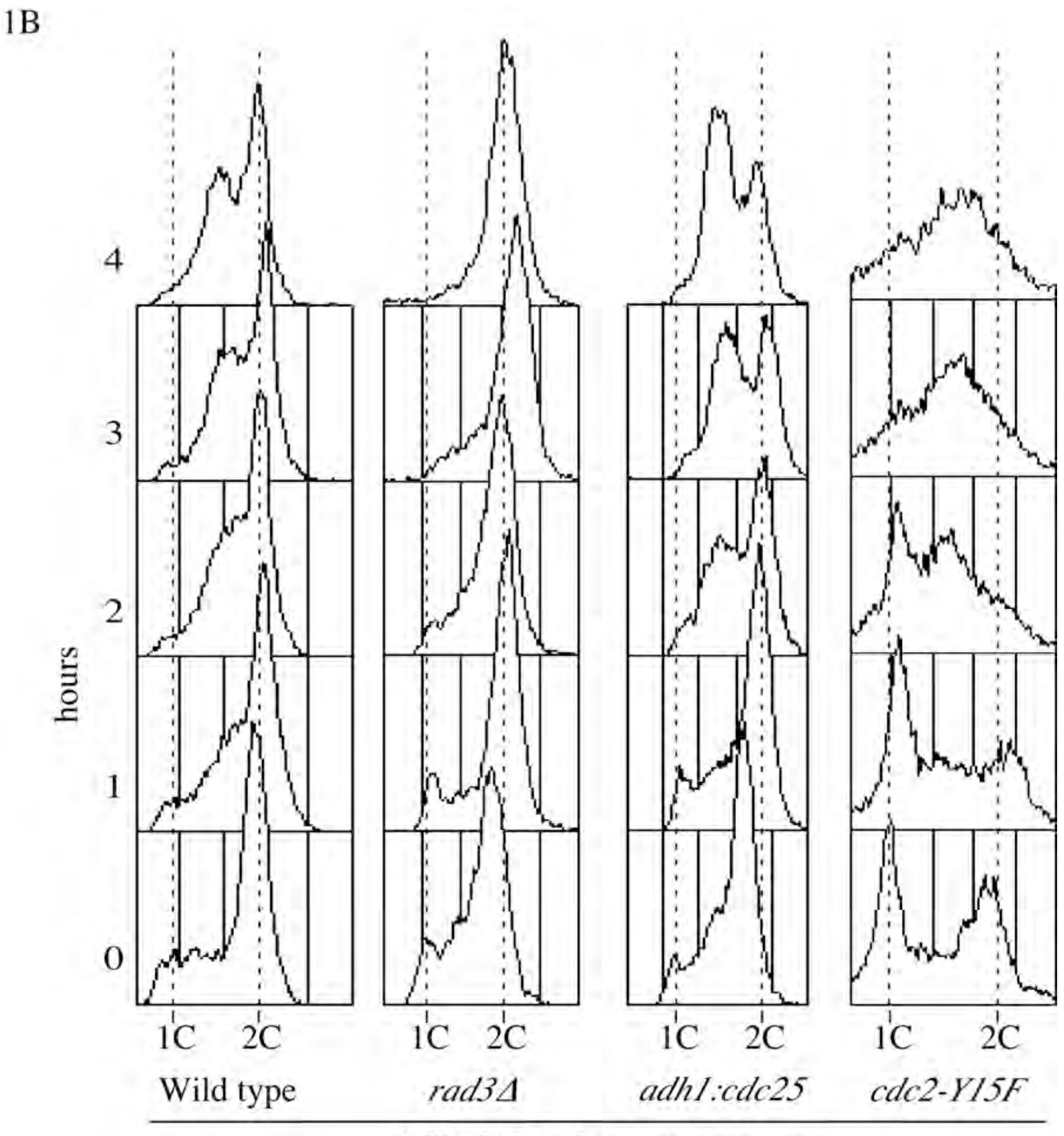


Figure II.1B S-phase DNA damage checkpoint analysis in asynchronous mutant cells



contrast, the *rad3*Δ strain, which is DNA damage checkpoint defective, showed no significant accumulation of sub-2C cells (Figure II.1B).

Although the asynchronous experiments show a robust checkpoint-dependent accumulation of sub-2C cells, it is difficult to infer cell-cycle kinetics from asynchronous experiments. To more carefully examine the effect of DNA damage on replication, we used synchronous cultures to analyze the progression of cells through S phase in the presence and absence of MMS. We synchronized cells by centrifugal elutriation, which isolates the smallest cells in a culture. Since in fission yeast cytokinesis is coincident with S phase, the smallest, newborn cells are in early G2. Thus, after elutriation we can follow a synchronous G2 population through mitosis into G1 and through S phase back to G2. As in the asynchronous experiment, we used 0.015% MMS because 0.03% MMS causes a significant fraction of the culture to arrest in G2 (our unpublished observation). Most untreated cells replicated between 80 and 120 minutes post elutriation (Figure II.2A). The MMS-treated cells begin replicating about the same time as untreated cells but do not complete replication by 180 minutes. This MMS-induced slowing is abrogated in *rad3*Δ, confirming that it is a checkpoint response (Figure II.2B).

As a third approach, we synchronized cells in G1. G1 synchronization has two advantages: the cells are past the G2/M transition, allowing us to use 0.03% MMS without evoking the G2 checkpoint, and the cultures are more

Figure II.2 S-phase DNA damage checkpoint analysis in G2 synchronized cells.

(A) Wild-type cells (yFS104) were synchronized in G2 by centrifugal elutriation, 0.015% MMS or 10 mM HU were added immediately and samples were collected every 20 minutes for flow cytometry. (B) G2 synchronized cultures of wild type (yFS104), *rad3*Δ (yFS189), *adh1:cdc25* (yFS357) and *cdc2-Y15F* (KGY14) were treated and collected for flow cytometry as in panel A; for clarity, only the MMS treated samples are shown. (C) *cdc25*Δ *cdc2-Y15F* (yFS445) cells synchronized in G2, 0.015% or 0.03% MMS or 10 mM HU was added immediately and samples were collected every 20 minutes for flow cytometry.

Figure II.2A S-phase DNA damage checkpoint analysis in G2 synchronized wild type cells.

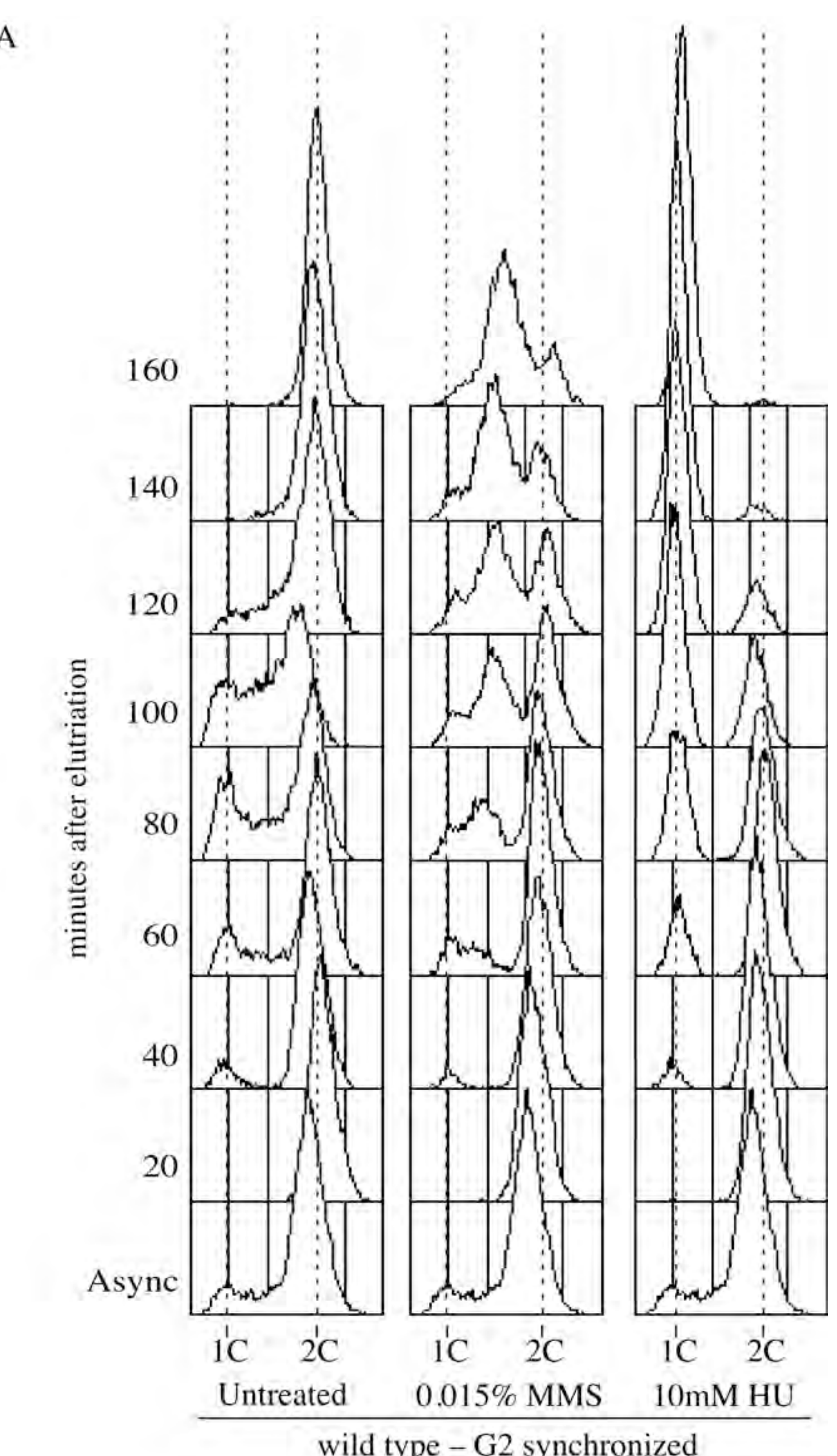


Figure II.2B S-phase DNA damage checkpoint analysis in G2 synchronized mutant cells.

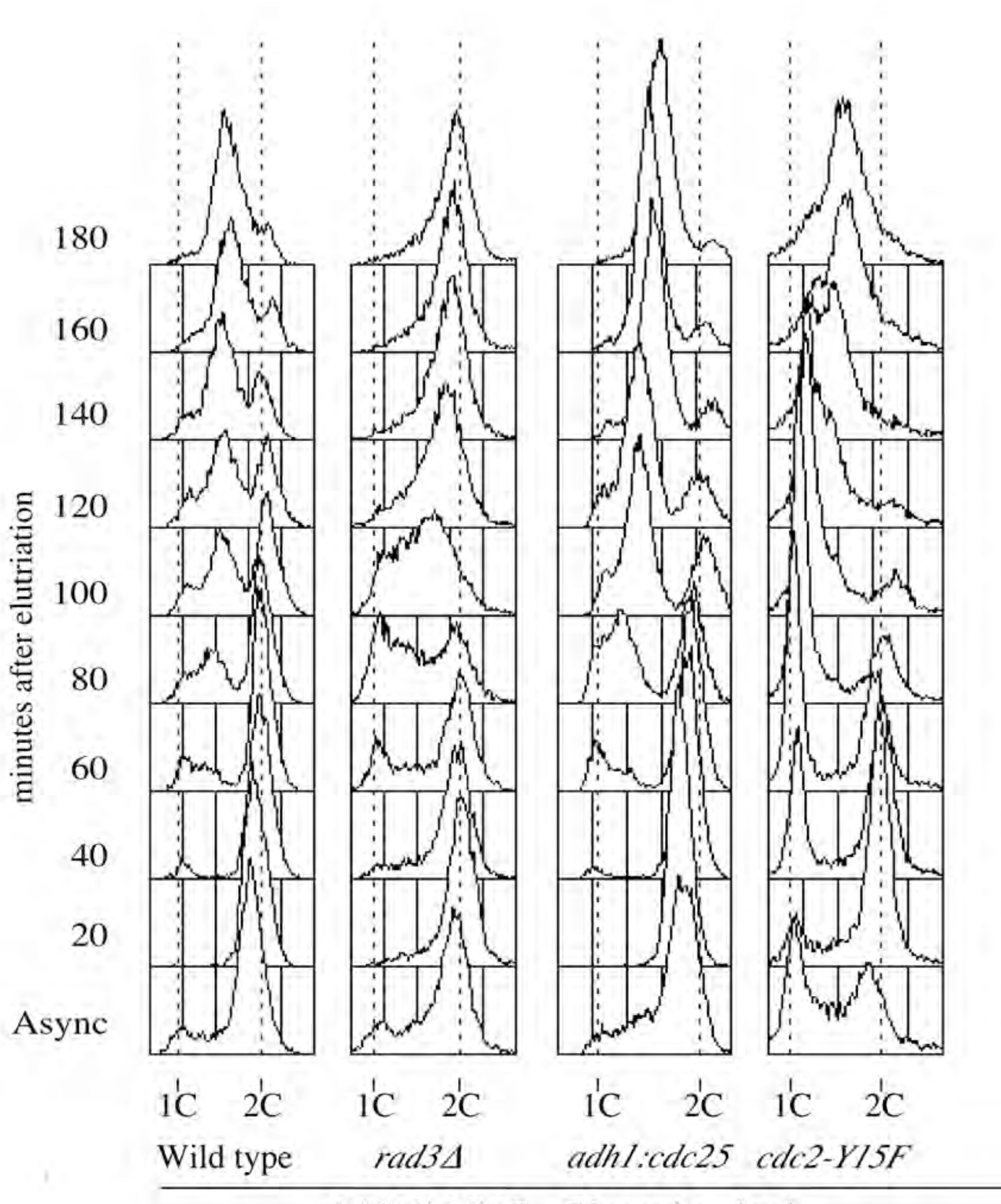
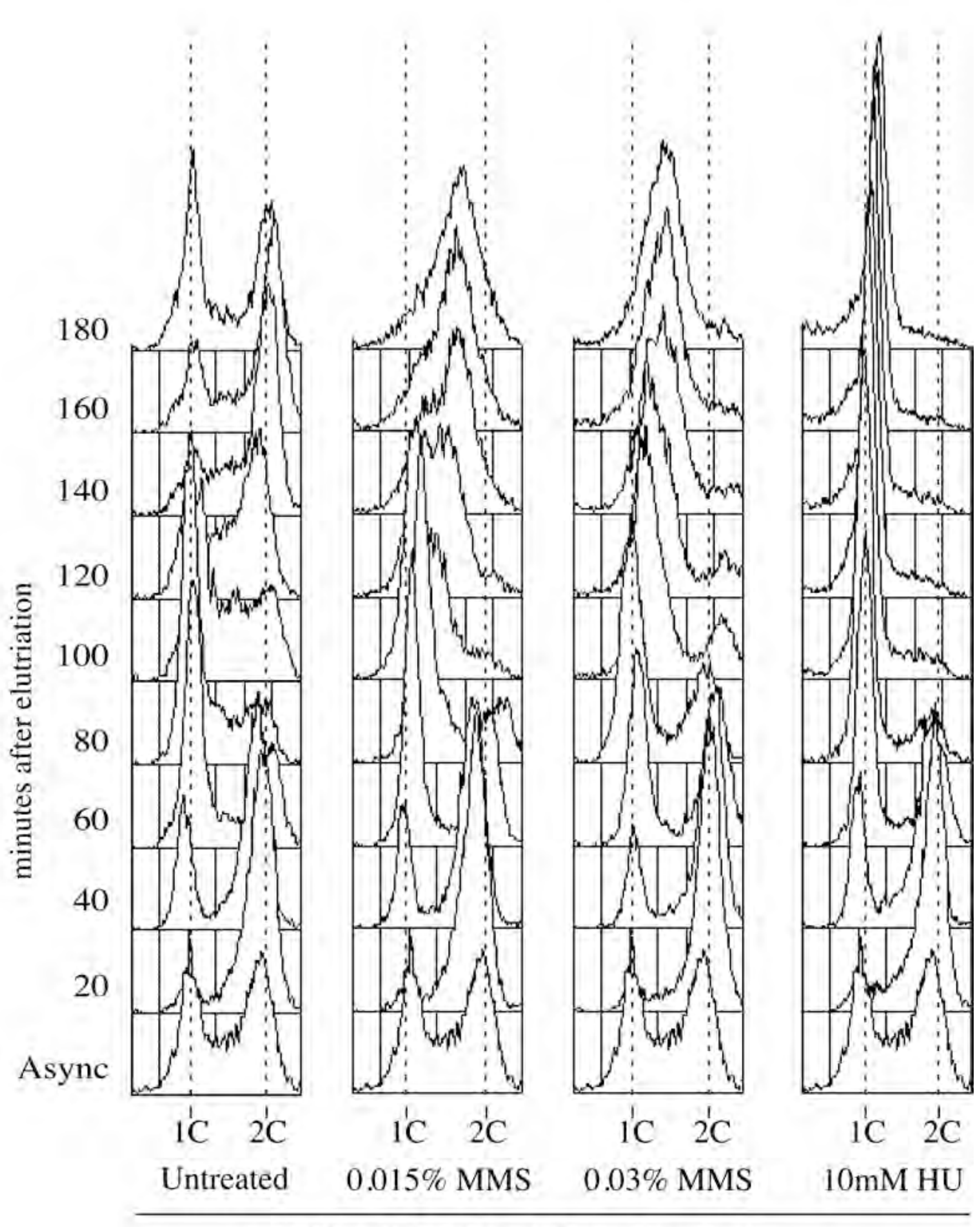


Figure II.2C S-phase DNA damage checkpoint analysis in G2 synchronized *cdc25∆ cdc2-Y15F* cells.



synchronous, allowing for meaningful quantitation (Figure II.3C). We employed a *cdc10-M17ts* temperature sensitive allele, which at 35˚C inactivates the fission yeast S-phase transcription factor, to block cells in G1. To avoid prolonged G1 arrest, we incubated asynchronous *cdc10-M17ts* cells at 35˚C for 90 minutes, and selected the smallest cells by elutriation. These cells will have just divided, and thus only recently entered G1. We estimate that the cells we isolate spend about 30 minutes arrested in G1 before they are released at the beginning of the time course. After elutriation, we observed a 1C peak showing that cells were arrested in G1. Cells were then released into the cell cycle and S-phase progression was assayed by flow cytometry (Figure II.3A). Untreated cells replicated between 40 and 80 minutes after release. MMS-treated cells did not complete replication by 180 minutes, and this slowing was dependent on Rad3 (Figure II.3B, C).

Over-expressing Cdc25 fails to override the S-phase DNA damage checkpoint

As an initial test of the role of Cdc25 in the S-phase DNA damage checkpoint, we examined if we could override the checkpoint by over-expressing Cdc25. Such over-expression efficiently overrides the Cdc25-dependent replication checkpoint arrest in G2 (Enoch and Nurse 1990). We used a strain in which Cdc25 was over-expressed from the strong, constitutive *adh1* promoter (Russell and Nurse 1986). If no difference in the cytometry profiles of cells with or without damage was seen, it would indicate that over-expressing Cdc25 had

overcome the S phase DNA damage checkpoint. However, in asynchronous culture, we observed sub-2C DNA content in the presence of damage, indicating that the checkpoint was still active (Figure II.1B). In fact, *adh1:cdc25* cells accumulate in a sub-2C population to a greater extent than wild-type cells, presumably because some of the wild-type cells arrest in G2, while the *adh1:cdc25* cells, lacking the G2 checkpoint, do not.

We also observed an MMS-induced delay of S-phase progression in synchronized *adh1:cdc25* cells. In both G2 and G1 synchronous experiments, the wild-type and *adh1:cdc25* strains demonstrated a similar degree of MMS-induced slowing of replication (Figures II.2B, 3B and 3C). Results from these experiments indicate that Cdc25 over-expression is not sufficient to override the S-phase damage checkpoint.

Inhibitory phosphorylation of Cdc2 is not required for the S-phase DNA damage checkpoint

Although the Cdc25 over-expression results suggest Cdc25 inhibition is not the mechanism for slowing of S phase, it is possible that the S-phase DNA damage checkpoint is able to inhibit even the over-expressed Cdc25. To directly test the role of Cdc2 tyrosine-15 phosphorylation in the checkpoint, we used an allele of *cdc2*, *cdc2-Y15F*, in which tyrosine-15 is mutated to phenylalanine, preventing its phosphorylation. Because Cdc2-Y15F cannot be inhibited by

tyrosine phosphorylation, it should bypass any Cdc25-dependent S-phase checkpoint, in the same manner that it overrides the G2 checkpoints (Rhind *et al.* 1997; Rhind and Russell 1998). Contrary to that prediction, asynchronous *cdc2-Y15F* cells treated with 0.015% MMS accumulated in a sub-2C peak, showing no defect in the checkpoint. As with the *adh1:cdc25* cells, *cdc2-Y15F* cells actually accumulate as sub-2C cells to a greater extent than wild-type cells, presumably due to the lack of a G2 checkpoint (Figure II.1B).

Synchronous experiments using the *cdc2-Y15F* strain also showed no defect in S-phase slowing in response to DNA damage. Because the Cdc2-Y15F cannot be inhibited by tyrosine phosphorylation, *cdc2-Y15F* cells go very quickly through G2. They compensate for this short G2 by expanding G1; this effect can be seen in the large G1 peak in asynchronous *cdc2-Y15F* cells. Thus *cdc2-Y15F* cells begin replication later than wild-type cells. Untreated G2 synchronized *cdc2-Y15F* cells began replicated around 100 minutes and completed replication by 160 minutes (data not shown). In the presence of MMS, cells started replicating at the same time as untreated samples but did not complete replication by 180 minutes (Figure II.2B).

G1 synchronized *cdc2-Y15F* cells begin replication at the same time as wild-type cells, because the arrest is at the end of G1. The MMS induced S-phase slowing is comparable between *cdc2-Y15F* and wild-type cells, beginning at around 60 minutes and not finishing by 180 minutes (Figures II.3B and 3C).

Figure II.3 S-phase DNA damage checkpoint analysis in G1 synchronized cells

(A) Flow cytometric analysis of S-phase DNA damage checkpoint in G1 synchronized *cdc10-M17ts* cells (yFS280). 0.03% MMS or 10 mM HU were added immediately after elutriation and samples collected after every 20 minutes. (B) G1 synchronized cultures of *cdc10-M17ts* (yFS280), *cdc10-M17ts rad3Δ* (yFS260), *cdc10-M17ts adh1:cdc25* (yFS430) and *cdc10-M17ts cdc2-Y15F* (yFS437) were treated and collected as in panel A; for clarity, only the MMS treated samples are shown

(C) Quantification of the data of A and B. The previously reported minor, checkpoint independent slowing is evident in the *rad3Δ* culture (Rhind and Russell 1998). Each point is the average of two experiments; the error bars represent the range of the data.

Figure II.3A S-phase DNA damage checkpoint analysis in G1 synchronized wild type cells

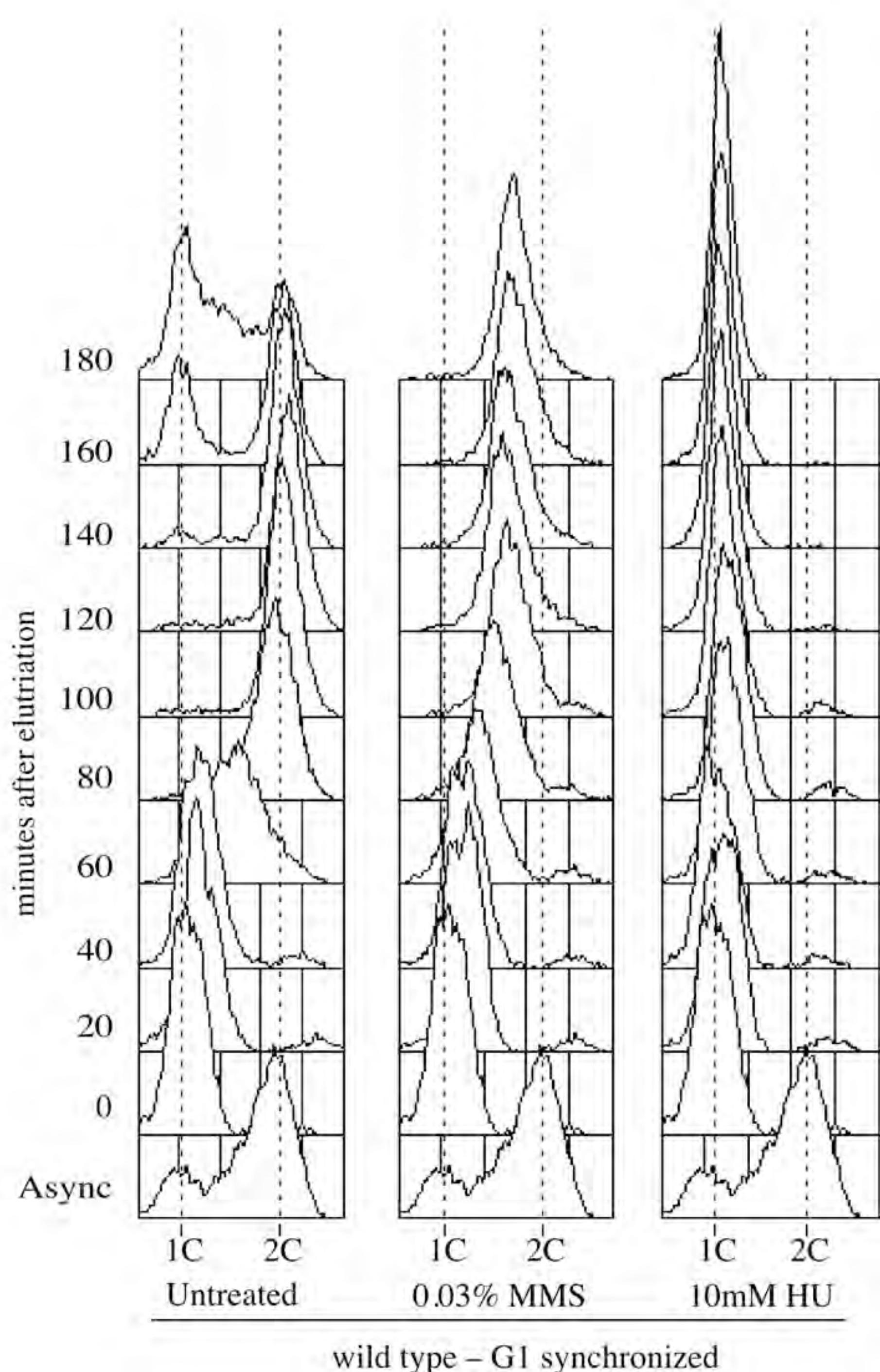


Figure II.3B S-phase DNA damage checkpoint analysis in G1 synchronized mutant cells

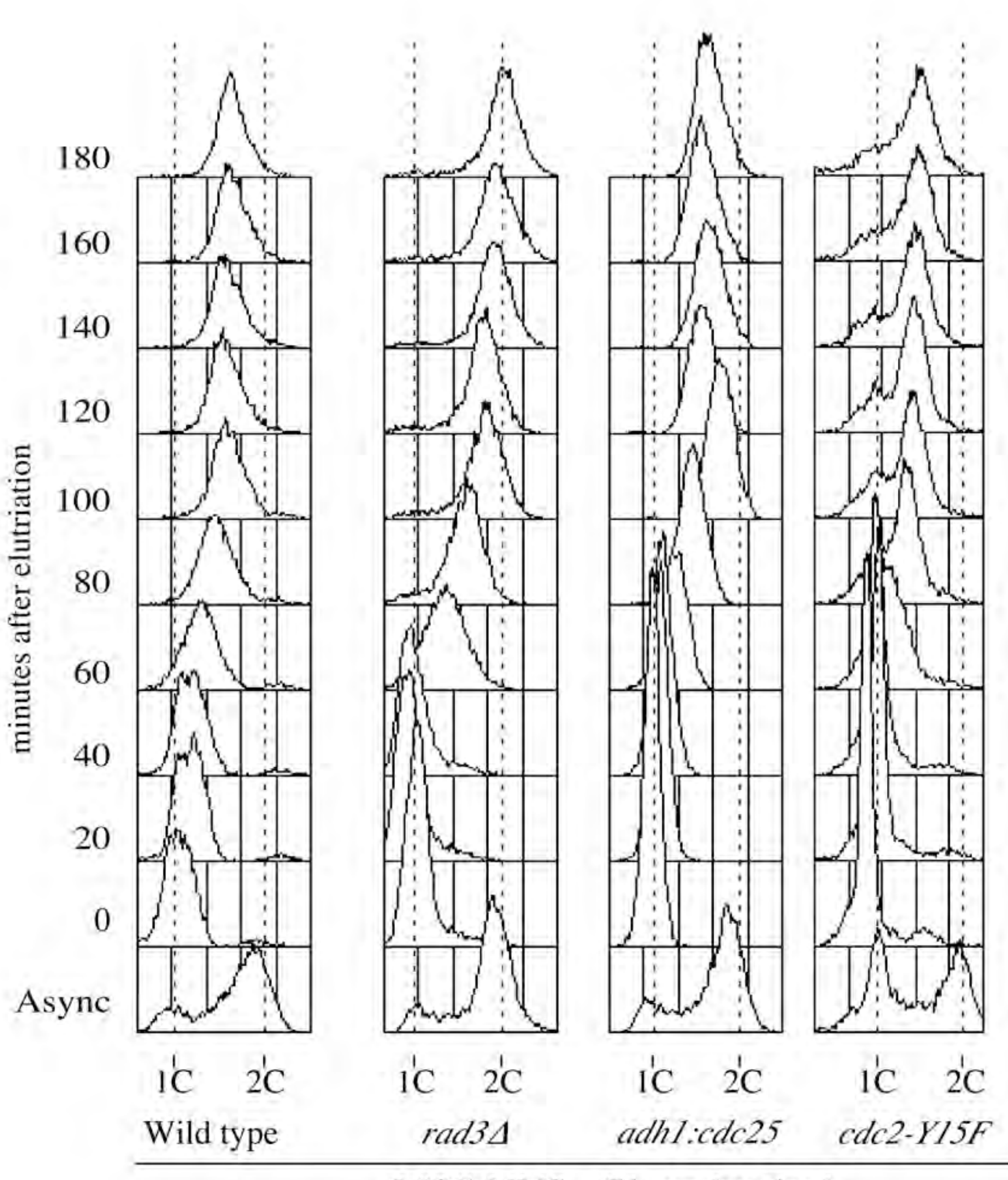
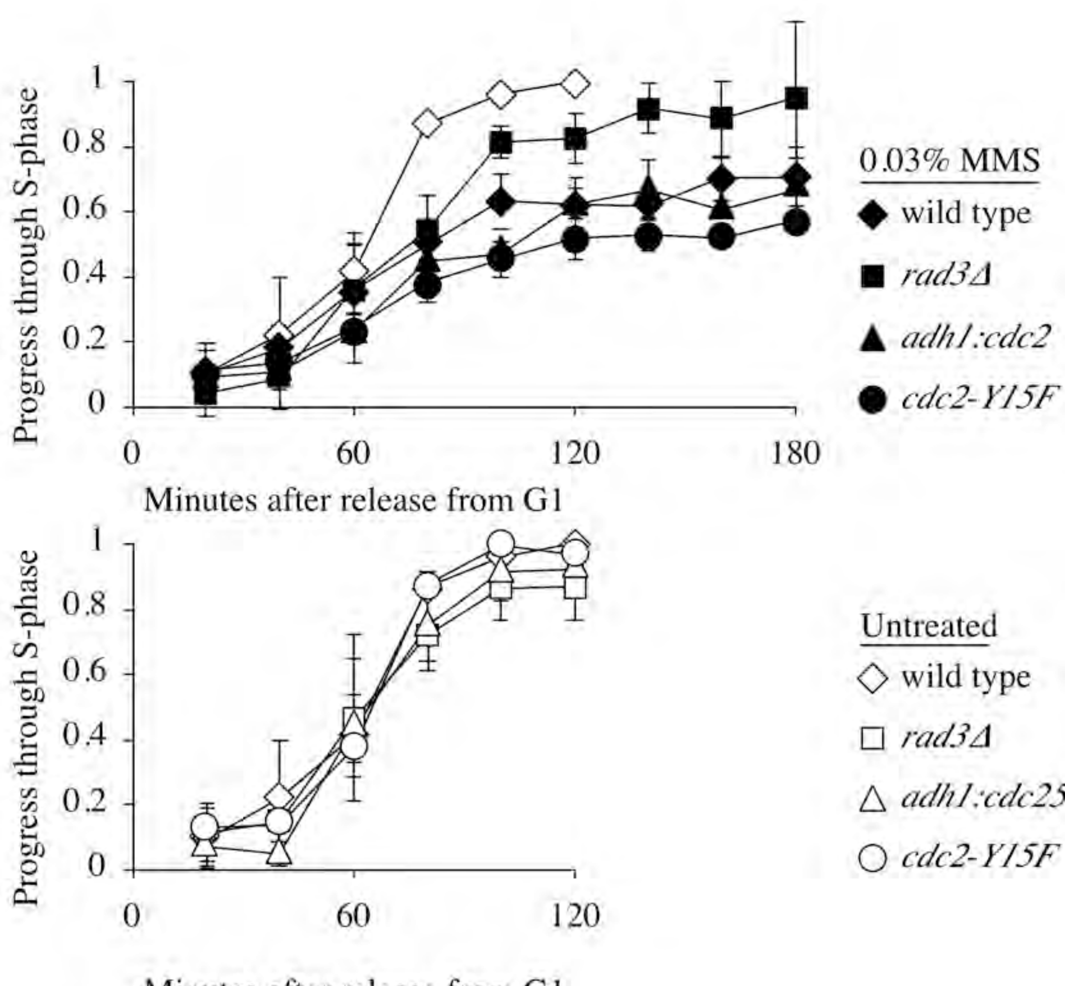


Figure II.3C Quantification of slowing in G1 synchronized cells

3C



These results show that Cdc2 tyrosine-15 phosphorylation is not required for cells to slow replication in response to MMS-induced DNA damage.

Cdc25 is not required for the checkpoint

The previous results show that Cdc2 tyrosine-15 phosphorylation is not required for the S-phase DNA damage checkpoint, but they leave open the possibility that Cdc25 is required to regulate another target besides Cdc2. To test this possibility directly, we wanted to study the S-phase progression in the absence of Cdc25. Since Cdc25 is an essential gene, we created a strain in which the essential function of Cdc25 - the dephosphorylation of Cdc2 - is bypassed by Cdc2-Y15F. *cdc25Δ cdc2-Y15F* cells were synchronized in the G2 phase by elutriation and their progress through S-phase in the presence or absence of MMS was monitored. Since these cells lack a G2 checkpoint, we were able to use 0.03% MMS without arresting the cells in G2; we also used 0.015% for comparison with the other G2 synchronization experiments.

As for *cdc2-Y15F* cells, untreated *cdc25Δ cdc2-Y15F* cells replicate later than wild-type, in this case between about 80 and 140 minutes (Figure 2C). *cdc25Δ cdc2-Y15F* cells treated with 0.015% MMS did not complete replication by 180 minutes; cells treated with 0.03% MMS replicated even more slowly. *cdc25Δ cdc2-Y15F cdc10-M17* cells are inviable, precluding G1 synchronization. These results show that S-phase damage checkpoint operates normally in the absence of Cdc25.

***Cdc25Δ nmt1:Pyp3* cells have a partial slowing in the presence of damage**

To study the effect of *cdc25Δ* we also used a phosphatase overexpression, which has been shown to rescue *cdc25Δ* lethality. Cdc2 is dephosphorylated by Pyp3 as well as Cdc25 (Millar *et al.* 1992). Pyp3 is a tyrosine phosphatase and plays a minor role in mitotic control. Pyp3 is not regulated by the G2 DNA damage checkpoint (Rhind and Russell 2001). Pyp3 overexpression rescues *cdc25Δ* lethality. In our studies *pyp3* was overexpressed by using the strong *nmt1* promoter for continuous expression of Pyp3.

Asynchronous experiments were performed for Pyp3 overexpression strain with or without Cdc25. Pyp3 overexpression cells, where *cdc25* has not been deleted, slowed S phase in the presence of damage similar to wild type fission yeast. However, DNA damage had only a slight slowing of S phase in a *cdc25Δ* background (Figure II.4). G2 synchronized cells slowed replication in the presence of damage even when *cdc25* was deleted. The slowing seems to be partial compared to wild type and similar to the asynchronous data (Figure II.5).

Our results are consistent with a strictly Cdc2-Y15 phosphorylation independent mechanism for the S-phase DNA damage checkpoint. However, results from the pyp3 studies are not consistent with our conclusions about Cdc25 is not the target of the S phase DNA damage checkpoint. We talk about the possible explanations for this result is the discussion.

Figure II.4 S-phase DNA damage checkpoint analysis in asynchronous Pyp3 cells.

Flow cytometric analysis of S-phase DNA damage checkpoint in asynchronous *nmt1:pyp3* cells (yNR248), *nmt1:pyp3 cdc25*Δ (yNR253). A mid-log, asynchronous culture was split three ways and incubated in the presence or absence of 0.015% MMS; samples were taken for flow cytometry every hour. For clarity, only the MMS treated samples are shown. The data shows that *nmt1:pyp3* slow in the presence of 0.015% MMS whereas *nmt1:pyp3 cdc25*Δ slows only partially.

Figure II.4 S-phase DNA damage checkpoint analysis in asynchronous Pyp3 cells.

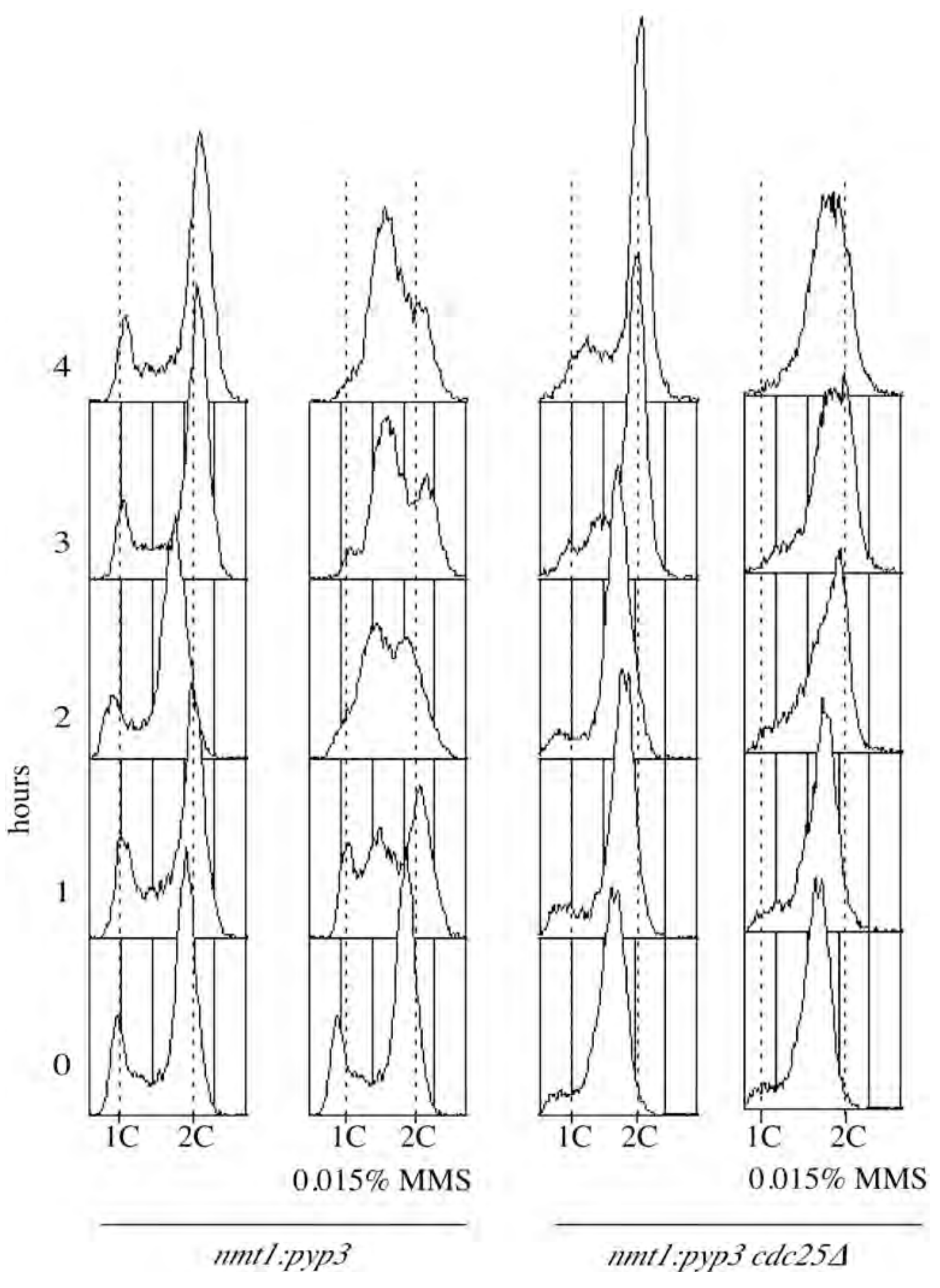


Figure II.5 S-phase DNA damage checkpoint analysis in G2 synchronized Pyp3 cells

Flow cytometric analysis of S-phase DNA damage checkpoint in G2 synchronized *nmt1:pyp3* cells (yNR248), *nmt1:pyp3 cdc25*Δ (yNR253). Data is shown for cells in the absence and presence of 0.015% MMS panel. MMS was added immediately after elutriation and samples collected after every 20 minutes. The data shows that *cdc25* deletion leads to a partial slowing in the presence of 0.015% MMS.

Figure II.5 S-phase DNA damage checkpoint analysis in G2 synchronized Pyp3 cells

DISCUSSION

We have investigated the role of inhibitory tyrosine-15 phosphorylation of Cdc2 and of the Cdc2 tyrosine-15 phosphatase, Cdc25, in the S-phase DNA damage checkpoint. Inhibition of Cdc25, and thus inhibition of Cdc2 tyrosine-15 dephosphorylation, is the mechanism by which fission yeast arrest in G2 in response to DNA damage or replication blocks (Rhind *et al.* 1997; Rhind and Russell 1998). Recent work has suggested that a similar mechanism may also slow replication in response to DNA damage (Kumar and Huberman 2004). We have tested this idea and the bulk of our results indicate no involvement of Cdc25 or Cdc2 tyrosine phosphorylation in the fission yeast S-phase DNA damage checkpoint.

Neither the overexpression of Cdc25, nor the mutation of tyrosine-15 to an unphosphorylatable phenylalanine, impairs the S-phase checkpoint; yet both override the G2 checkpoint (Enoch and Nurse 1990; Rhind *et al.* 1997; Rhind and Russell 1998). Furthermore, cells lacking both Cdc25 and tyrosine-15 of Cdc2 slow replication normally in response to MMS-induced DNA damage. This result rules out checkpoint mechanisms that involve Cdc2-independent targets of Cdc25, and Cdc25-independent regulation of Cdc2 tyrosine-15 phosphorylation.

Although the above cited results support our conclusion, cells lacking Cdc25 in a Pyp3 background seem to have only a partial slowing of replication in the presence of DNA damage. This result is not consistent with our observation

that deleting cdc25 does not have an effect on the slowing of replication. There are a couple of possibilities that can explain the contradictory results seen for *cdc25* deletion in a *cdc2-Y15F* and the Pyp3 overexpression background.

First, there can be something wrong with the genotype in *nmt1: pyp3 cdc25*Δ strain which is having an effect on the results. There can be a mutation in one of the upstream checkpoints leading to a lack of slowing in the presence of damage. Second, data from the *nmt1: pyp3 cdc25*Δ suggests that there is a Cdc25 dependent target which is independent of Pyp3. It is possible that Cdc25 has a role in S-phase slowing in a Cdc2-Y15 independent manner. However, replication slowing in the *cdc25*Δ Cdc2-Y15F strain rules out such a possibility. Slowing of replication in the presence of damage when Cdc25 is overexpressed is also contradictory to this theory. Third, slowing of replication in the *Cdc2-Y15F* background can be due to some background mutations. Mutating Cdc2 maybe having an effect on the S phase progression of the cells. However normal progression through S phase in unperturbed cells is inconsistent with such a possibility. To completely rule out this possibility, other Cdc2 mutants can be used for looking at the effect of inducing DNA damage.

Finally, Cdc2 may also be slowing replication at the site of DNA damage in a checkpoint independent manner. To show that the slowing of replication in the mutant strains is checkpoint dependent, these experiments can be done in rad3Δ and cds1Δ cells. Majority of our data shows that Cdc25 is not the target of the checkpoint and it is hard to reconcile the Pyp3 data with the other experiments.

We have drawn our conclusions based on the known literature and the majority of the data showing that Cdc25 is not the target of the checkpoint. It is possible that a more complicated model may exist.

These results contradict those of Kumar and Huberman, who, using similar approaches, concluded that *adh1:cdc25* and *cdc2-Y15F* cells lack the S-phase DNA damage checkpoint. There is technical difference between the two studies that may explain the discrepancy. Kumar and Huberman used a whole-cell flow cytometry protocol, in which cytoplasmic background contributes significantly to the total signal, reducing the sensitivity of the assay. We used an isolated-nuclei protocol, which removes the cytoplasm before analysis. This approach greatly increases the resolution of the assay and allows for quantitation of the data. In addition, for their G1 synchrony experiments, cells were arrested in G1 for up to 4 hours, which allows the cells to elongate, further reducing the sensitivity of the whole-cell assay. We find that the combination of the four hour arrest and the whole cell flow-cytometry analysis compromises detection of the checkpoint delay (Nick willis, unpublished results). We used centrifugal elutriation to isolate cells that had been arrested for only 30 minutes, allowing for a more sensitive analysis of the checkpoint. We believe that these technical differences are responsible for the different conclusion drawn.

Recently, another paper from the same lab gave a number of reasons for the discrepancy between the two datasets (Kumar and Huberman 2009). Use of

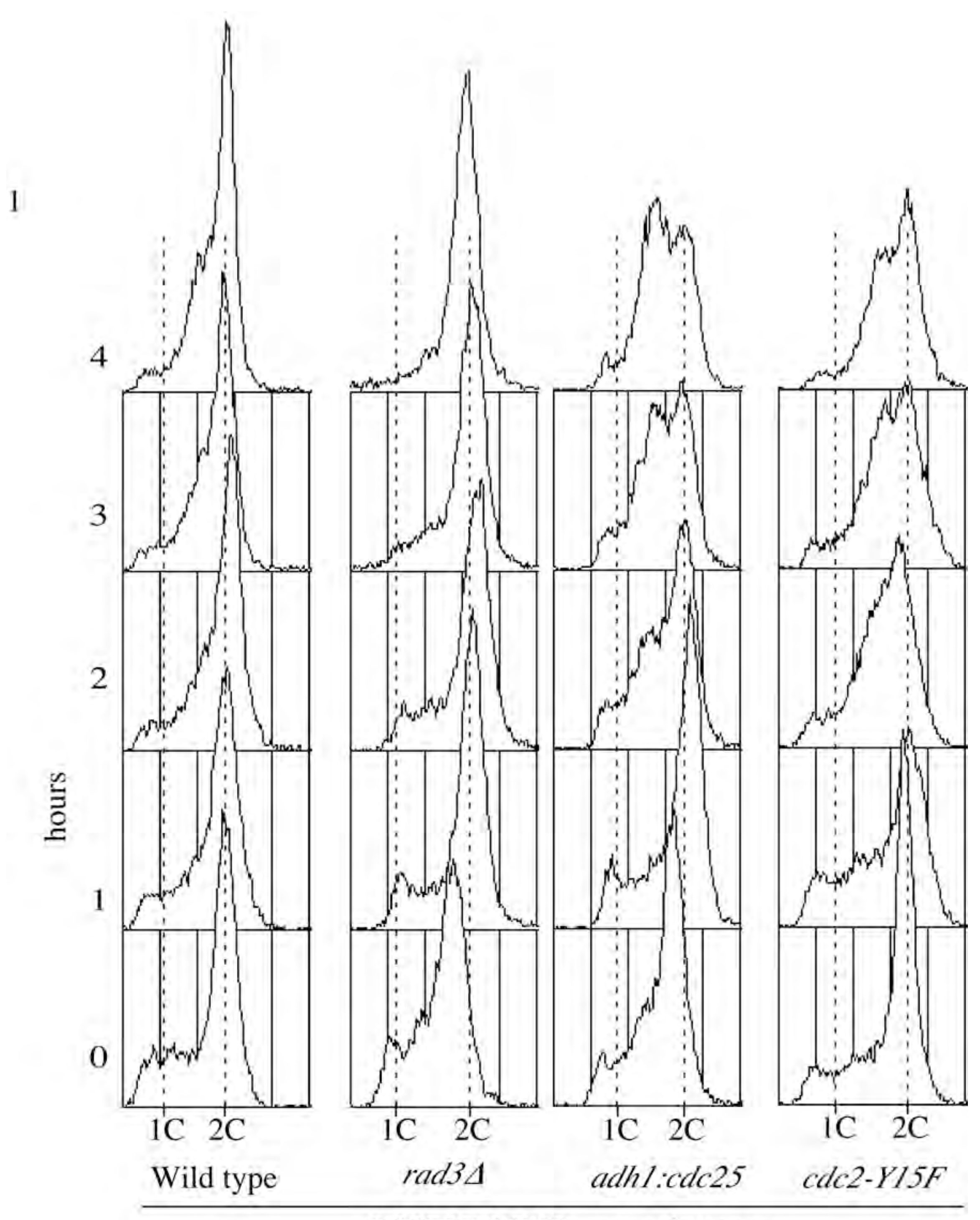
a higher concentration of MMS (0.03%) in some of our experiments was cited as a reason for slowing of S phase in the mutant strains. Figure II.6 shows that even at 0.0075% MMS concentration the level of slowing is similar in the mutant strains compared to the wild type. The data from the G2 elutriation experiments at the lower concentration also show a similar effect of slowing in the mutant strains (Kommajosyula and Rhind 2006). Hence, using 0.03% for a subset of the experiments does not effect our conclusions. The data for 0.015% was not presented in our paper, since the cells have a shorter time for entering S phase from G1 synchrony, making the response to damage subtle as opposed to a more pronounced effect at a higher concentration. Indeed, previous work from our lab has checked the effect of different concentrations of MMS for inducing checkpoint effects and we find that 0.03%MMS works best for studying checkpoints in our hands. Nonetheless, as shown in figure II.1.1 all doses of MMS have the same pattern of slowing.

Batch to batch variation in effective concentrations of MMS was also cited as a reason for the results. However, we have used the same absolute concentration of 0.03% MMS in our lab over the last several years and have observed no such variations. The fact that Kumar et al. have to use varying concentrations to observe a slowing of S-phase seems to suggest an issue with their FACS preparation, an observation that we elaborated upon in the discussion of our paper.

Figure II.6 S-phase DNA damage checkpoint analysis in asynchronous cells using low MMS

Flow cytometric analysis of S-phase DNA damage checkpoint in asynchronous wild type (yFS104), *rad3*Δ (yFS189), *adh1:cdc25* (yFS357) and *cdc2-Y15F* (KGY14). The 0.0075% MMS panel is from the same experiments shown in Figure II.1. A mid-log, asynchronous culture was split three ways and incubated in the presence or absence of 0.0075% MMS; samples were taken for flow cytometry every hour. For clarity, only the MMS treated samples are shown. The data shows that *adh1:cdc25* and *cdc2-Y15F* slow in the presence of 0.0075% MMS.

Figure II.6 S-phase DNA damage checkpoint analysis in asynchronous cells using low MMS



Another objection made in the paper was our observation that MRN is a target of S-phase damage checkpoint. Work from the Huberman lab suggests that MRN does not play a role in the damage checkpoint (Marchetti *et al.* 2002). The explanation given was that a higher concentration of MMS would lead to such an effect. However, in our opinion such a opposite result for the same strains between our labs can be explained by the technical difference in flow cytometry described in the discussion section of chapter II.

The Huberman lab claimed that the magnitude of inhibition of progression in our work was consistently higher for the same concentration of MMS used by both labs. These differences can be explained by a higher resolution of our FACS protocols. We were also able to observe significant slowing in our asynchronous and G2 synchronized populations showing that the results we observed were not due to high MMS concentration (Kommajosyula and Rhind 2006).

The paper also argues that *cds1* deletion should have been tested for an effect on MMS induced checkpoints instead of *rad3* deletion, which we used. Studies from our lab show that *cds1*Δ has the same effect as a *rad3*Δ and there is a complete absence of S phase slowing in the presence of MMS (Nick Willis, unpublished data).

Our results implicate a Cdc2 tyrosine phosphorylation independent target of the S-phase DNA damage checkpoint in fission yeast. Precedent for such a

target exists. In mammals, the S-phase checkpoint appears to have two branches: one which acts through inhibition of Cdc25A to inhibit origin firing, and one which is Cdc25-independent and requires the MRN complex and XRCC3 (Falck *et al.* 2002; Henry-Mowatt *et al.* 2003). Since the fission yeast checkpoint requires MRN, it may be mechanistically similar to the MRN-dependent branch of the mammalian checkpoint (Chahwan *et al.* 2003). Little is known about the mechanism or purpose of the MRN-dependent branch of the pathway, except that its loss leads to severe cancer-predisposition in humans (Petrini 2000). The possible role of MRN and XRCC3 in slowing replication fork progression through induction of replication-coupled recombinational repair provides a model that can be tested in fission yeast (Rhind and Russell 2000). Whatever the case, fission yeast provides an tractable system for the study of this checkpoint mechanism.

Chapter III

Measure if origin efficiency increases during S phase

INTRODUCTION

Replication needs to be completed with high fidelity and in a finite amount of time to ensure a faithful transmission of genetic material to the next generation and to avoid genomic instability. Eukaryotes initiate replication from multiple sites known as origins (Dutta and Bell 1997; Gilbert 2001; Bell and Dutta 2002). Many origins are licensed but only a few fire. It is unclear how only a fraction of licensed origins are allowed to fire. Specific origins that fire during any one cell cycle do not necessarily fire during the subsequent round. Any origin that fires during a specific cell cycle does so in only a fraction of the total population. Fission yeast origins don't fire at the beginning of the S phase concurrently but rather fire throughout the S phase (Patel *et al.* 2006). *Xenopus* and *Drosophila* embryos are able to initiate replication on any DNA sequence (Hyrien and Méchali 1993; Shinomiya and Ina 1994). Similarly, in humans large regions appear to sustain replication initiation at random sites inside the sequence (Krysan and Calos 1991). Budding yeast have well defined origins that fire at particular times during the cell cycle and hence are defined as early or late firing origins (Kelly and Brown 2000; Gilbert 2001).

Genome-wide analysis of fission yeast origins was initially done bioinformatically (Segurado *et al.* 2003). Origins were found to contain AT rich tracts. However, unlike budding yeast, which has a well defined ARS consensus sequence (ACS), no consensus sequence was identified for fission yeast. Fission yeast origins are inefficient as shown in work from various labs including ours.

Studies from our lab have shown that origins fire in a stochastic manner in fission yeast (Patel *et al.* 2006). The efficiency of fission yeast origins averages about 30%. The origins are present in the intergenic region in general and they are about 1kb in length.

Random gap problem

Inefficient origins and random firing of origins will increase the time taken by cells to complete replication in metazoans and fission yeast. Theoretically, random firing may lead to long inter-origin gaps. These gaps may be too long to be passively replicated by replication forks traveling from neighboring origins in a timely manner. Such gaps would prevent cells from completing replication within the defined and short time of S phase. This phenomenon has been described as the random gap problem (Lucas *et al.* 2000; Herrick *et al.* 2002; Jun *et al.* 2004). Cells must overcome this potentially serious problem since replication is completed in a finite amount of time (Hyrien *et al.* 2003). However the mechanism by which cells overcome this potential problem is unclear.

Several models have been proposed to explain how cells complete replication (Legouras *et al.* 2006). The origin redundancy model suggests that there is an excess of origins distributed throughout the genome (Chapter I Figure I.4) (Legouras *et al.* 2006). These origins are closely spaced. Random origin firing would lead to some origins firing early and most being passively replicated and not firing. The origins that are not passively replicated would have the

potential to fire and thus solve the random gap problem. Using this model it is hard to explain efficient replication in fission yeast and mammals due to several restrictions including intergenic location of origins in fission yeast to chromatin structure and ongoing transcription in mammals.

The coordinated model describes a defined distance between origins that fire and prevent large unreplicated sequences from causing random gap problem during S phase (Hyrien and Méchali 1993; Blow *et al.* 2001). This model is based on plasmid studies in various organisms. An exponential distribution of inter-origin distances, where a majority of the fired origins are spaced close together but some are very large distances apart, shown by our lab disproves this model in fission yeast (Patel *et al.* 2006). Data from *Xenopus* embryos also prove that this model is incorrect (Herrick *et al.* 2000).

The increasing origin efficiency/Redistribution model describes an increase in the firing efficiency of origins as cells progress through S phase (Lucas *et al.* 2000). Thus, the efficiency of origins that have not fired or been passively replicated increases. Thus, origins present in regions in which no origins have fired are more likely to fire allowing the cell to overcome the random gap problem. We propose a mechanism to explain this model involves a rate-limiting factor, which allows only a subset of origins to fire at any given time. Since this rate-limiting factor is limited in quantity, it will only be able to activate a few origins at any given moment. Once these origins have fired, this factor freely

diffuses to provide firing elsewhere. We have shown recently that Dfp1, the catalytic subunit of Hsk1, is the rate-limiting factor which determines as to which origin fires during the S phase (Patel *et al.* 2008). The study to identify the rate-limiting step is shown in Appendix III.1.

The increasing origin efficiency model seems the most plausible mechanism by which fission yeast overcomes the random gap problem. An increase in origin firing through S phase is seen in budding yeast and *Xenopus* embryo extracts (Herrick *et al.* 2000; Lucas *et al.* 2000; Raghuraman *et al.* 2001; Yabuki *et al.* 2002; Goldar *et al.* 2009). Identifying Dfp1 as the rate-limiting factor responsible for determining which origins fire strongly supports this model. A similar model has been proposed to explain the efficiency of origins in *Xenopus* embryos (Lucas *et al.* 2000).

Although studies from our lab support the redistribution model, we need to measure if there is an increase in efficiency of a late firing origin in S phase to test our hypothesis. However, no late firing origins have been identified in fission yeast making it hard to look at a specific origin throughout our studies. Late replicating sequences have been identified (Kim and Huberman 2001). These late replicating sequences are passively replicated by replication forks traveling from distant early firing origins or by unidentified late firing origins. Due to these reasons, the chances of any one particular origin firing during late S phase are very low since it will be replicated before it gets a chance to fire.

To measure if the efficiency of an origin increases during the S phase we blocked passive replication of a late replicating sequence from neighboring origins, and forced that region to be replicated by a late/inefficient internal origin, which allowed us to test the efficiency of that origin. To block passive replication we used the replication termination sequence 1 (RTS1), a unidirectional fork-blocking sequence originally identified in the mating-type locus (Dalgaard and Klar 2001). RTS1 arrests forks coming from the proximal side. The fork arrest is replication termination factor 1(Rtf1) dependent and studies show that RTS1 arrests at the Mat locus and at other regions of the genome (Eydmann *et al.* 2008). RTS1 is 859bp in length with a 60bp partially conserved sequence motif (Codlin and Dalgaard 2003).

AT2062 is a late replicating sequence present on chromosome II and has a firing efficiency of about 10%. In this study we used RTS1 on either side of *AT2062* to prevent passive replication from neighboring origins and allow us to measure the firing efficiency of the origin. We are also integrating RTS1 to prevent passive replication of an early firing origin, *AT3003*. Various methods have been used to look at the firing efficiency of the origin.

Our studies show that there are a number of small very inefficient origins that seem to fire to ensure the replication of the region flanked by RTS1. We were unable to see origin firing at *AT2062*. The possible reasons for this are discussed later.

MATERIALS AND METHODS

All strains were grown in yeast extract with supplements (YES) at 25°C or 30°C and manipulated using standard methods (Forsburg and Rhind 2006).

Table III.1 - Strain list

Strain	Genotype	Source
yPP113	*h+ leu1-32 ura4-D18 rts1(R):hph AT2062 rts1(L):kanMX*	This study
yFS105	*h- leu1-32 ura4-D18*	Lab stock
yKN18	*h+ ura4-D18 ade6-M210 rts1(R):hph rts1(L):kanMX cdc25-22*	This study
yFS128	*h- leu1-32 ura4-D18 cdc25-22*	Lab stock
yNW239	*h- smt0 leu1-32 ura4-D18 his3-D1 cdc10-M17 sfr1::ura4 swi5::ura4*	From Nick Willis
yKN20	*h- ura4-D18 ade6-M210 rts1(R):hph rts1(L):kanMX rad51::nat*	This study
yKN22	*h+ leu1-32 ura4-D18 his7-366 AT3003 rts1(L):kanMX leu1 adh1 :hENT1 his7 adh1 :tk*	This study
yKN23	*h- leu1-32 ura4-D18 AT3003 rts1(R):hph leu1*	This study

	adh1 :hENT1 his7 adh1 :tk	

Strain construction

RTS1 was integrated on the left side of *AT2062* using PP192 (GTTTTAACTGTCAGCAATACTACACTACGCTATGATACTCCACGTTGCATAT CACTATATGTCACATGTTCACAATGTCGATGAATTCGAGCTCGTTTA) and PP193 (CTTATAACTGAACTGAGGGACGAGGTTCAGTTGTTCTCAATTTATAATATTT GAAGTAGTAAGAATTATATCTGATAGAACGCGGCCGCCAGCTGAAGC). RTS1 was integrated on the right side of *AT2062* using PP196 (ATAACAGCGTTTAAGAATTAGTTACTTATAAAGACCGAAGCGATCTTCCAGA TAATGAATAGCAATACATTAGATGTGAACGCGGCCGCCAGCTGAAGC) and PP197 (TATACCGTTGTAACAGCAAGAGCTTAATTGTTTCAACAATCCAACTTACGCG TTAGGCGGAGTCAGTAAGTCACCTAACGATGAATTCGAGCTCGTTTA). The two resulting constructs were mated to get yPP113. RTS1 has been integrated on the left side of *AT3003* using NK53 (ATTTACAGCCGCCAAACGTGGCTTATTCAAAGCCCCACTTGAGAACCAATG CGAGCGCATCTGGAAAAAGGGCTATCGGTGAATTCGAGCTCGTTTAAAC) and NK54 (GCCATGGCAGCTAGGTAACAACCACGAGGCCGATCGCTGCTTCGGCGGAT

TTAGGCTGACGTAAGATGAGACTGTTTGTTTTAGGTGACACTATAGAACG). RTS1 has been integrated on the right side of *AT3003* using NK55 (TTTGCGAATCCAAAGTCTGCCAATGGATATACGCTCTACTTTCGACCACTGA TTGTTTCCTGCATTTCTCAAATAATAGGTTAGGTGACACTATAGAACG) and NK56 (AAACCAACACCACTGCACATACGACCGATAAGAATTAAATACAGCCATTGT GCACGGTACGCTAGTGAATACAGTAAATGGAATTCGAGCTCGTTTAAAC)

Synchronization experiments

For two-dimensional gels, cells were grown to an OD_{600} 0.5. 10 mM HU was added to the culture and a quarter of cells harvested and frozen at 1.5 hours and 3 hours. The remaining culture was pelleted, HU washed, and the pellet resuspended in a HU free media. Half the culture was collected after 30 minutes post release from HU arrest and the rest collected 90 minutes post release.

For the microarray HU experiment, cells were grown to an OD_{600} 0.5 and collected for G2 phase sample. 10 mM HU was added to the culture and the cells were collected after 3 hours for the S phase sample.

For the microarray time course experiments, cells were synchronized by centrifugal elutriation and then kept at the restrictive temperature of 35°C to synchronize cells in G2 using cdc25-22 mutation for 3.5 hours. The cells are then shifted to the permissive temperature of 25°C. Samples were collected at the indicated times.

Two-dimensional gel electrophoresis

Genomic DNA was isolated using cesium chloride gradients and two-dimensional gel electrophoresis performed as described (Noguchi *et al.* 2003). To study replication at *AT2062*, genomic DNA was digested with BamHI and XhoI. For analysis at the RTS1 integration site, DNA was digested with KpnI and SacI. Southern blotting was performed as described (Noguchi *et al.* 2003).

Micro-array design

Probes were designed to cover 128 kb region with *AT2062* in the middle using the Arraydesigner 4.2. The average distance between the probes was 250 bp and 448 probes were designed with an average length of 60 bp. The slides for microarrays were printed in the Rando lab and the slides were post-processed using the protocols followed in their lab.

DNA preparation and microarray experiment

Genomic DNA was isolated using cesium chloride gradients as described (Noguchi *et al.* 2003). DNA was indirectly labeled with Cy3 and Cy5 dyes using the Amino-allyl labeling protocol from the DeRisi lab with a few modifications used in our lab (http://derisilab.ucsf.edu/data/microarray/protocols.html) (Dutta *et al.* 2008). Experimental DNA was mixed with the reference DNA, which was the G2 samples for all our experiments, for differential hybridization. The sample was hybridized onto the microarray slides for 16 hours at 65°C. Slides were scanned

using Genepix5000b scanner and the data was acquired using Genepix pro 6.0 software. The data was normalized and replication profiles created using excel.

Deep sequencing experiments

Cells were synchronized using centrifugal elutriation. One half of the culture was collected immediately after elutriation as the G2 sample for sequencing. To the other half 10 mM HU was added and cells kept at 25 °C for four hours and the cells collected as the S phase sample for sequencing. 1 OD was also collected every 20 minutes, pelleted and resuspended in 70% ethanol, and processed for flow cytometry. Genomic DNA was isolated using cesium chloride gradients as described (Noguchi *et al.* 2003). DNA samples were sent to Helicos Biosciences and the data collected analyzed by the Weng lab using Igor software.

Flow cytometry:

Cells were collected for flow cytometry and processed as described previously (Forsburg and Rhind 2006).

RESULTS AND DISCUSSION

Origin efficiency at *AT2062*

To measure the efficiency of a late firing origin we chose *AT2062*, a late replicating sequence present on chromosome II in fission yeast. *AT2062* was chosen because known adjacent neighboring origins are present very far away. Recent study from Nurse lab shows that *AT2062* fires with about 10% efficiency (Wu and Nurse 2009). We integrated RTS1 sites on either side 40Kb apart. The schematic representation is shown in Figure III.1.

To measure origin efficiency of *AT2062*, two-dimensional gel electrophoresis was employed. 2-D gels are a common method of distinguishing if a stretch of DNA being studied is getting passively replicated or is actively firing (Brewer and Fangman 1987) (Figure III.2). Signal from the 2-D gels may be quantified to determine the firing efficiency of an origin. Using 2-D gels we wanted to compare the origin firing efficiency of *AT2062* in a wild type strain and a strain containing *AT2062* flanked by RTS1 on either side. DNA was isolated and from culture arrested and released from HU.

There are three different possibilities that can come as a consequence of using RTS1 to block passive replication. First, we may see an increase in the origin firing efficiency at *AT2062*. Second, the firing efficiency of the origin does not increase. This would indicate that either the increase in origin efficiency model is incorrect or there is passive replication occurring at that region.

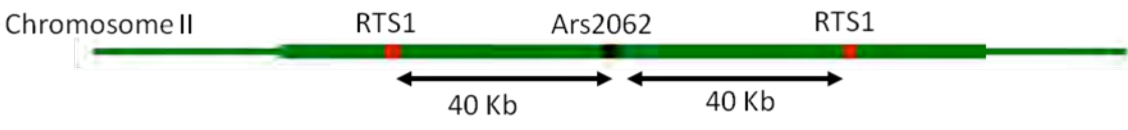
Figure III.1 Schematic representation of RTS1 integration

The regions where RTS1 was integrated are shown: *AT2062*, the late replicating sequence present on chromosome II and *AT3003*, an origin present in the ura locus. The RTS1 sites are 40 kb on either side of *AT2062* and 7 kb on either side of *AT3003*.

Figure III.2 A general description of two-Dimensional gel electrophoresis

DNA is digested with specific restriction enzymes and run in two dimensions with different conditions. In the first dimension, DNA is separated by size and in the second by shape. The common replication intermediates seen are described in the right-handed panel. If the region of interest has an active origin then we see a bubble arc whereas Y-arc is seen when replication forks from neighboring origins passively replicate the region. X-shaped intermediates are the recombination intermediates. Linear DNA is the majority of DNA containing no shape.

Figure III.1 Schematic representation of RTS1 integration



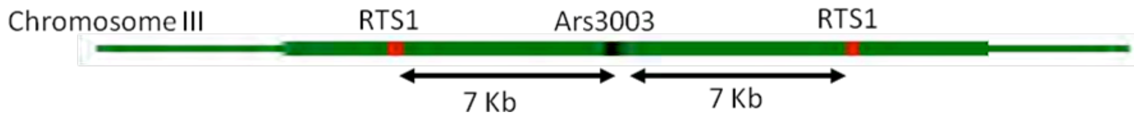
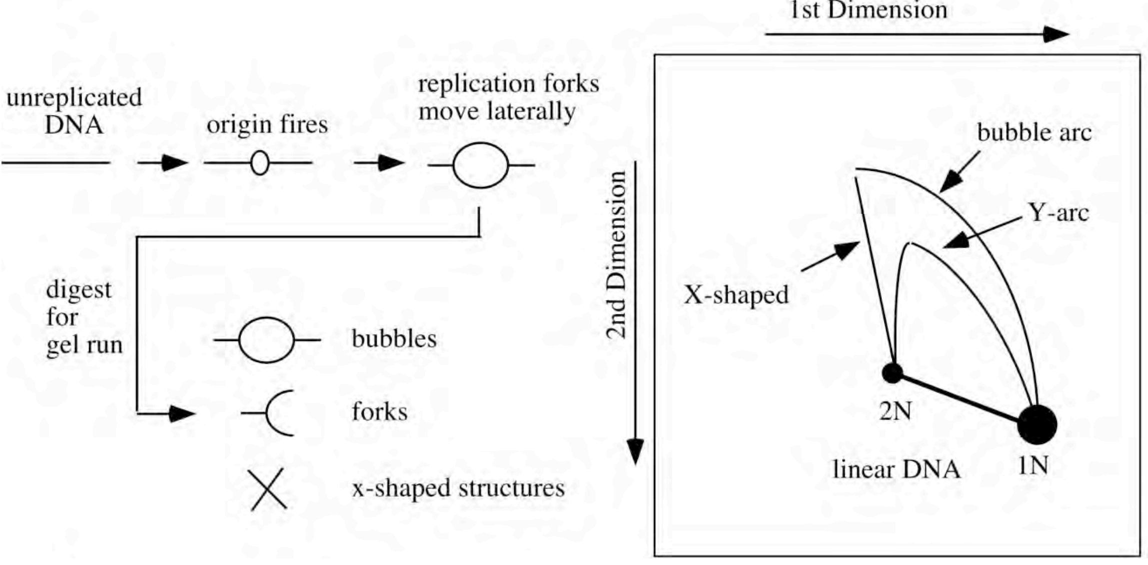


Figure III.2 A general description of 2-Dimensional gel electrophoresis



The final possibility is that cells will be unable to replicate the region between RTS1 due to the failure of *AT2062* to fire and hence it may lead to genomic instability and lethality. The result will show that our model is incorrect and the region can replicate only by passive replication. The increase in firing efficiency can be measured on the 2-D gels by measuring the percentage of bubble arcs, the shape that appears on the blots when the origin fires and comparing it between the wild type and the RTS1 strain.

HU arrest experiments were performed for 2-D gels. We observed no difference in the replication pattern between the wild type and the RTS1 strain. Passive replication seems to be occurring in both the strains at *AT2062* as indicated by the Y-arc (Figure III.3). We were unable to observe any bubble arcs in either strain. This passive replication indicates that *AT2062* is not firing during the S phase or it is firing at a low rate, which cannot be detected due to limitations of 2-D gels. To ensure that RTS1 sequences were blocking the forks, HU arrest and release was performed and 2-D gels were run to check for fork blockage at the termination site upstream of *AT2062* in Figure III.4. The spot on the 2-D gel shown on the Y-arc indicates fork arrest. These results demonstrate that forks appear to be blocked at the RTS1 sites.

The spots seen on 2-D gels at the RTS1 sites seem to be very weak which means that the block might not be very strong and would allow for forks to bypass it. Studies have shown that replication forks can bypass RTS1 blockage

in a recombination dependent manner. Passive replication observed at AT2062 could be due to forks bypassing the RTS1 sites and replicating AT2062 before it fires. *rad51*Δ prevents recombination-mediated fork bypass of RTS1 allowing us to observe if passive replication occurs at *AT2062*. Origin efficiency studies at any origin using RTS1 sites need to be done in a *rad51*Δ background to ensure that the neighboring forks are getting blocked.

Passive replication observed at *AT2062* may also be due to inefficient origins firing which have not been identified in the various genomic studies. Inserting RTS1 sites on either side of *AT2062* may force a number of these inefficient origins to fire, replicating the 80 kb region efficiently. We used two approaches to identify new origins in this region. First, we used microarray analysis and secondly deep sequencing.

Oligonucleotide arrays at *AT2062*

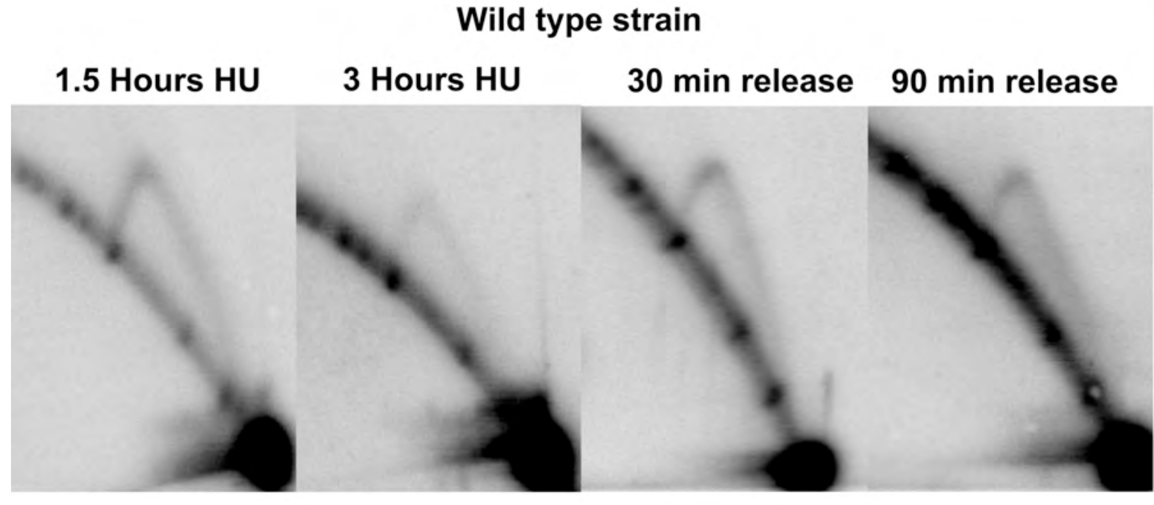
Micro-array analysis has been used in previous studies to look at origin firing (Raghuraman *et al.* 2001; Yabuki *et al.* 2002; MacAlpine *et al.* 2004; Woodfine *et al.* 2004). We used micro-arrays to measure changes in the copy number of the *AT2062* region during S phase. We designed oligos ~250bp apart spanning that 128 kb region on chromosome II. This 128 kb region contains the 80 kb region flanked by RTS1. Replication profiles were generated for the array experiments by fluorescently labeling S phase DNA and hybridizing it to the oligonucleotide arrays. The DNA copy number was measured by normalizing the

Figure III.3 *AT2062* replicates passively with or without RTS1

A) Passive replication is observed in wild type cells (yFS105) when cells are arrested in the beginning of S phase using 10 mM HU. The arrest was for either 1.5 hours or 3 hours and the cells released into a HU free media. The release was for 30 or 90 minutes. To study replication at *AT2062*, genomic DNA was digested with BamHI and XhoI. Southern blotting was performed as described (Noguchi *et al.* 2003).

B) Passive replication was also observed in RTS1 flanking *AT2062* (yPP113). There seems to be no difference in the Y-arc, which represents passive replication, between wild type and the RTS1 flanked strain indicating a failure of *AT2062* to fire.

Figure III.3 *AT2062* replicates passively with or without RTS1



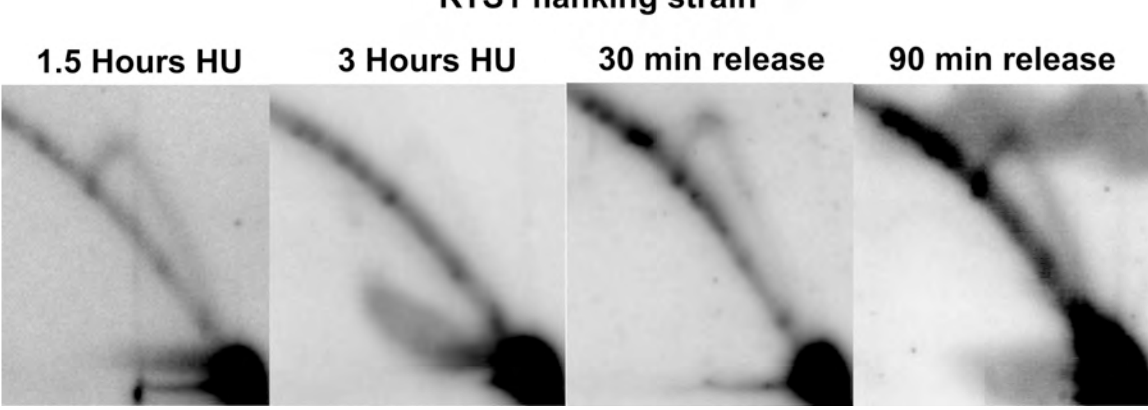
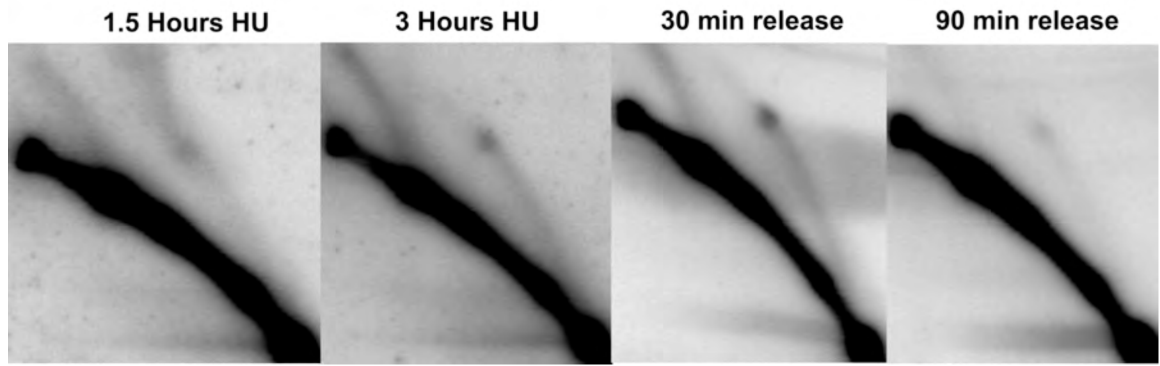


Figure III.4 Forks are getting blocked at RTS1 sites

Passive replication is observed in RTS1 cells (yPP113) when cells are arrested in the beginning of S phase using 10 mM HU. The arrest was for either 1.5 hours or 3 hours and the cells released into a HU free media. The release was for 30 or 90 minutes. DNA was prepared using cscl gradient method. For analysis at the RTS1 integration site, DNA was digested with KpnI and SacI. Southern blotting was performed as described (Noguchi *et al.* 2003). Fork blockage appears as a blob or big spot in the place where Y-arc is expected. The block does not seem to be very strong but appears in all the four conditions tested.

Figure III.4 Forks are getting blocked at RTS1 sites



data to G2 DNA. The array was designed to give us a resolution of 1 kb. Figure III.5 shows the replication profile generated during the arrest, effectively the beginning of S phase. Replication does not occur at *AT2062* consistent with previous studies showing that *AT2062* gets replicated later in S phase (Heichinger *et al.* 2006; Wu and Nurse 2009).

To study the increase in origin efficiency at *AT2062*, we looked at replication profiles as cells progressed through S phase in wild type cells by performing a microarray time course experiment. Samples collected every five minutes from 80-95 minutes show that the region has been mostly replicated by two or three potentially inefficient origins (Figure III.6). The timing of when cells enter S phase varies between elutriation and we were unable to capture the transition from a non-replicated locus to completely replicated locus in a strain lacking RTS1 sites. This limited our S phase progression experiments. We looked at origin efficiency in the RTS1 strain by collecting cells at two points- 95 and 105 minutes after release and saw similar replication profiles (Figure III.7) indicating that we are unable to study replication kinetics using microarrays. A higher resolution or more sensitive technique is required to understand the origin pattern in the 80 kb region.

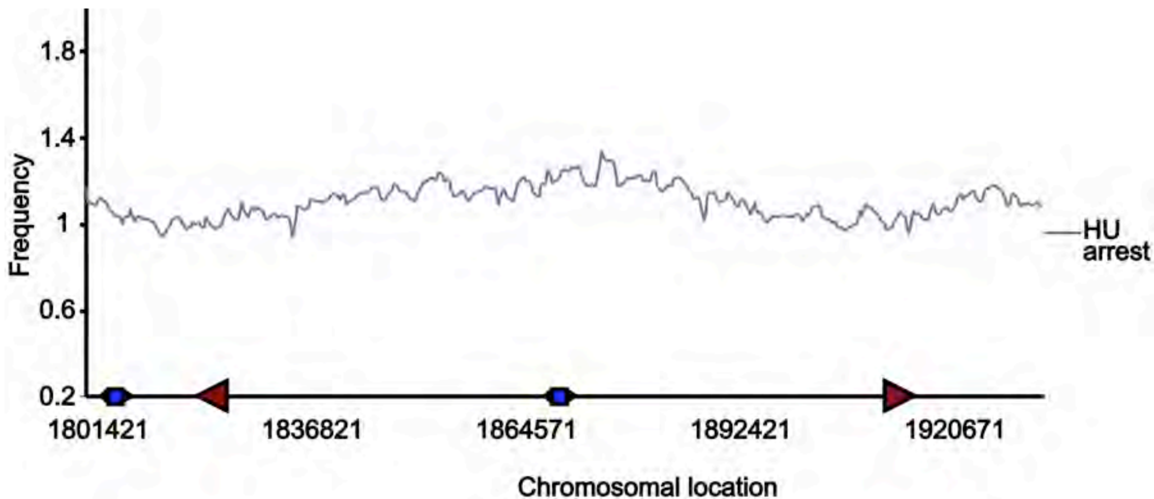
Using deep sequencing to look at AT2062 region

Data from micro-array analysis suggests that inefficient origins maybe present in the 80 kb region including *AT2062*. To confirm the presence of

Figure III.5 *AT2062* does not fire in early S phase

Cells were arrested in early S phase using HU in the strain flanking AT2062, yPP113. DNA was prepared and hybridized against G2 DNA. Replication profile was made as described in materials and methods. Known origins are represented with and there are only two known origins in the region on our microarray. *AT2062* is the origin in between the RTS1 represented by . As expected in HU, only early/efficient origins fire and we do not detect any signal from within the RTS1 region.

Figure III.5 *AT2062* does not fire in early S phase



inefficient origins, we used deep sequencing to identify origins in this region. We used Helicos single molecule sequencing technology. In this approach, genomic DNA is randomly fragmented to 100-200 bp and poly-A tail attached to the fragments with a fluorescently labeled A at the end of the tail. The poly-A tail attached fragments are hybridized to flowcell, a platform coated with immobilized poly-T oligomers. Sequencing consists of numerous cycles of replication. During each cycle, polymerase and a single labeled nucleotide which has a reversible fluorescence terminator attached to it is flowed in. The flow cell is imaged to locate the position and therefore identify fragments to which this nucleotide has attached. The fluorescent label is then cleaved and released and the cycle of polymerase and nucleotide is repeated for the remaining three nucleotides. This sequential cycle using reversible fluorescent labeling followed by imaging is repeated 200-250 times.

Similar to the micro-array analysis, we measured the copy number of S phase DNA and normalized it to G2 phase of the DNA. Sequencing was done on a strain without the RTS1 inserts. Cells were synchronized by centrifugal elutriation in G2 phase and released in a HU media for four hours. Figure III.8 shows the replication profile of the same 128 kb region. The replication profile at *AT2062* indicates that there is a presence of small inefficient origins in the 80 kb region flanked by RTS1. The inefficient origins can potentially fire and passively replicate the *AT2062* locus. We conclude that *AT2062* is not a particularly useful region for studying late origin efficiency.

Figure III.6 Replication profile during S phase at *AT2062* in wild type cells

cdc25-22 (yFS128) cells were synchronized in G2 by centrifugal elutriation and then synchronized a second time due to the *cdc25-22* arresting cells in G2 when kept at the restrictive temperature (35°C) for 3.5 hours. Cells were allowed to go through the cell cycle by switching back to the permissive temperature (25°C). Samples were collected for 80, 85, 90 and 95 minutes to look at the replication progression though late S phase as followed by flow cytometry. Known origins are represented with and there are only two known origins in the region on our microarray. *AT2062* is the origin in between the RTS1 represented by . The replication profile shows no progression through time. However, the array suggests that there may be atleast two more inefficient origins present between RTS1.

Figure III.6 Replication profile during S phase at *AT2062* in wild type cells

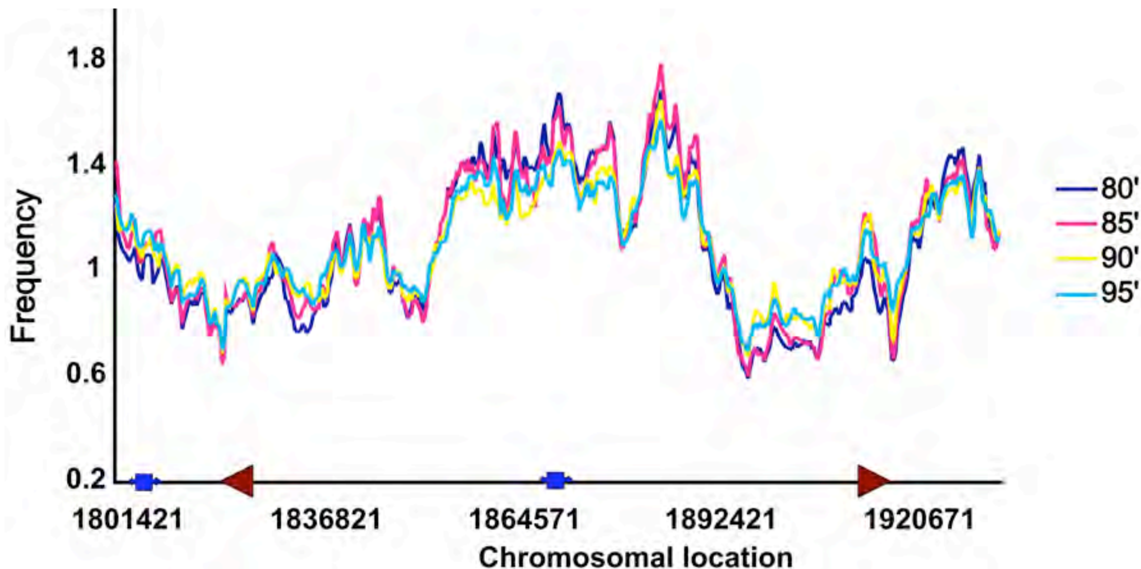
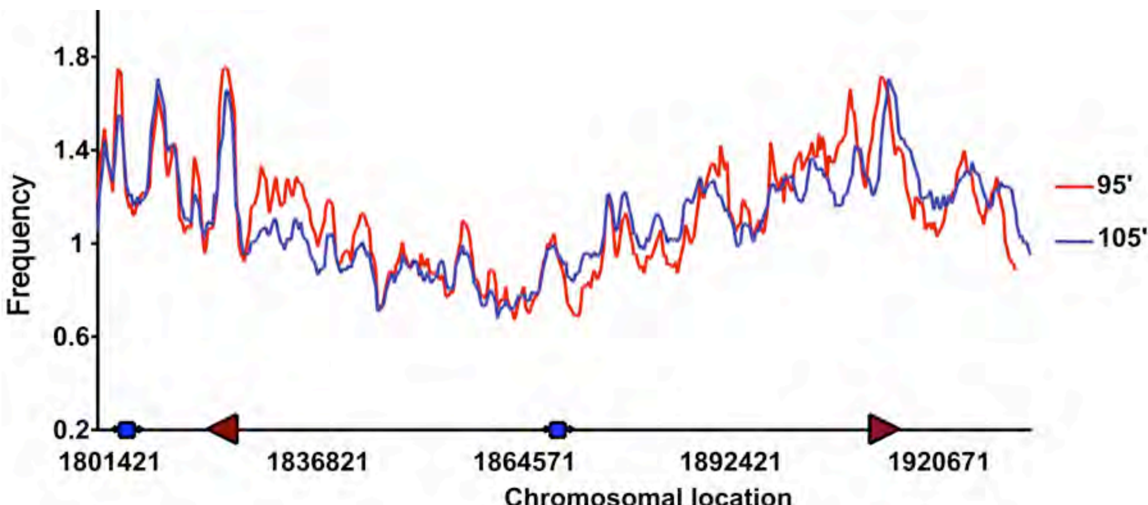


Figure III.7 Replication profile during S phase at *AT2062* in RTS1 strain

RTS1 flanking strain with a cdc25-22 background (yKN18) cells were synchronized in G2 by centrifugal elutriation and then synchronized a second time due to the *cdc25-22* arresting cells in G2 when kept at the restrictive temperature (35°C) for 3.5 hours. Cells were allowed to go through the cell cycle by switching back to the permissive temperature (25°C). Samples were collected for 95 and 105 minutes to look at the replication progression though late S phase as followed by flow cytometry. Known origins are represented with and there are only two known origins in the region on our microarray. *AT2062* is the origin in between the RTS1 represented by . The replication profile shows no progression through time. However, the array suggests that there may be atleast two more inefficient origins present between RTS1.

Figure III.7 Replication profile during S phase at *AT2062* in RTS1 strain



Origin efficiency at *AT3003*

AT3003 is one of the well-defined origins in the *ura4* locus on chromosome III in fission yeast (Kim and Huberman 1999). *AT3003* fires early during the S phase and the efficiency of the origin is about 30% (Patel *et al.* 2006). RTS1 sites were integrated on either side of *AT3003* in opposite directions thereby preventing the passive replication of the region. RTS1 sites are about 7.5 kb on either side of *AT3003*. Since there are no inefficient origins present in the region flanked by RTS1, *AT3003* must fire during every cell cycle for the region to replicate. Fork directional studies next to the origin within the RTS1 region will allow us to identify the direction in which the forks are traveling and determine if the efficiency of *AT3003* increases during S phase (Dalgaard and Klar 2001).

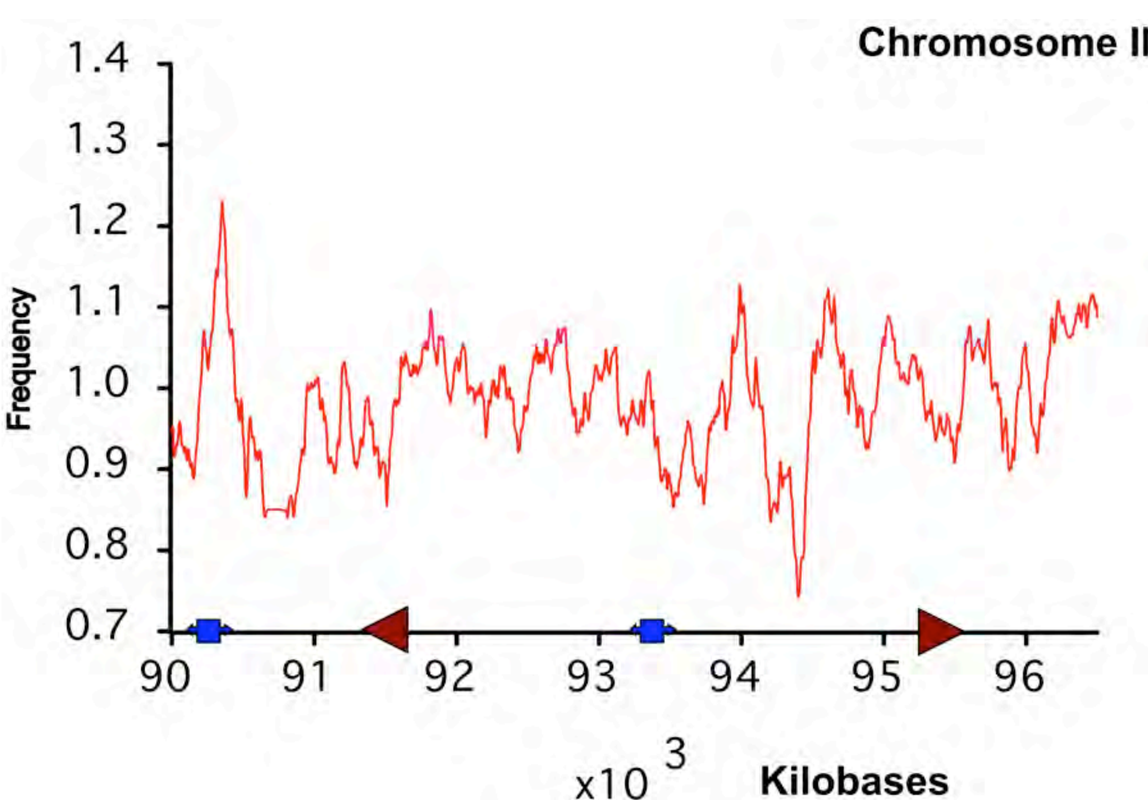
Fork bypass in a recombination dependent manner

Rad51 is the central mitotic recombination protein essential for homologous recombination in budding yeast. *rad51*Δ prevents recombination-mediated fork bypass of RTS1 allowing us to observe if passive replication occurs at *AT2062*. Studies show that a strain having *rad51*Δ background with a non origin stretch of DNA flanked by RTS1 on either side is lethal (Lambert *et al.* 2005). *AT3003* flanked by RTS1 in *rad51*Δ background will be lethal unless *AT3003* fires everytime. If forks bypass replication blocks in a recombination dependent manner, then deleting *rad51* will prevent the bypass. In the future, we will compare the efficiency of the origin in the presence or absence of *rad51*.

Figure III.8 Replication profile at *AT2062* in HU arrest using deep sequencing

S.pombe (yNW239) was synchronized in G2 by centrifugal elutriation. 10 mM HU was added and cells kept at 25°C for four hours. HU arrested sample was collected after four hours. G2 sample was collected after elutriation. Raw data from sequencing was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for the region on our microarray. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. Known origins are represented with and there are only two known origins in the region on our microarray. *AT2062* is the origin in between the RTS1 represented by . There seem to be about four more inefficient origins present within the RTS1 and these inefficient origins may be responsible for passively replicating *AT2062*

Figure III.8 Replication profile at *AT2062* in HU arrest using deep sequencing



CONCLUSION

Origin efficiency at *AT2062* was studied using 2-D gels, microarrays and deep sequencing. We were unable to force *AT2062* to fire using RTS1 sites. However, as shown above the passive replication observed at the origin was due to the presence of a few inefficient origins within the 80 kb region flanked by RTS1. This rendered *AT2062* unsuitable for testing our hypothesis of an increase in origin firing efficiency through S phase progression.

We have started to study the origin firing efficiency of *AT3003*. *AT3003* has the RTS1 sites close enough to make sure that there is no origins present in the region which can passively replicate *AT3003* region. We plan to study this origin in detail using the various methods used in the study at *AT2062*. All the studies at *AT3003* will be done in *rad51*Δ background to prevent the forks from bypassing the RTS1 block sites. Viability of this strain will show that the *AT3003* region flanked by RTS1 is only replicated by *AT3003*. We are also planning to study the origin efficiency using a *rad51* shutoff strain, which allows us to shut off *rad51* during the course of our experiment.

Chapter IV

Genome-wide analysis of origins in *Schizosaccharomyces* group

INTRODUCTION

Origin studies using a variety of model organisms have been ongoing for the last three decades. However these studies have been primarily limited to the study of a few well defined origins or a low resolution search for new origins. Only recently have genome-wide analysis experiments been feasible for global origin identification and characterization. These studies have been helpful not only in defining origin location but also in defining replication timing of origins. These origin studies also allow for the identification of common origin features. These features may then be used to identify putative origins in additional organisms in which these origin studies have not been performed.

Different origin identification methods

A number of methods have been employed to identify origins in fission yeast and other organisms. Hydroxyurea arrest has been used to identify origins that fire early in S phase for some of the genome-wide studies.These methods include:

1) Plasmid stability assays: Plasmid stability assays identified genomic regions capable of maintaining plasmid copy numbers termed as autonomous replication sequences (ARS) (Clyne and Kelly 1995). These regions capable of maintaining this activity, ranged in size from 100-150 bp for budding yeast, to 1 kb fission yeast, and to more than 10 kb for human cells.

2) Two-dimensional gel electrophoresis: Two-dimensional gel electrophoresis was also used early for identification of origins. Origins located on the smallest budding yeast chromosome III, were originally identified using 2-D gels (Reynolds *et al.* 1989).

3) Density transfer experiments: Density transfer approach using heavy isotope labeling of newly replicated DNA was the first genome-wide method used to identify origins in budding yeast. Heavy isotope labeling was followed by hybridization to microarrays to identify origins firing throughout S phase (Raghuraman *et al.* 2001).

4) Copy number change: Measuring copy number (replicated versus unreplicated DNA) using microarrays has been used in the recent past to identify the regions where origins are present (Yabuki *et al.* 2002; Heichinger *et al.* 2006).

5) ChIP-microarrays: Chromatin immunoprecipation followed by hybridization to microarrays can be used to map the binding sites of various pre-replicative complex (pre-RC) components. The binding sites are the origins where the pre-RC is formed (Wyrick *et al.* 2001).

6) BrdU pulse-microarray: Asynchronous or synchronized cells are pulsed with BrdU and flow cytometry is used to isolate BrdU labeled cells. BrdU labeling represents cells present in the S phase. DNA is then isolated from the samples, enriched by immunoprecipitation using BrdU specific

antibodies and hybridized to microarrays (Schubeler *et al.* 2002; Woodfine *et al.* 2004).

7) Single stranded DNA microarrays: Single stranded DNA produced upon HU arrest is hybridized on open reading frame (ORF) microarrays in the presence of HU have been used to map origins in fission yeast (Feng *et al.* 2006). The resolution for this study was about 12 kb.

8) Bioinformatic analysis: Bioinformatic analysis has also been used to propose putative origins based on AT rich islands in fission yeast (Segurado *et al.* 2003). The putative origins were validated using 2-D gels.

Studies in fission yeast and other organisms have identified origins at the genome-wide level (Raghuraman *et al.* 2001; Yabuki *et al.* 2002; MacAlpine *et al.* 2004; Jeon *et al.* 2005; Feng *et al.* 2006; Heichinger *et al.* 2006). However, the resolution of these studies is not very high. Budding yeast ARS consensus sequences (ACS) is very generic and cannot be used to actually map origins. On the other hand, no such motifs have been identified in any other organism and known origins have not given enough information to identify additional origins based on sequence homology. Therefore, precise identification of origin sequences on a genome-wide scale still needs to be done (MacAlpine and Bell 2005).

Origin sites are not conserved across species. Sequences important for origin activity are conserved across the *Saccharomyces* genus (Nieduszynski *et al.* 2006). Identifying origins across the *Schizosaccharomyces* genus will help

identify essential sequence or regions, which will in turn allow the identification of additional putative origins across the genome. Due to similar nature of origins in metazoans and fission yeast, identification of essential or signature sequences of origins may also allow identification of origins in metazoans with greater accuracy. In this chapter, we discuss our efforts at using a new technique of single molecule sequencing for identifying origins.

Sequencing to identify origins

Recent advances in sequencing technology have lead to improvements in the time taken to sequence DNA and also the cost of sequencing. There are various methods by which high throughput sequencing is done (Shendure and Ji 2008). One of the sequencing methods developed recently is the single molecule sequencing technique developed by Helicos Biosciences.

Single molecule sequencing does not use an amplification step like other sequencing methods, such as Solexa. DNA from the samples to be sequenced is randomly fragmented into 200 bp fragments. Fragmented DNA is then labeled at the 3’ end with a poly-A tail. This library of the fragmented poly-adenylated DNA is tethered to a surface coated with poly-T oligomers known as flow cell producing a disordered array of primed sequencing templates. The flow cell is imaged to identify the position of each tethered DNA strand. Sequencing consists of numerous cycles of strand replication, which allows for the sequence identification of the DNA strand. At each cycle, polymerase and a single labeled

nucleotide is added which has a reversible fluorescence terminator attached to it. Flow cell is again imaged to locate the position and therefore identify fragments to which the nucleotide has attached. The fluorescent label is then cleaved and released and the cycle of polymerase and a nucleotide is repeated for all four nucleotides. This sequential cycle using reversible fluorescent labeling followed by imaging is repeated 200-250 times. The average read length is about 25 bp since four cycles are needed for each and every base pair sequenced.

We show that deep sequencing may be used to identify origins in various organisms. Cells can be synchronized and samples collected during S phase. Samples are also collected from G2 phase and the DNA is sequenced. Regions that have replicated will have twice the amount of reads compared to regions that have not. Replication profiles can be created based on the number of reads.

We use deep sequencing to identify origins in three fission yeast- *S. pombe*, *S. octosporus* and *S. japonicus*. Although origins have been identified in *S. pombe*, no origins are yet to be identified in the other fission yeasts like *S. octosporus* and *S. japonicus*. The aim of this project is to identify the inefficient origins that have not shown up in the previous studies and to map the already known origins more precisely. This study shows that single molecule sequencing can be used to identify origins. In collaboration with the Weng lab, bioinformatic analysis is currently underway to identify signature sequences defining origins across the *Schizosaccharomyces* genus.

MATERIALS AND METHODS

Strain maintenance

All strains were grown in yeast extract with supplements (YES) at 25°C or 30°C and manipulated using standard methods (Forsburg and Rhind 2006).

Table IV.1 - Strain list

Strain	Genotype	Source
yFS101	*h-*	Lab stock
yNW239	*h- smt0 leu1-32 ura4-D18 his3-D1 cdc10-M17 sfr1::ura4 swi5::ura4*	From Nick Willis
yFS128	*h- leu1-32 ura4-D18 cdc25-22*	Lab stock
yFS275	Wild type *Schizosaccharomyces japonicus*	Lab stock
yFS286	Wild type *Schizosaccharomyces octosporus*	Lab stock

G2 synchronization

For the first dataset, the cells were grown at 25°C. *S.octosporus* and *S.japonicus* were grown at 30°C. Cells were grown to OD_{600} 1.4 for *S. pombe* and 0.8 for *S. octosporus* and *S. japonicus*. Cells were synchronized in G2 using centrifugal elutriation. A fraction of the cells were collected for the G2 sample and the rest incubated at 25°C for four hours in the presence of 10 mM hydroxyurea

(HU). 1 OD was also collected every 20 minutes, pelleted and resuspended in 70% ethanol, and processed for flow cytometry. Cells were also collected and DNA prepared for sequencing by cesium chloride gradient centrifugation. DNA samples were sent to Helicos Biosciences and the data collected analyzed by the Weng lab using Igor software.

Time-course experiment

Cells were synchronized in G2 using a *cdc25-22* temperature sensitive mutant. The culture was grown to an OD_{600} 0.5 and the culture shifted to the restrictive temperature of 35°C for 3.5 hours. A fraction of the cells were collected as G2 sample control for sequencing. Cells were then shifted to 25°C and samples collected every five minutes for flow cytometry. For sequencing samples were collected at time points 65, 75, 85, 95, 105, 115 and 125 minutes. DNA was prepared for sequencing using Qiagen G/20 columns as previously described (Wu and Gilbert 1995).

Flow cytometry:

Cells were collected for flow cytometry and processed as described previously (Forsburg and Rhind 2006).

Deep sequencing experiments

Cells were synchronized using centrifugal elutriation or using *cdc25-22* ts strain. One half of the culture was collected immediately after elutriation as the G2 sample for sequencing. To the other half 10 mM HU was added and kept at

permissive temperature for different times and the cells collected as the S phase sample for sequencing. 1 OD was also collected every 20 minutes, pelleted and resuspended in 70% ethanol, and processed for flow cytometry. Genomic DNA was isolated using cesium chloride gradients as described (Noguchi *et al.* 2003). DNA samples were sent to Helicos Biosciences and the data collected analyzed by the Weng lab using Igor software.

Alignments

To align the reads that we get, an alignment strategy is employed where the read (from sequencing) is aligned using the genomic sequence as a reference. Only uniquely mapped reads are used for the mapping study to eliminate repeat sequences in the genome. For each alignment all putative alignments to the reference genome are considered and alignments are considered unique if the best alignment has a normalized score greater than 4.2 (out of 5) and the next best alignment is at least .5 worse. Normalized Alignment scores are calculated as follows: sum (5*matches - 4*mismatches) / ReadLength. The reads are assembled into contigs or chromosomes at Helicos. The number of hits at each nucleotide in the genome is counted, histograms made and normalized for S and G2 samples. Normalized G2 hits for each nucleotide are subtracted from S phase hits and the frequency at each nucleotide is plotted giving us the replication profile for each chromosome.

RESULTS AND DISCUSSION

Various labs have done genome-wide analysis of origins in fission yeast (Segurado *et al.* 2003; Feng *et al.* 2006; Heichinger *et al.* 2006; Mickle *et al.* 2007). However, there have been no reports of genome-wide search for origins in either *S. japonicus* or *S. octosporus.* We have used deep sequencing to not only identify number of origins in the three *Schizosaccharomyces* species but also identify efficient and inefficient origins. The difference in efficiency can be measured by building replication profiles made by plotting the number of hits at each nucleotide across the genome against the nucleotide position. The height of the peaks at each origin gives us the efficiency of each origin.

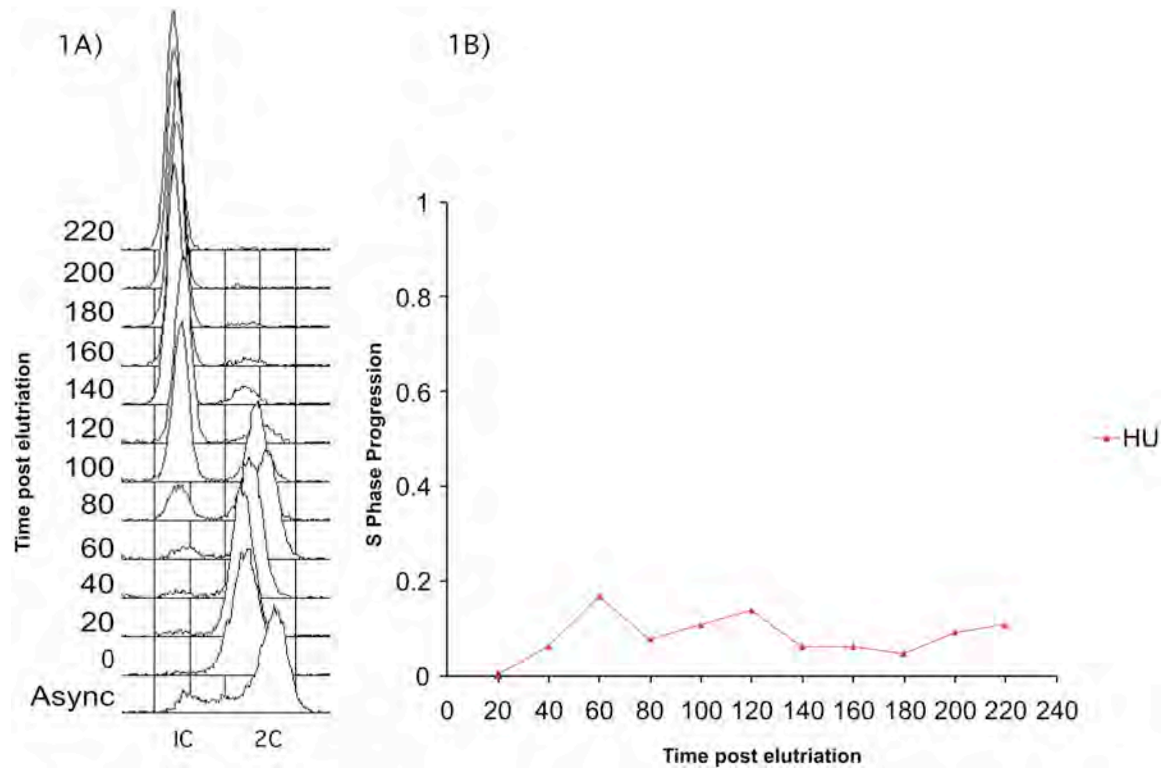
Identifying *Schizosaccharomyces pombe* origins

To identify origins in *S. pombe* we used HU to arrest cells at the beginning of S phase. For the first experiment, we synchronized an *S. pombe* strain in G2 phase of the cell cycle by centrifugal elutriation. The synchronized cells were then arrested in HU for four hours. HU arrest in early S phase was monitored by flow cytometry and S phase progression plotted {figure IV.1}. Flow cytometry shows that the forks have traveled about 15% in S phase. In the presence of HU the forks have not traveled far from the origins. Deep sequencing generated about 12 million reads for both G2 and S phase samples and the aligned reads were about 6 million. The number of reads for each point in the genome were measured and normalized to aligned G2 counts allowing us to generate high resolution replication profiles for early S phase. These profiles were smoothed

Figure IV.1 Replication arrest in the presence of HU for *S.pombe*

S.pombe (yNW239) was synchronized in G2 by centrifugal elutriation. 10 mM HU was added and cells kept at 25°C for four hours. HU arrested sample was collected after four hours. G2 sample was collected after elutriation. Cells were fixed every 20 minutes and nuclear DNA content measured by flow cytometry. A) S-phase flow cytometry histogram stacks shows that at the end of the time course cells are arrested in the beginning of S phase. B) S-phase progression is plotted over time by measuring the shifting of the mean of S-phase peaks from unreplicated 1C towards fully replicated 2C values. S-phase progression curve shows that cells have replicated about 15% showing that HU has arrested cells in the beginning of S phase.

Figure IV.1 Replication arrest in the presence of HU for *S.pombe*



and a peak finding algorithm used to identify the origins in the dataset. Figure IV.2 shows the process of identifying origins from the raw sequence reads. Figure IV.3 shows the data for all the three chromosomes of *S. pombe*. Using a peak finding model we identified origins. To verify the peaks we identified were origins, we compared this sequencing data with our microarray data for *ura4* gene cluster located on chromosome III. As seen in figure IV.4, the sequencing data correlates very well with our microarray data. Resolution of the origins using sequencing is greater than our array data and work from other labs (Heichinger *et al.* 2006). Origins identified in the previous studies were also identified in this study indicating that our technique is capable of identifying previously characterized origins as seen in figure IV.5. Rigorous analysis has been carried out by Weng lab to ensure that peaks identified in our studies are not random noise. There are peaks that have not been identified as origins in the previous studies due to the low resolution of those studies and these peaks are the ones that we are interested in exploring further to understand the complex nature of origin efficiency and location.

We were able to see peaks on chromosome III but on chromosome I and II the peaks seemed to be in regions near the centromere. Recent work shows that the region near the centromeric region seem to replicate early in a swi6 and dfp1 dependent manner (Hayashi *et al.* 2009). To see if the peak effect that we observed is due to the pericentromeric effect, we have done a similar experiment on a dfp1-3A mutant, which does not allow dfp1 to localize in the pericentromeric

Figure IV.2 Process of identifying origins on Chromosome I

yNW239 cells were synchronized using elutriation and cells collected for G2 phase and after 4hours in HU. Samples were sent for sequencing and the reads aligned to the three chromosomes using the known genomic sequence of *S.pombe* as the reference. A) The Raw data from sequencing the G2 and S samples of yNW239 are represented here for chromosome I. X-axis is the chromosomal location. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. B) The reads for the G2 and S phase samples are then normalized. C) The G2 reads are subtracted from the S phase reads. D) The resulting data is smoothed to give us the potential origins. E) Peak finding algorithm is used in the Igor software and identifies the peaks.

Figure IV.2 Process of identifying origins on Chromosome I

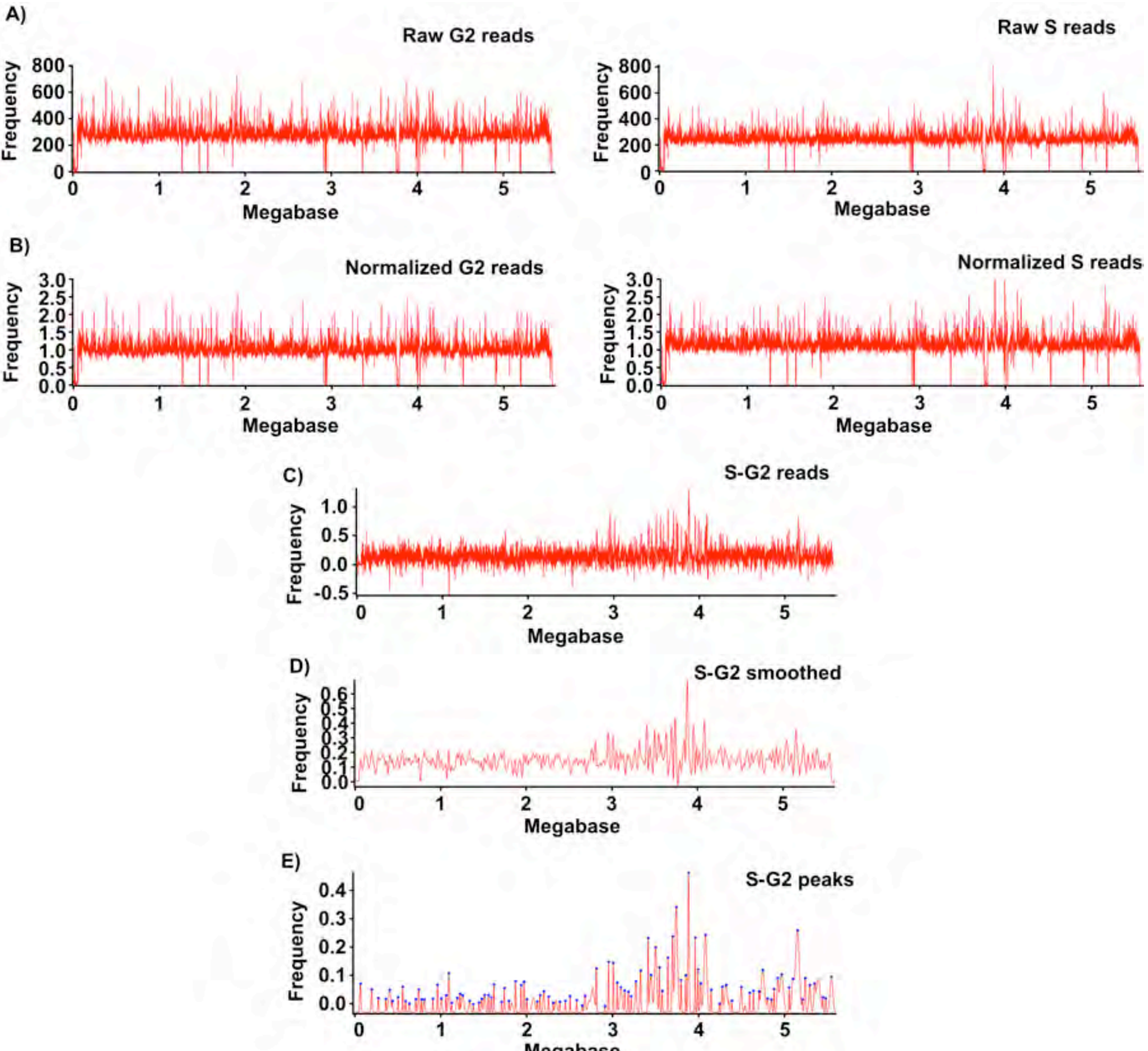


Figure IV.3 Replication profiles of *S. pombe* chromosomes

Raw data from sequencing of yNW239 was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for *S. pombe* chromosomes. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. Chromosome III shows lots of peaks representing origins. Chromosome I and II have only few peaks in regions centered around the centromere. Centromere is represented with .

Figure IV.3 Replication profiles of *S. pombe* chromosomes

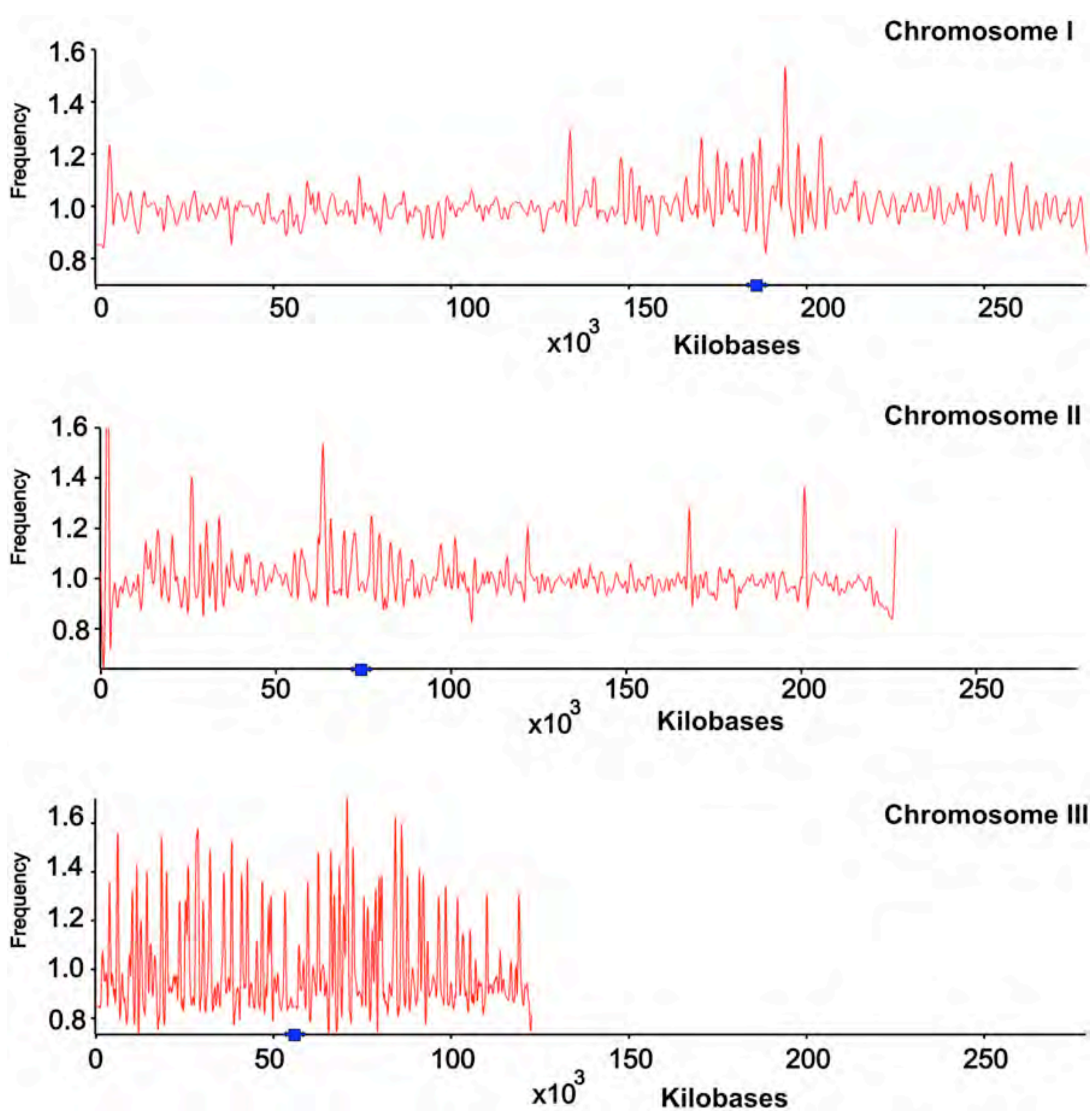
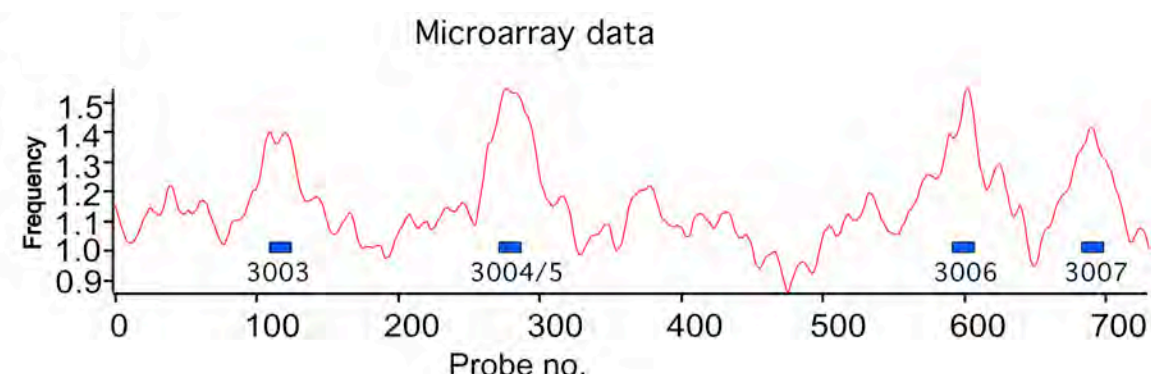


Figure IV.4 comparison of *ura4* locus between array and sequencing data

To analyze the difference in resolution between the microarray method and the deep sequencing method we compared the *ura4* region. A) Microarray data for the *ura4* region from HU arrest experiment in yFS240 B) Deep sequencing data for the same region from yNW239 cells arrested in HU.

The greater resolution of the sequencing data can be seen by the separation of *AT3004* and *AT3005* which in the microarray data appear as one origin. *AT3004* and *AT3005* are identified as separate peaks by the peak finding program establishing deep sequencing as a higher resolution method. There also seems to be the presence of some inefficient origins not yet identified in other studies.

Figure IV.4 comparison of *ura4* locus between array and sequencing data



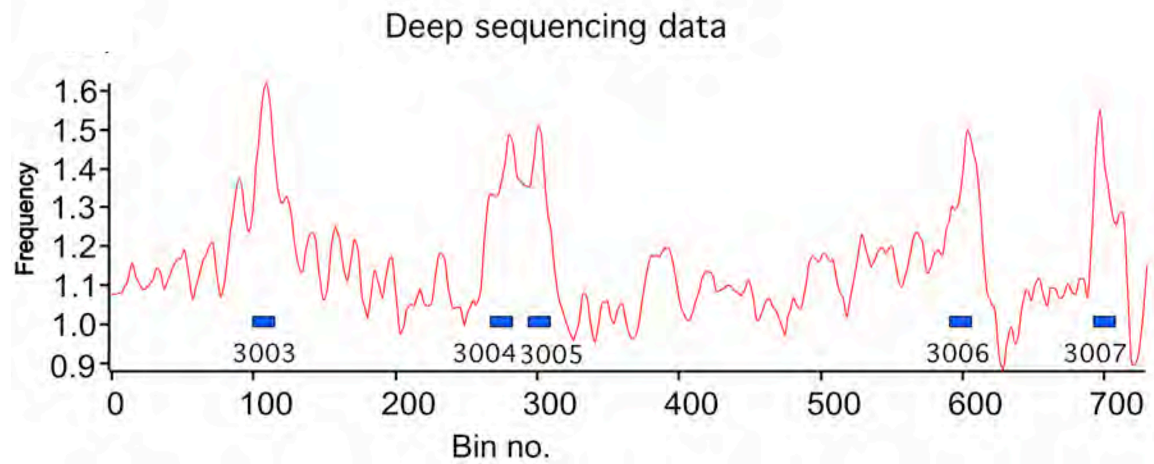
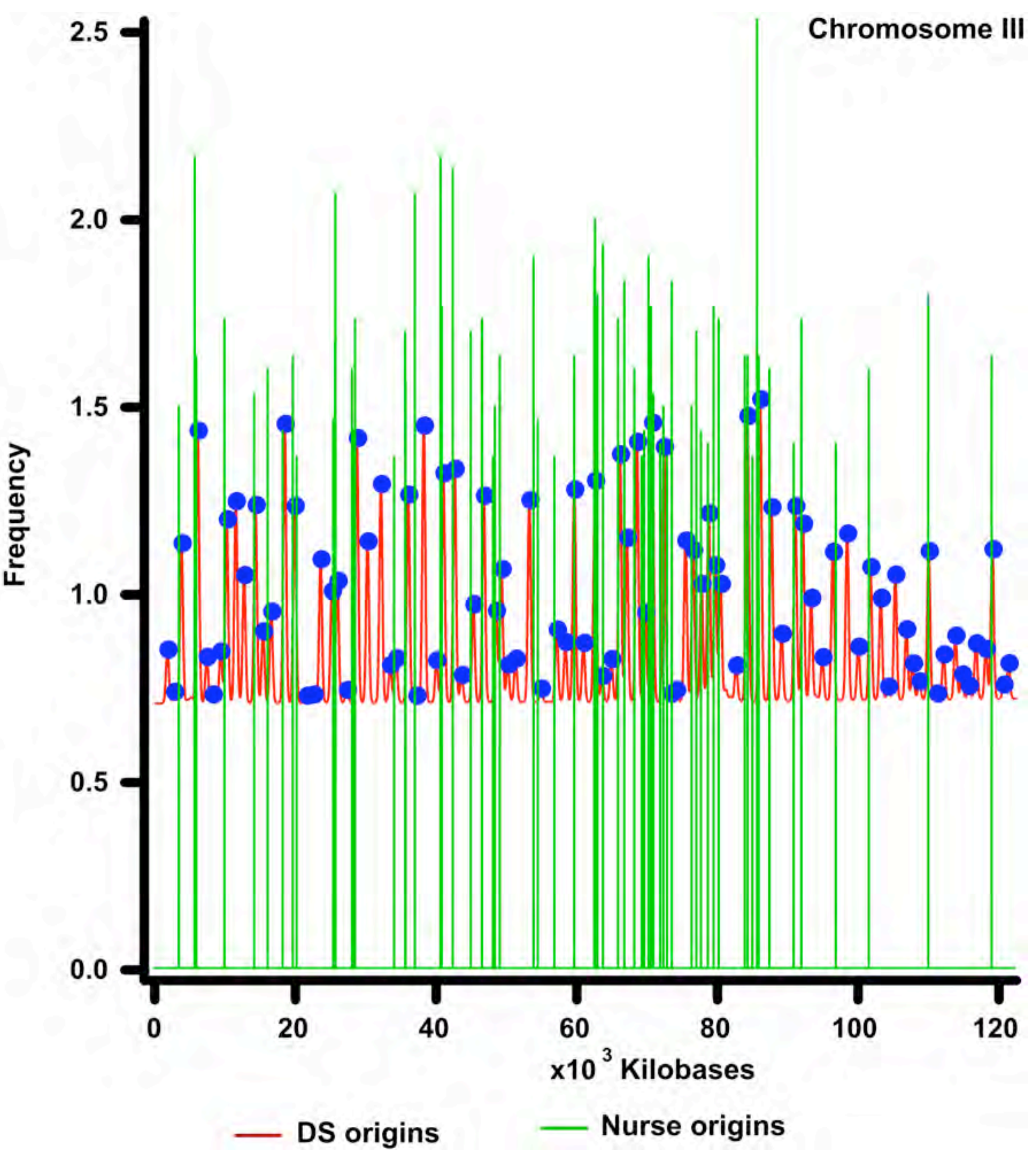


Figure IV.5 Comparison of origins identified with previous studies

To analyze the difference in resolution between previous studies that used different methods from the deep sequencing method we compared all the three chromosomes with origins identified from Nurse lab. Origins identified at a lower stringency on chromsome III are shown as a representation of the origins identified. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. The figure shows that we are able to not only identify majority of the origins identified in the Nurse paper but also a few more. Detailed bioinformatics analysis is still ongoing on these origins

Figure IV.5 Comparison of origins identified with previous studies



region. The DNA is presently being sequenced.

To overcome the centromeric effect observed in the first experiment we did another experiment where the forks had traveled 30% in another *S.pombe* strain, measured by flow cytometry, at the time of collection for sequencing (figure IV.6). Figure IV.7 shows the replication profiles for the three *S. pombe* chromosomes. Compared to the previous dataset, peaks were broader and interpreted as forks progressing further. Direct comparison between the two datasets for chromosome III is shown in figure IV.8. This is a result of the cells starting to leak through from the HU arrest. To look at the noise between two independent experiments we subtracted the G1 reads, obtained from the two experiments, from each other and observed a noise level of about 10%.

The microarray data from Nurse lab did not show a pericentromeric effect and to compare and measure the differences in resolution between deep sequencing and the microarray data, we synchronized the cells using the temperature sensitive *cdc25-22* for 3.5 hours and released to arrest cells in HU. Cells were collected after 90 minutes in HU, and this served as the S phase sample. The G2 sample was collected at the end of the synchronization with *cdc25-22*. These samples are presently being sequenced and the replication profile expected from this dataset should show us the sensitivity of our origin identification strategy since the experiment is similar to previous studies allowing us to directly compare the origins identified (Heichinger *et al.* 2006).

Figure IV.6 Replication arrest in the presence of HU in wild type *S. pombe*

yFS101 was synchronized in G2 by centrifugal elutriation. 10 mM HU was added and cells kept at 30°C for four hours. HU arrested sample was collected after four hours. G2 sample was collected after elutriation. Cells were fixed every 20 minutes and nuclear DNA content measured by flow cytometry. A) S-phase flow cytometry histogram stacks shows that at the end of the time course cells are arrested in the beginning of S phase. B) S-phase progression is plotted over time by measuring the shifting of the mean of S-phase peaks from unreplicated 1C towards fully replicated 2C values. S-phase progression curve shows that cells have replicated about 30% showing that HU has arrested cells in the beginning of S phase. The % replicated was twice when compared to the previously sequenced strain. The difference in % replicated could be due to the difference in the strain genotype where yFS101 is completely wild type but yNW239 has *sfr1* and *swi5* deletion, which might effect the ability of the forks to travel and replicate DNA.

Figure IV.6 Replication arrest in the presence of HU in wild type *S. pombe*

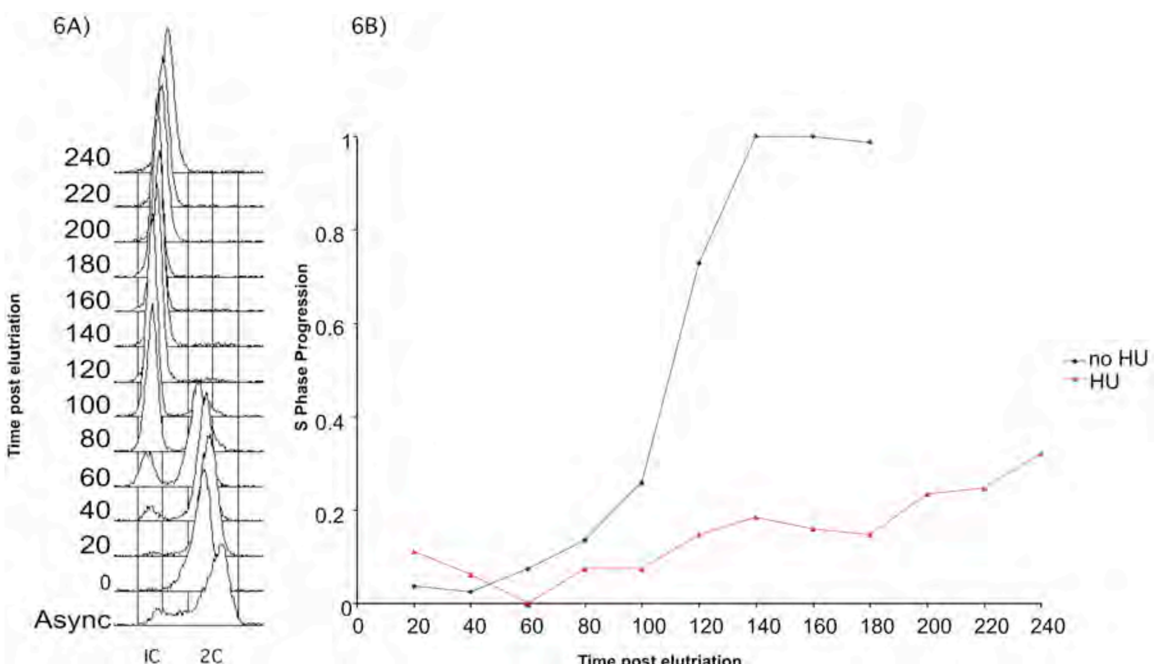


Figure IV.7 Forks progress slowly during HU arrest in *S. pombe*

Raw data from sequencing of yFS101 was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for *S. pombe* chromosomes. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. The peaks are broader in all the three chromosomes owing to a 30% replication seen by S-phase progression. Chromosome III shows lots of broad peaks, which can be one origin having traveled far or two origins firing close by and merging. Chromosome I and II have only few peaks in regions centered around the centromere. Centromere is represented with .

Figure IV.7 Forks progress slowly during HU arrest in *S. pombe*

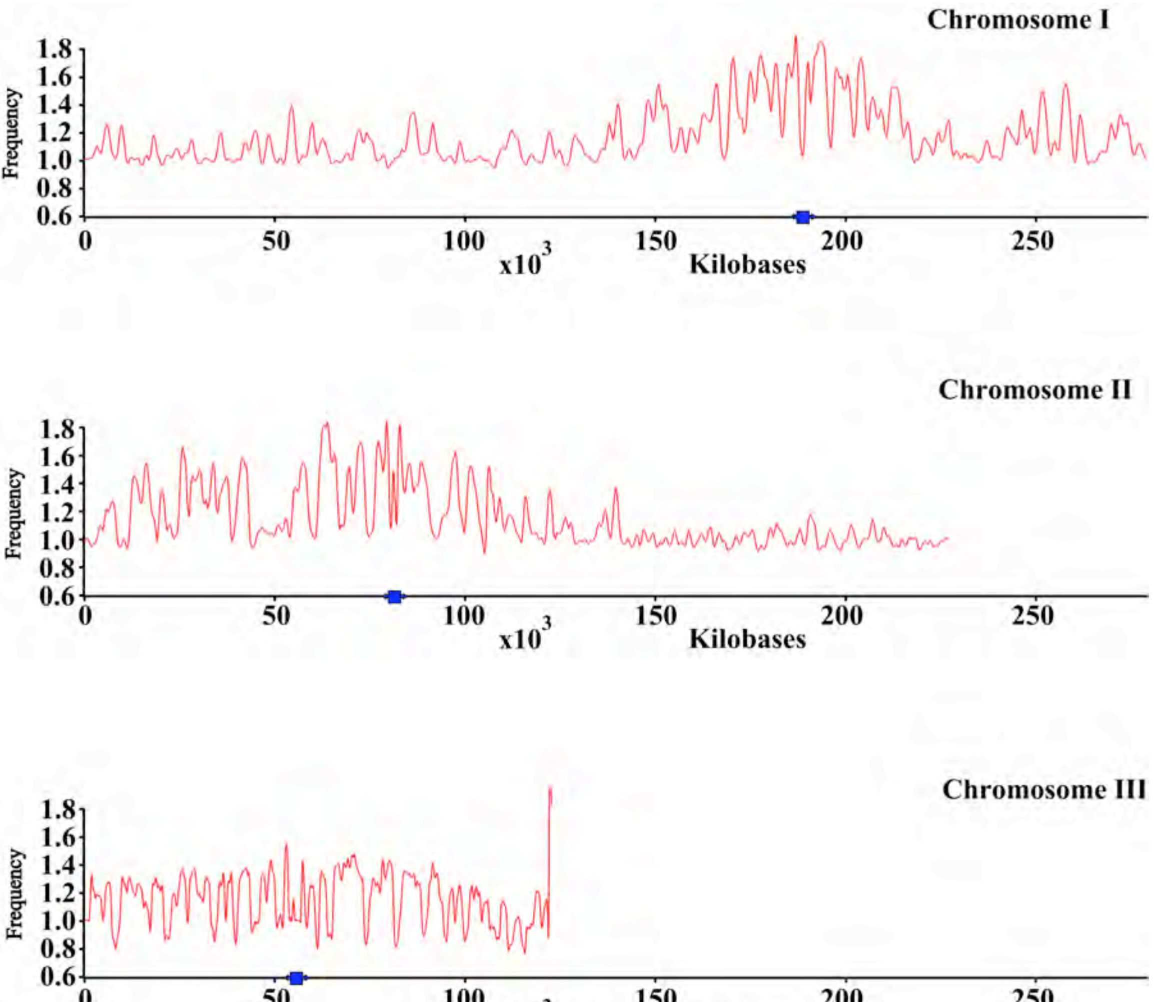
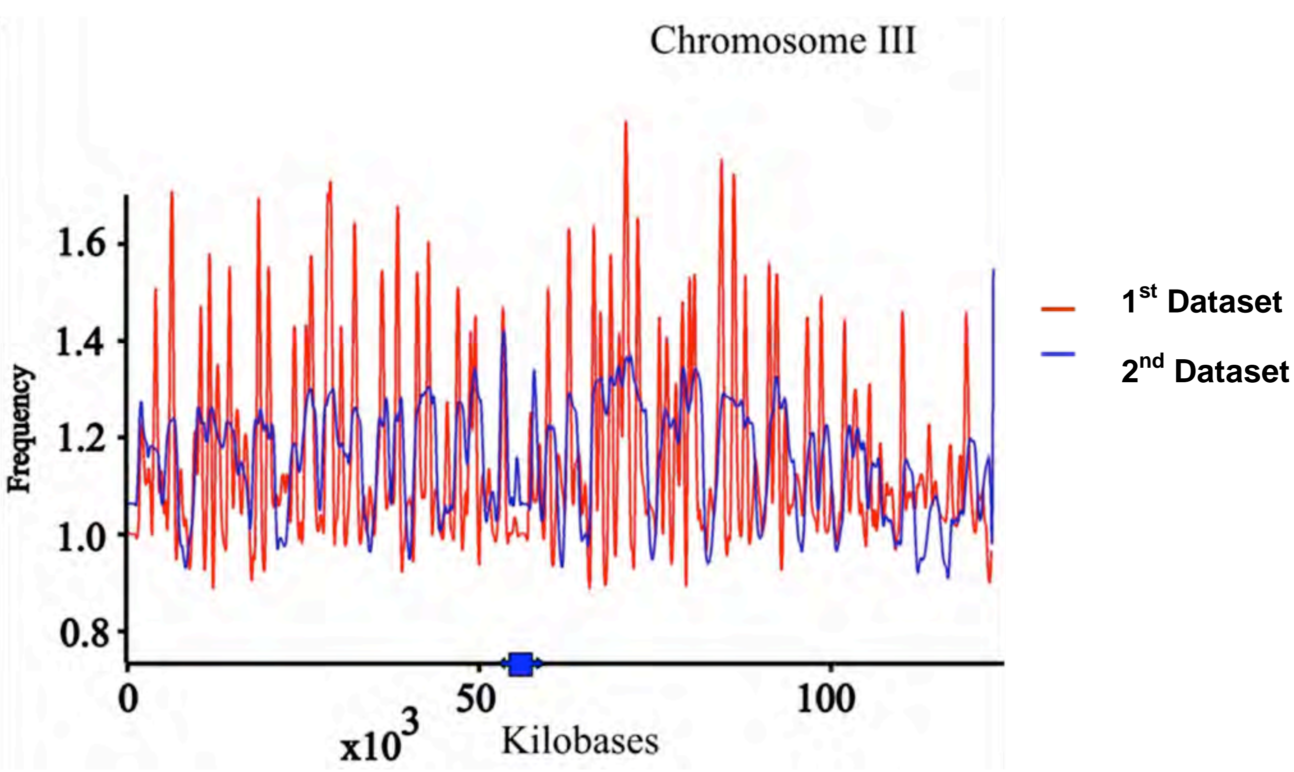


Figure IV.8 Comparison between two independent *S. pombe* datasets

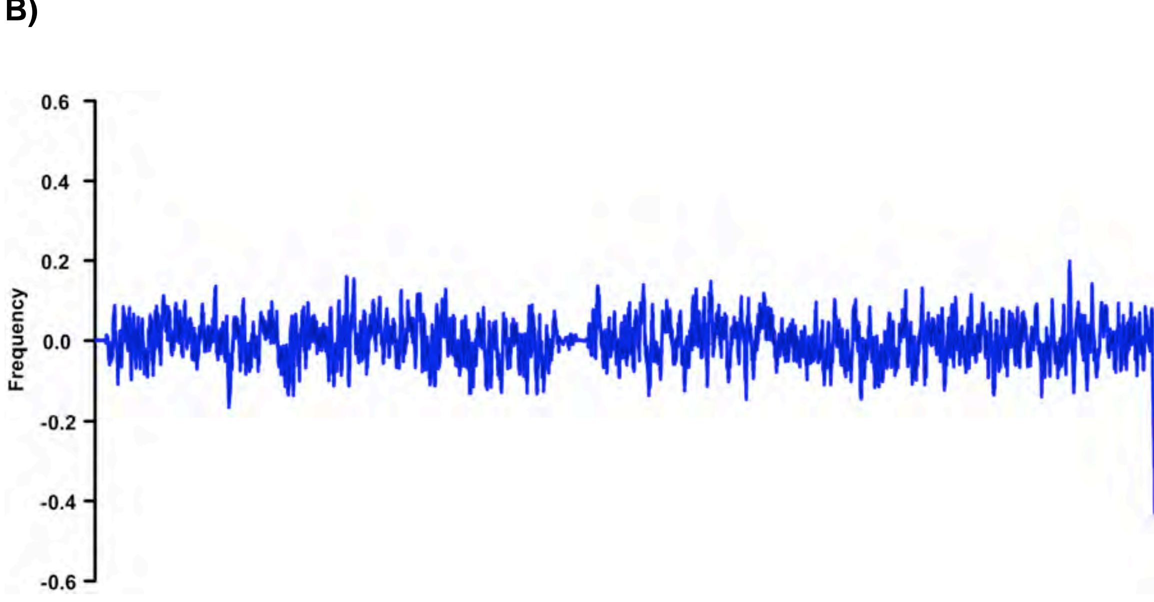
A) Raw data from sequencing was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for *S. pombe* chromosomes. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. Replication profile from Chromosome III for both the S. pombe experiments was overlayed. The peaks from the second dataset are broader and flatter due to several forks merging together.

B) The G2 sequence reads for the two wild type replicates were subtracted from each other, normalized and then smoothed. The G2-G2 control was done for chromosome III. The data shows that there is very little noise.

Figure IV.8 Comparison between two independent *S. pombe* datasets



B)



The experiments performed till now were done by synchronizing cells in early S phase using HU. This prevents us from identifying origins that will fire during late S-phase. Also, HU activates the replication checkpoint. The activation of checkpoints can itself have an effect on the genomic replication profile. It is imperative to perform experiments in the absence of HU to compare the results and identify if HU is having an effect on the replication profile. We have performed an S-phase time course experiment to understand the kinetics of replication timing in *S. pombe*. The cells progressed through S phase without HU which allows us to look at the replication profiles without the activation of any checkpoint. We can compare the origins identified in the time-course with those identified in our previous experiments. *cdc25-22* mutant strain was synchronized by incubating cells at restrictive temperature (35°C) for 3.5 hours. Cells were then released at the permissive temperature (25°C) and samples collected every 10 minutes to measure the progression of replication throughout the S phase. These timepoints will be assembled into a kinetic profile which will allow us to observe the replication kinetics of all the origins in the genome over time. This dataset will allow us to compare the origin efficiency between sequencing and microarray analysis.

Identifying *Schizosaccharomyces octosporus* origins

S. octosporus is a fission yeast similar to *S. pombe* except for having eight-spored ascii. *S.octosporus* genome has recently been sequenced in collaboration with the Broad Institute. No information is available regarding

origins in this fission yeast species. Origin identification will contribute greatly to understanding how replication occurs in *S. octosporus* and whether there are any similarities between the species. Origin identification in these different species will also be helpful in identifying the signature motifs, which could be employed to identify origins *de novo.* To collect samples for sequencing, cells were synchronized in G2 by centrifugal elutriation and a fraction was collected as the G2 sample. The culture was arrested in HU for four hours to synchronize cells in early S phase {figure IV.9A}. Replication profiles were made for *S. octosporus* chromosomes in the same way as for *S. pombe*. Using sequencing we were able to identify origins. Figure IV.9B shows the profiles for the three chromosomes.

Identifying *Schizosaccharomyces japonicus* origins

Rounding out the three fission yeast investigated, *S. japonicus* is different from both *S. octosporus* and *S. pombe* displaying invasive hyphal growth form. Hyphal growth is a virulence trait of pathogenic fungi. Those interested in understanding fungal diseases can take advantage of *S. japonicus* as a model organism. Similar to *S. octosporus*, no origin information is available for *S. japonicus* and we wished to address this point by identifying origins using deep sequencing. The HU experiment used for *S. japonicus* was similar to the ones used for *S. pombe* and *S. octosporus*. However, we observed that HU treated cells did not arrest in early S phase. Samples were still sequenced and we observed some peaks, which could be potential origins {figure IV.10}. This approach needs to be repeated using enough HU to efficiently arrest these cells

Figure IV.9 Replication profiles of *S. octosporus* chromosomes

A) S-phase flow cytometry histogram stacks shows that at the end of the time course cells are arrested in the beginning of S phase. B) Raw data from sequencing of yFS286 was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for *S. pombe* chromosomes. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each chromosomal position. The three chromosomes have plenty of peaks, which are potential origins. There does not seem to be any centromeric effect on any of the chromosomes. The potential origins do not have AT rich islands. The peaks identified as origins in this study are currently being experimentally verified using 2-D gels.

Figure IV.9A HU arrest of *S. octosporus*

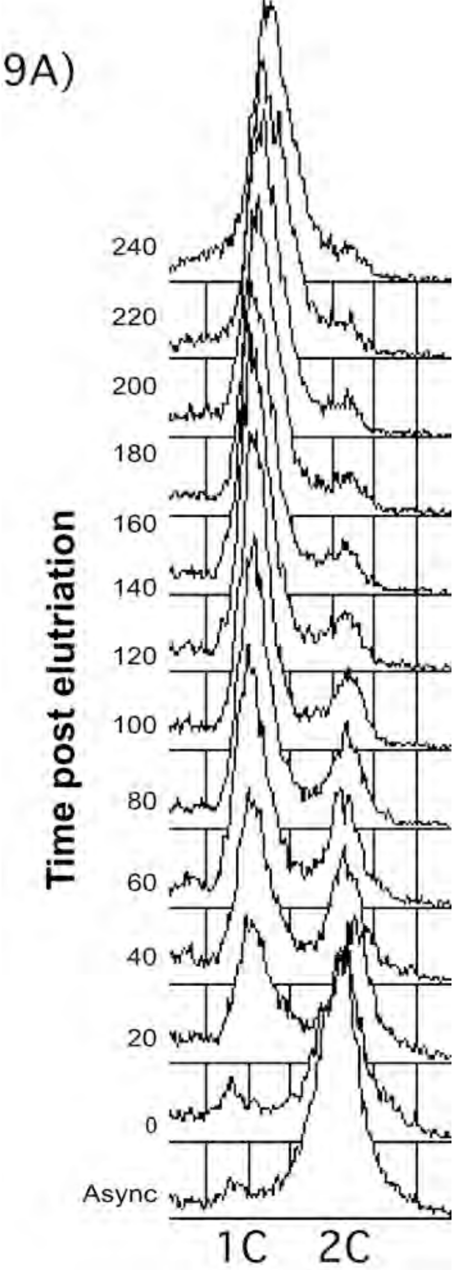


Figure IV.9B Replication profiles of *S. octosporus* chromosomes

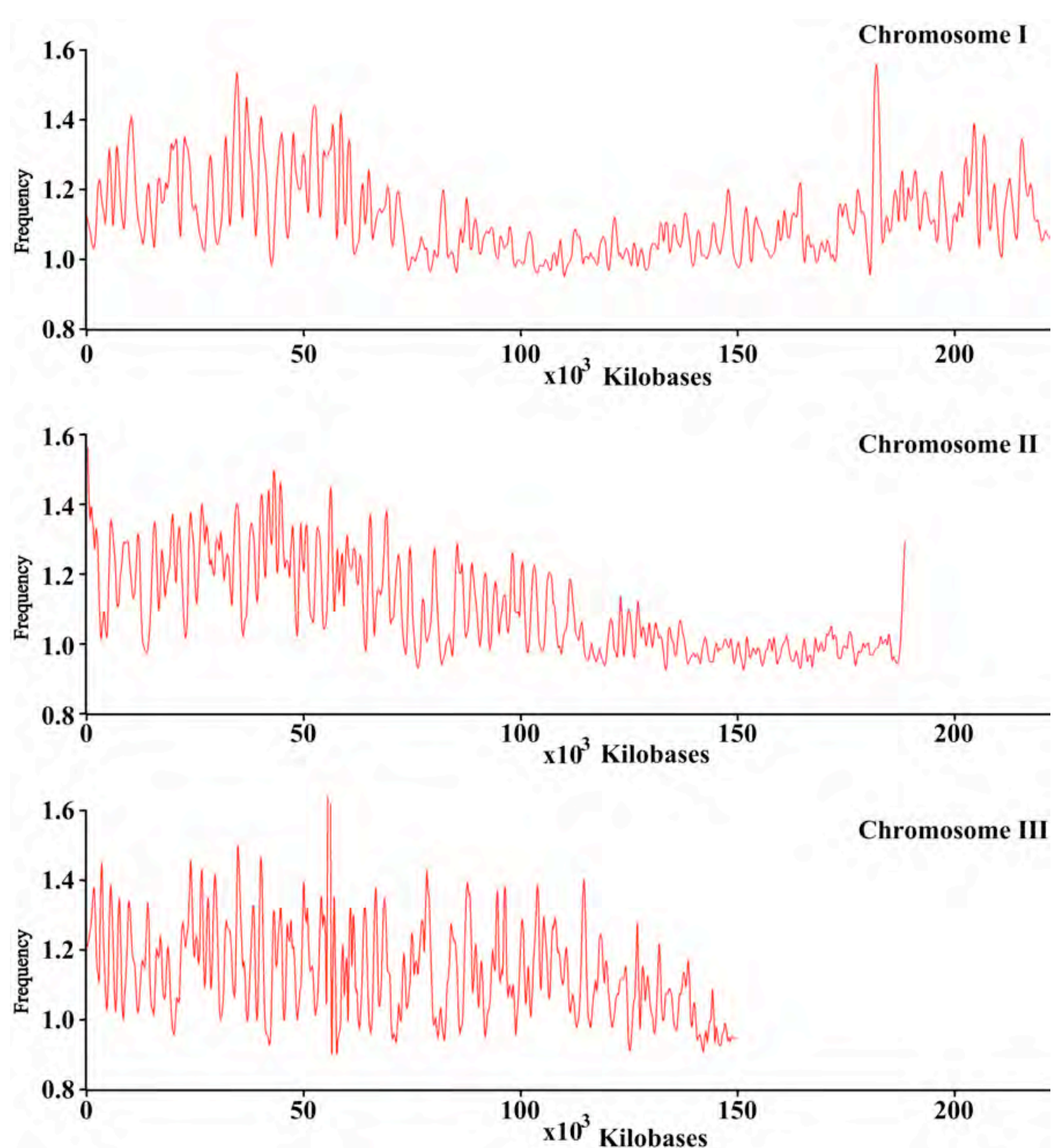
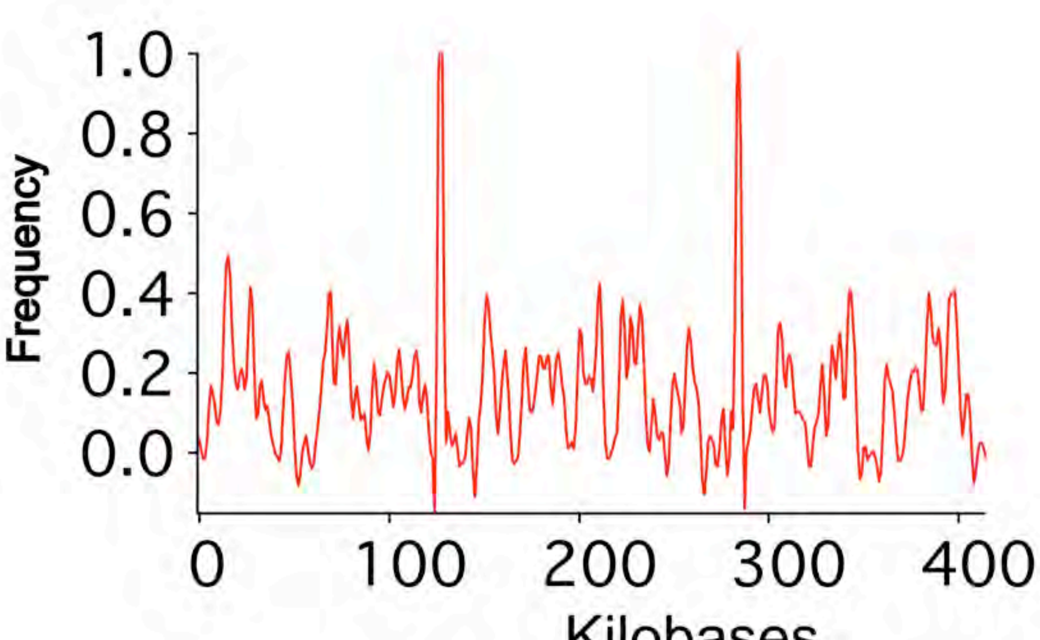
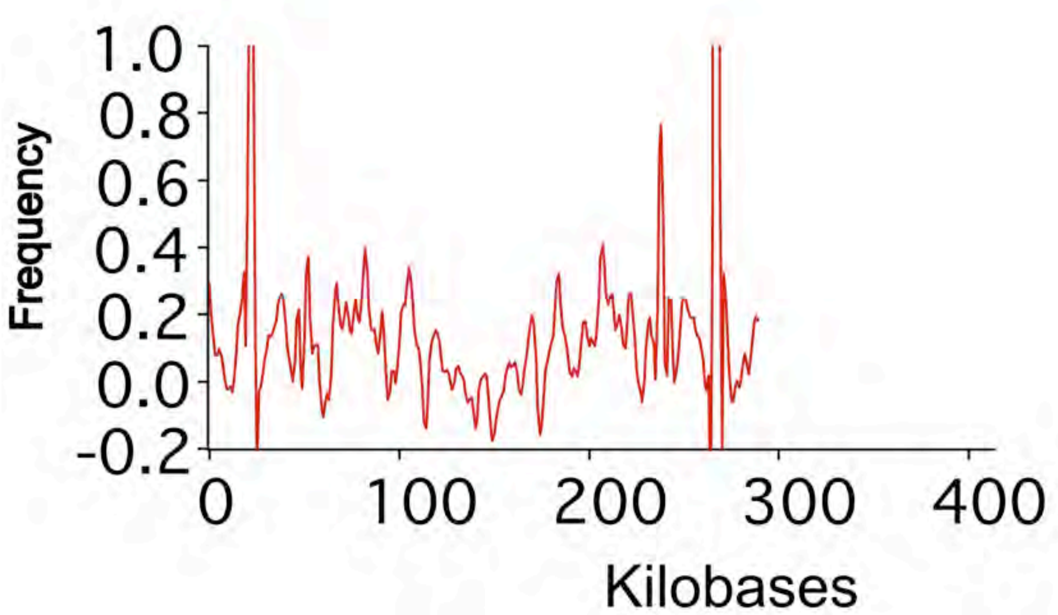


Figure IV.10 Replication profiles of *S. japonicus* scaffolds

Raw data from sequencing of yFS275 was taken for the G2 and S phase samples and normalized. The G2 peaks were subtracted from the S phase and the resulting data was smoothed to give the replication profile for *S. pombe* chromosomes. X-axis is the window index of 200 bp. Y-axis is the height of the peak and gives the number of reads at each scaffold position. The number of chromosomes in *S. japonicus* is not known and the data shown is from two of the contigs. Due to a lack of S phase arrest in HU, the number of S phase peaks after removing the G2 peaks is low. There are potential peaks identified by the peak finding software but the amplitude is low.

Figure IV.10 Replication profiles of *S. japonicus* scaffolds





S phase. It appears *S. japonicus* yeast requires a greater HU concentration (100 mM) for efficient arrest in S phase. We repeated the experiment with the HU arrest for only one hour to prevent cells from leaking through. We have also used an asynchronous culture using higher dose of HU and collected cells after 3 hours. Sequencing of the samples is currently going on.

CONCLUSIONS

Initial bioinformatic analysis shows that the AT rich islands prevalent in *S. pombe* origins do not seem to be determinants of origins in the other two fission yeast (Segurado *et al.* 2003). There seem to be no similarities between the origins identified on the different chromosomes of *S. pombe* and *S. octosporus* on which initial analysis has been performed. Further analysis is going on to identify the characteristics of origins to be able to identify them de novo across the genome.

We have shown that single molecule sequencing is an effective way of identifying origins across the genome and to make replication profiles, which can help in measuring the efficiency of an origin.

Table IV.2 – List of experiments performed for deep sequencing

Experiment	Strain	Origins	Species	Method
1) G2 Synchronized + 2)HU arrest 4 hours	yNW239	Sharp peaks identified as origins	*S.pombe*	Deep sequencing
1) G2 Synchronized + 2)HU arrest 4 hours	yFS101	Peaks are fewer, broad, flat	*S.pombe*	Deep sequencing
1) G2 Synchronized + 2)HU arrest 2 hours	yFS718	Sequence awaited	*S.pombe*	Deep sequencing
1) Cdc25-22 arrest 35C-3.5hrs + 2) 90' HU arrest	yFS128	Sequence awaited	*S.pombe*	Deep sequencing
1) Cdc25-22 arrest 35C-3.5hrs. 2-8) timecourse through S phase	yFS128	Sequence awaited	*S.pombe*	Deep sequencing
1) G2 Synchronized + 2) HU arrest 4 hours	yFS286	Sharp peaks- peaks being verified as origins experimentally	*S.octosporus*	Deep sequencing
1) G2 Synchronized + 2) HU arrest 4 hours	yFS275	Not many peaks due to lack of HU arrest	*S.japonicus*	Deep sequencing
HU arrest 2 hours	yFS275	Sequence awaited	*S.japonicus*	Deep sequencing

Chapter V

Unpublished data

APPENDIX V.1. ORIGIN INHIBITION BY DNA DAMAGE CHECKPOINT

INTRODUCTION

DNA damage during S phase leads to a slowing of replication. Studies in mammals have shown that slowing occurs by inhibition of origin firing and a slowing of fork progression (Falck *et al.* 2002). However, it is unclear as to how the DNA damage checkpoint slows replication in fission yeast. Recent work by Nick Willis in our lab shows that in the presence of DNA damage, replication forks are slowed and hence S-phase progression is slowed. This result does not rule out the possibility that replication origins are also prevented from firing once the cells encounter damage. Recent work has suggested that a combination of replication fork slowing and inhibition of origin firing is the mechanism by which DNA damage checkpoint slows replication (Kumar and Huberman 2009). I performed an experiment to look at *ars3001*, known efficient origin present in multiple copies in the rDNA loci, in the presence or absence of damage (0.03% MMS) using two-dimensional gel electrophoresis. In the presence of DNA damage *ars3001* stopped firing as observed by the absence of bubble arcs, which are present when there is no damage.

MATERIALS AND METHODS

yFS128, a *cdc25-22* mutant was grown to an OD_{600} 0.6 and kept at the restrictive temperature of 35°C for four hours which arrests and synchronizes cells in G2. The cells were then shifted to the permissive temperature of 25°C in the presence of HU for two hours to synchronize and arrest cells at the beginning of S phase. The first time point was collected at the end of the HU arrest. The remaining culture was pelleted, HU washed, and the pellet resuspended in a HU free media. The culture was divided into two and 0.03% MMS was added to one of the cultures. The cells were kept at 25°C to allow for progression through S phase and samples were collected at 30 and 60 minutes from + or – MMS cultures. Cells were also fixed every 20 minutes and nuclear DNA content measured by flow cytometry.

Two-Dimensional gel electrophoresis

Genomic DNA was isolated using cesium chloride gradients and two-Dimensional gel electrophoresis performed as described (Noguchi *et al.* 2003). To study origin firing at *ars3001*, genomic DNA was digested with KpnI and HindIII. Southern blotting was performed as described (Noguchi *et al.* 2003).

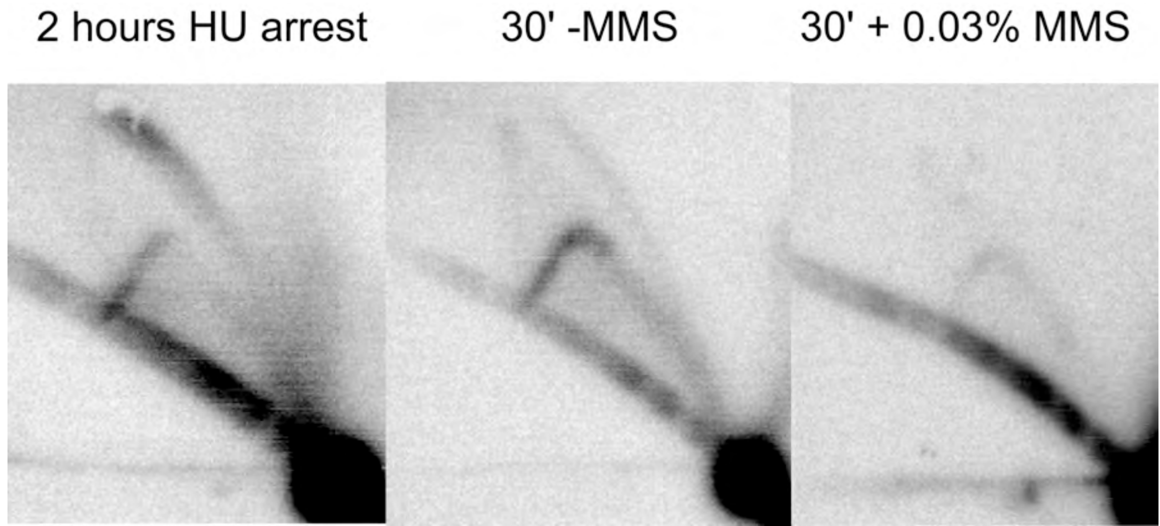
RESULTS

Figure V.1 shows the 2-D gels for the *ars3001* origin in the presence or absence of MMS. *ars3001* is an efficient origin and fires early represented by the presence of bubble arcs on the 2-D gel. In the absence of any insult, when cells are released from HU arrest, replication is still on going at 30 minutes evident by the presence of the bubble arc at the 30' –MMS 2-D gel. However, in the presence of MMS, origin firing is inhibited inferred from a loss of bubble arc on the 30' + MMS 2-D gel. Replication is still continuing because Y-arcs can be seen which implies that the forks that have already fired are passively replicating the genome but no further origins are being fired.

CONCLUSIONS

The data shown here suggests that origins are inhibited from firing in the presence of DNA damage. Origin inhibition might not be responsible for bulk slowing but it certainly may play a role in slowing DNA replication. Further experiments involving different drugs and ionizing radiation (IR) can be done using 2-D gels allowing us to understand the mechanism of replication slowing.

Figure V.1 Origin inhibition in the presence of damage



APPENDIX V.2 USING MICROARRAYS TO MEASURE REPLICATION ORIGIN FIRING EFFICIENCY

INTRODUCTION

Studies from our lab showed that Hsk1-Dfp1 kinase regulates origin efficiency (Patel *et al.* 2008). Hsk1-Dfp1 is rate limiting and is required at each origin to fire. Level of Dfp1, the catalytic subunit of Hsk1, increases at the beginning of the cell cycle and it is a freely diffusible factor. DNA combing studies from our lab showed that modulating the levels of either Dfp1 or Hsk1 affects the origin efficiency (Patel *et al.* 2008). We sought to show a similar effect on origin efficiency using microarrays. We designed probes to cover the well studied *ura4* locus present on chromosome III of *S. pombe.* Experiments were performed on wild type and Dfp1 overexpression strains to measure the difference in origin efficiencies when Dfp1 is constitutively active or localized to an origin.

MATERIALS AND METHODS

Synchronization experiments

For the microarray HU experiments, cells were grown to an OD_{600} 0.5 and collected for G2 phase sample. 10 mM HU was added to the culture and the cells were collected after 2, 3 or 4 hours for the S phase sample for different experiments.

For the microarray time course experiments, cells were synchronized by centrifugal elutriation followed by synchronization using the *cdc25-22* mutation which arrests cells in G2 at restrictive temperature of 35°C. Cells were kept at 35°C for 3.5 hours. The cells are then released from the block by shifting to the permissive temperature of 25°C. Samples were collected at the indicated times and cell cycle progression followed by flow cytometry.

Micro-array design

Probes were designed to cover 200 kb region around the *ura4* locus using the Arraydesigner 4.2. The average distance between the probes was 250 bp and 768 probes were designed with an average length of 60 bp. The slides for microarrays were printed in the Rando lab and the slides post-processed using the protocols followed in their lab.

DNA preparation and microarray experiment

Genomic DNA was isolated using cesium chloride gradients as described (Noguchi *et al.* 2003). DNA was indirectly labeled to cy3 and cy5 dyes using Amino-allyl labeling protocol from DeRisi lab with a few modifications used in our lab (http://derisilab.ucsf.edu/data/microarray/protocols.html) (Dutta *et al.* 2008). Experimental DNA was mixed with the reference DNA, which was the G2 samples for all our experiments, for differential hybridization. The sample was hybridized onto the microarray slides for 16 hours at 65°C. Slides were scanned using Genepix5000b scanner and the data was acquired using Genepix pro 6.0 software. The data was normalized and replication profiles created using excel.

RESULTS AND DISCUSSION

To measure the change in origin efficiency when Dfp1 was constitutively expressed, we designed probes for 200 kb region of chromosome III including the *ura4* locus which has well defined origins (Dubey *et al.* 1994). We used microarrays to study origin efficiency as measured by the change in DNA copy number described previously in chapter III. To study if we could look at origin efficiency using this technique, we did HU arrest experiments where cells were arrested for different periods of time. HU arrests cells in early S phase where early firing origins have fired and traveled about 10 kb and then arrested. We did HU arrest experiments for four hours using wild type cells (figure V.2A). Although we observed the firing of all the origins known in *ura4* locus, the peaks were very broad and it seemed that forks are able to travel longer making it hard to estimate the exact location of origin firing. We proceeded to shorter HU arrests of three and two hours (figure V.2B and 2C). Figure V.2C shows that two hour HU arrests gave us sharp peaks closer to the known origins and we used two hour HU arrests for subsequent experiments.

The positions of peaks identified in our microarrays correspond to already known origins in the *ura4* region. To ensure that the peaks observed in our experiments are not random noise, we label G2 samples with two different dyes and hybridize on the array to give us self-self hybrizidization. There are no peaks in the G2 control array indicating that the peaks identified are actual origins and not random noise (figure V.3). Hybridizing two G2 samples with different labels

served as the negative control for all experiments and also indicated the level of noise for each array experiment.

Dfp1 effect on origin firing was studied using two different strains. We used a strain in which *dfp1* was constitutively expressed using an *adh1* promoter. Studies from our lab showed that origin efficiency increased globally when Dfp1 was expressed continuously indicating an effect on origin efficiency. Dfp1 was also tethered near the origin *AT3003* using Gal4 DNA binding domain (DBD) as described in appendix III.1. To measure the increase in origin efficiency in the *adh1:dfp1* cells, we performed the two hour HU arrest experiment in the wild type and *adh1:dfp1 cells*. Replication profile of the *ura4* locus in the wild type cells is shown in Figure V.4A. The replication profile shown is an average of three independent experiments and shows all the known origins represented by peaks in the *ura4* locus. Replication profile of *adh1:dfp1* has a similar pattern to the wild type cells (figure V.4B). Overlays of the replication profiles of wild type and *adh1:dfp1* indicate no change in origin efficiency when *dfp1* is constitutively expressed (figure V.4C).. We also looked at the replication profiles when *dfp1* is tethered to *AT3003*. Although we observed the known origins firing in the *ura4* locus, the data is noisy and it is difficult to make conclusions from it (Data not shown). Based on our combing data, we expected an increase in origin efficiency by around 15-20%, but it is possible that the microarray data has a low signal to noise ratio to detect that magnitude of change.

Figure V.2 Timing of HU arrest to look at replication profile of *ura4* locus

Cells were arrested in early S phase using 10 mM HU in wild type *S.pombe* yFS240. Known origins are represented with ▬. 2A) Cells were arrested in HU for four hours. 2B) Cells were arrested in HU for three hours. 2C) Cells were arrested in HU for four hours. The peaks representing the origins become sharper as the HU arrest time is reduced. *AT3004/3005* is a combination of two origins, which we cannot separate on our array due to the resolution. The data shows that two hours is the suitable time for HU arrest experiments.

Figure V.2 Timing of HU arrest to look at replication profile of *ura4* locus

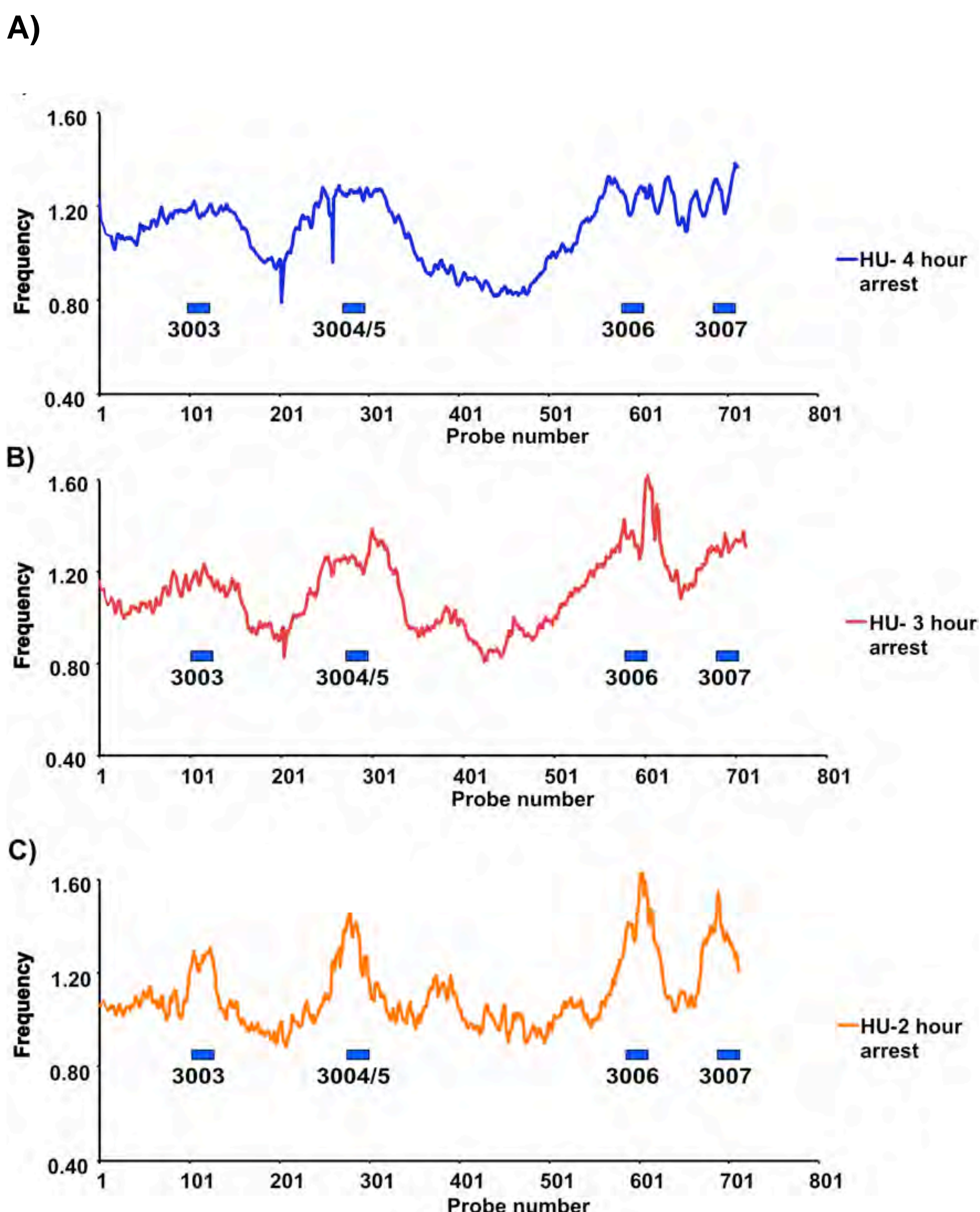
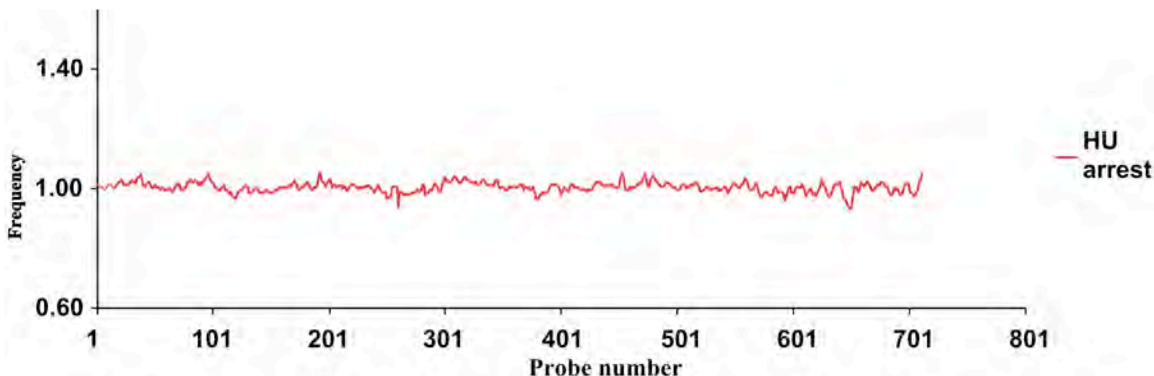


Figure V.3 Control for HU arrest profiles using Self hybridizations

Cells were arrested in early S phase using HU in yFS240 and is the same experiment performed in the previous figure. The replication profile is from the control array where two G2 samples were labeled was Cy3 and Cy5 and hybridized to the array. The array shows the noise level for the HU experiment. The profile has been created by averaging the control arrays for all the three HU experiments.

Figure V.3 Control for HU arrest profiles using Self hybridizations



A time course synchronization experiment was performed and replication profiles for cells during S phase were made. The S phase progression was followed by flow cytometry. Figure V.5 shows the replication profiles from 80 to 95 minutes. Although the origins in the *ura4* locus seem to fire during this time course, there is no noticeable change in the profiles itself at the various points. The lack of change in the replication profiles can be due to a lack of synchrony or due to the resolution of the microarrays itself. We used a double synchronization protocol to ensure maximum synchrony and the absence of any difference in the profiles indicates that the change in efficiency we are looking for cannot be observed using microarrays. We have also done early time course points (60-75 minutes) and see no peaks at all indicating that the origins have not fired yet (Data not shown). We tried to capture the transition of an origin from no firing, to firing, and collected cells every 10 minutes or 20 minutes. However, these datasets were very noisy and we have been unable to reach any conclusions from them (Data not shown).

Figure V.4 No change in replication profile when dfp1 is overexpressed

Cells were arrested in early S phase using 10 mM HU in wild type yFS240 and *adh1:dfp1* yFS458. Cells were arrested in HU for two hours. 4A) Replication profile of wild type cells. 4B) Replication profile of *adh1:dfp1* cells. The known origins are represented by . There seems to be no increase in origin efficiency in these cells. 4C) Replication profile overlays of yFS240 and yFS458 show no noticeable difference in the origin efficiency.

Figure V.4 No change in replication profile when dfp1 is overexpressed

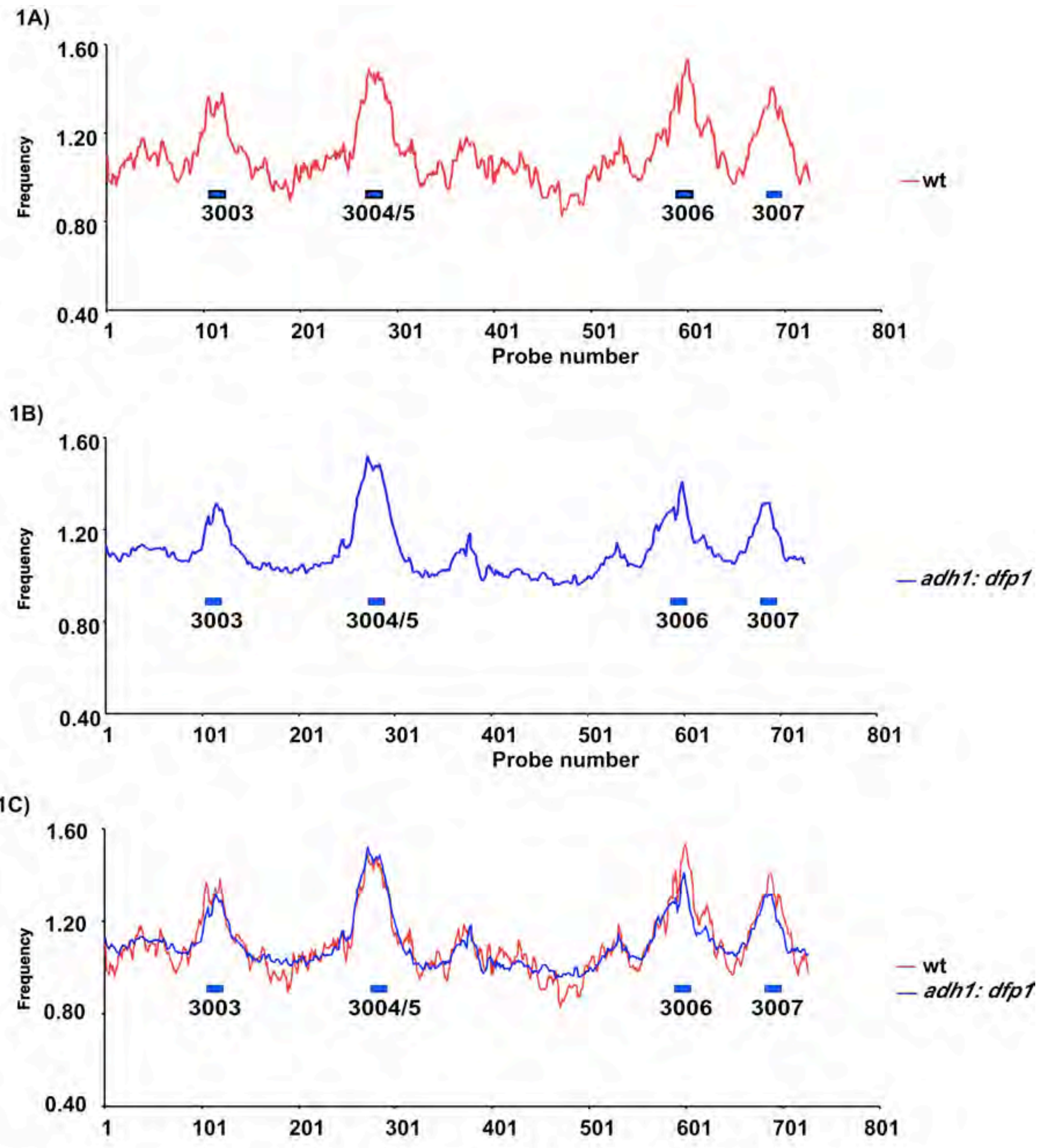
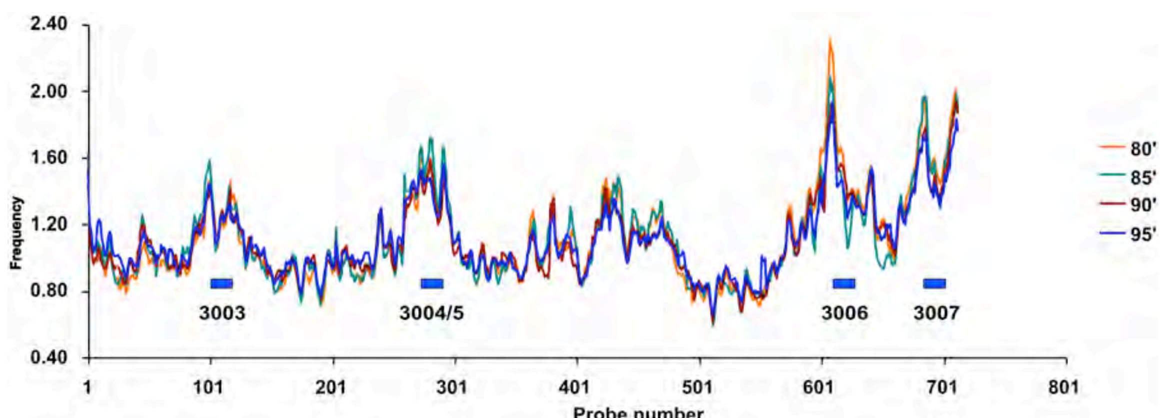


Figure V.5 Replication kinetics at *ura4* locus

cdc25-22 mutant cells were synchronized in G2 by centrifugal elutriation. The cells were then arrested at 35 C for 3.5 hours. Cells were collected at 80', 85', 90' and 95' and replication profile made as described earlier. The replication profile shows no change in the origin firing over time as represented by no change in the peaks indicating that we are unable to see the small change in origin efficiency observed using DNA combing.

Figure V.5 Replication kinetics at ura locus



CONCLUSION

Microarrays have been used as a method to measure a change in copy number to study change in origin efficiencies in various mutant strains. Although, we are able to identify known origins from the replication profiles, we have been unable to notice a change in origin efficiency. The absence of any changes in the replication profiles during various time courses shows that our setup has a low signal to noise ratio. Microarrays can be used to measure changes in origin efficiency using density transfer method instead of measuring the change in DNA copy number. The data shows that microarrays can be used for origin studies and different origin studies can still employ microarrays. The resolution of my setup was also significantly higher than the previous origin mapping studies and such a high-resolution array can be made for the entire genome. We are also able to see a couple of putative origins, which have not been identified in the previous genomic studies and we can use this system to look at other regions for identifying new origins.

APPENDIX V.3. EFFECTS OF MODULATING DFP1 LEVELS

INTRODUCTION

Studies from our lab showed that origin firing is stochastic in fission yeast. We have been working on identifying the factor responsible for the stochasticity and randomness of origin firing. We were able to identify Hsk1-Dfp1 as the kinase that regulates origin efficiency (Patel *et al.* 2008). Hsk1-Dfp1 is rate limiting and is required at each origin to fire. Level of Dfp1, the catalytic subunit of Hsk1, increases at the beginning of the cell cycle and it is a freely diffusible factor. DNA combing studies from our lab showed that modulating the levels of either Dfp1 or Hsk1 affects the origin efficiency (Patel *et al.* 2008). This appendix presents my work towards showing that Hsk1-Dfp1 is responsible for regulating origin efficiency.

RESULTS AND DISCUSSION

To increase the activity of the Hsk1-Dfp1 kinase, we overexpressed Dfp1 from the constitutive *adh1* promoter, leading to an approximately 3-fold increase in Dfp1 protein and Hsk1-Dfp1 kinase activity relative to wild-type S-phase levels (Figure V.6). The *adh1:dfp1* cells grow normally and have normal bulk replication kinetics by flow cytometry.

To test if over-expression of Dfp1 interferes with or activities the replication checkpoint in fission yeast, we assayed HU sensitivity and Cds1 kinase activity in *adh1:dfp1* cell. We find no evidence of HU sensitivity, Cds1 inhibition or Cds1 activation, suggesting that the effects of Dfp1 over-expression are not due to indirect effects on the replication checkpoint (Figure V.7).

Although the simplest explanation for the effect of tethering Hsk1-Dfp1 on the efficiency of local origins in that Hsk1-Dfp1 is directly activating the origins by phosphorylating MCM, it is also possible that the local high concentration of Hsk1-Dfp1 affects local chromatin structure, which in turn indirectly affects local origin efficiency. We reasoned that any effect on local chromatin structure that would affect origin efficiency would also affect transcription. Therefore, to test for local chromatin affects, we used genome-wide transcriptional profiling to assay transcript levels in cells with and without Hsk1-Dfp1 tethered near the ura4 locus (Oliva et al. 2005). We find no significant difference in transcript levels near ura4 between wild-type cells and cells with Gal4-Dfp1 tethered at AT3003

Figure V.6 Expression levels of *dfp1* alleles.

A) Protein levels measured by Western blot. Cells were elutriation synchronized and harvested in S phase as determined by septation index and flow cytometry. 150 µg of whole cell lysate was separated by SDS-PAGE on a 10% gel, transferred to a PVDF membrane and visualized using anti-Dpf1antibodies as previously described (Takeda et al., 1999). The bands representing Dpf1 and the Dpf1-2xGFP fusion are indicated; asterisks indicate non-specific bands. The membrane was reprobed with anti-tubulin antibodies. When normalized to the tubulin control, the adh1-expressed Dfp1 is approximately 3-fold more abundant that the wild-type Dfp1 and the Dfp1-2xGFP is approximately equal.

B) Protein activity measured by *in vitro* kinase assay. Cells were elutriation synchronized and harvested in S phase as determined by septation index and flow cytometry. IP kinase assay was performed as described, using polyclonal anti-Dfp1 antibodies and myelin basic protein as substrate (Takeda et al.1999). Lanes 1 and 2 are wild type (yFS240) cells; lane 3 is *adh1:dfp1* (yFS458) cells. Lane 1 is a mock IP, using no antibody; lanes 2 and 3 are Dpf1 IPs. Quantitation of activity is shown below the figure in arbitrary units with the background in Lane 1 subtracted.

Figure V.6 Expression levels of *dfp1* alleles

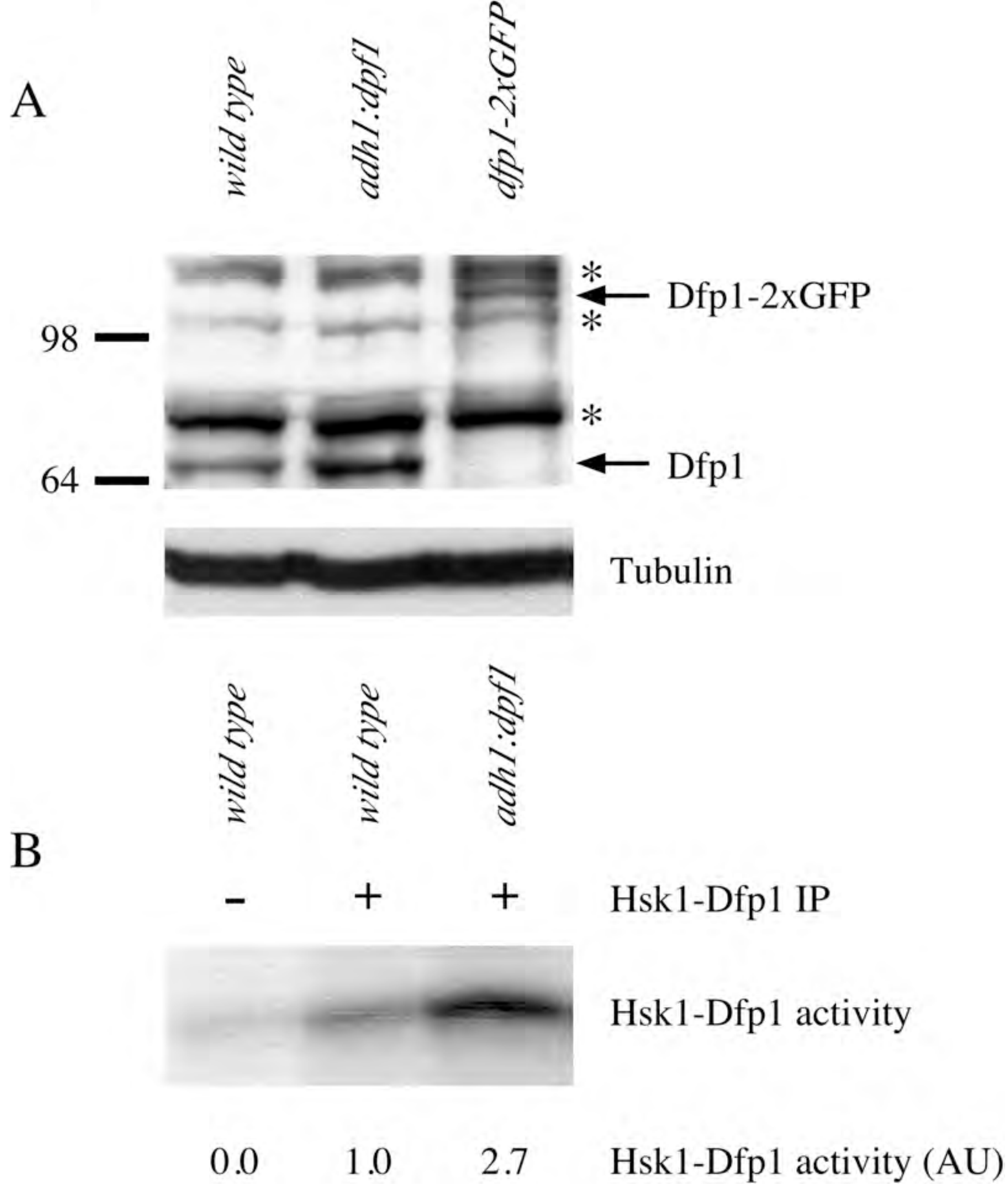
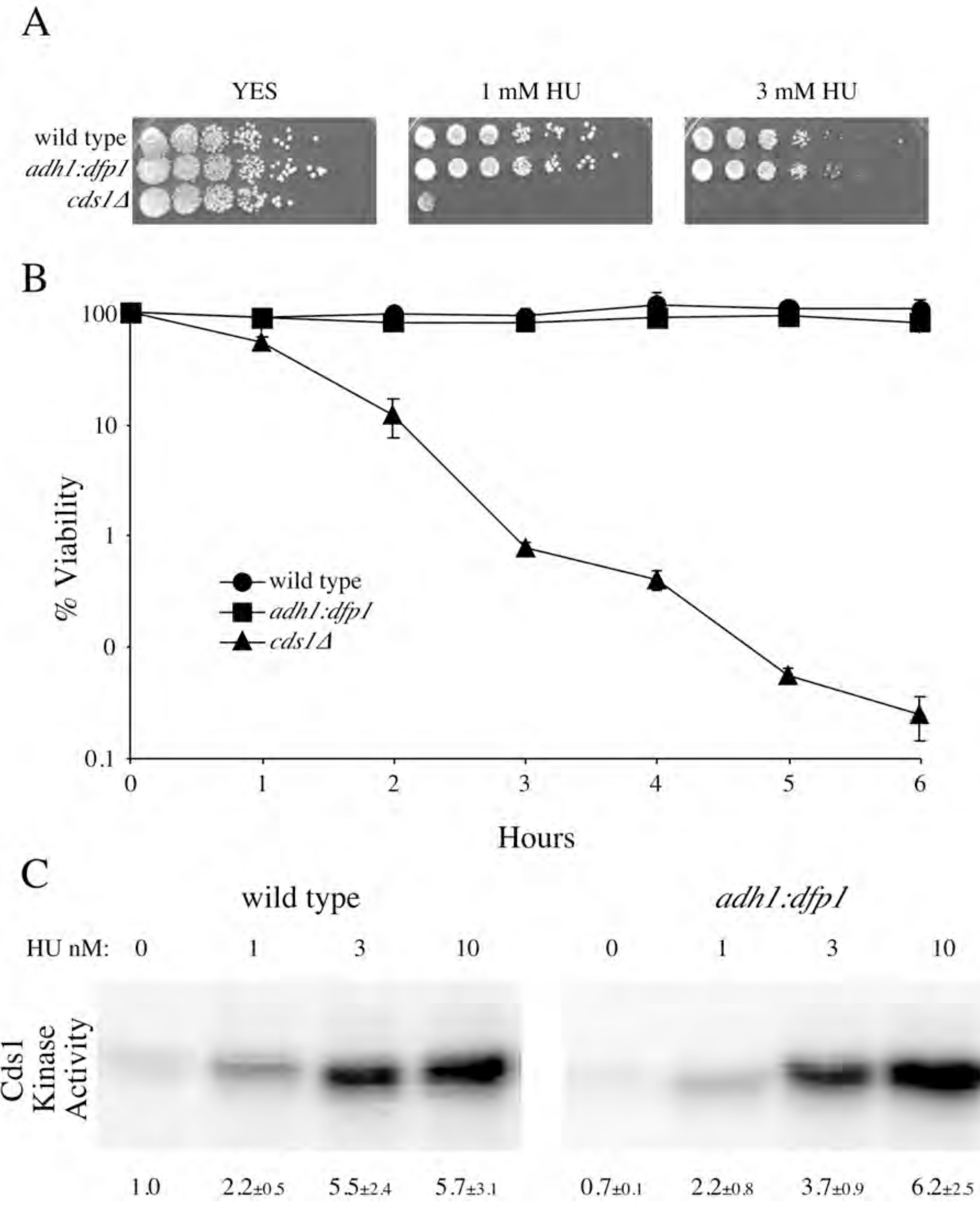


Figure V.7 Over-expression of Dfp1 does not activate or inhibit the replication checkpoint.

A) Dfp1 over-expressing cells are not sensitive to chronic exposure to HU. Wild-type (yFS240), *adh1:dfp1* (yFS458) and *cds1::ura4* (yFS199) cells were grown to mid-log, 10-fold serially diluted, spotted onto YES plates containing 0, 1 or 3 mM HU and grown for 5 days.

B) Dfp1 over-expressing cells are not sensitive to acute exposure to HU. Wild-type (yFS240), *adh1:dfp1* (yFS458) and *cds1::ura4* (yFS199) cells were grown to mid-log, transferred to YES containing 10 mM HU, grown for the indicated time, plated on YES, grown for 5 days and counted. Data points represent mean +/- s.e.m.; n = 4. C) Dfp1 over-expressing cells activate Cds1 normally in response to HU. Wild-type (yFS240), *adh1:dfp1* (yFS458) and *cds1::ura4* (yFS199) cells were grown to mid-log, transferred to YES containing 10 mM HU for 4 hours and harvested. Cds1 was immunoprecipitated from 10 OD pellets and assayed by *in vitro* kinase assay using myelin basic protein as a substrate (Lindsay et al., 1998). Quantitation is mean ± SEM; n is 3 or 4.

Figure V.7 Over-expression of Dfp1 does not activate or inhibit the replication checkpoint.



(Figure V.8). Specifically the change in transcript levels for the 24 genes with in the 50 kb around AT3003 is 1.02 fold, as compared to a genome wide change of 1.01 fold ($p > 0.2$). These results suggest that the increased local concentration of Hsk1-Dfp1 is not affecting origin efficiency indirectly through local chromatin effects.

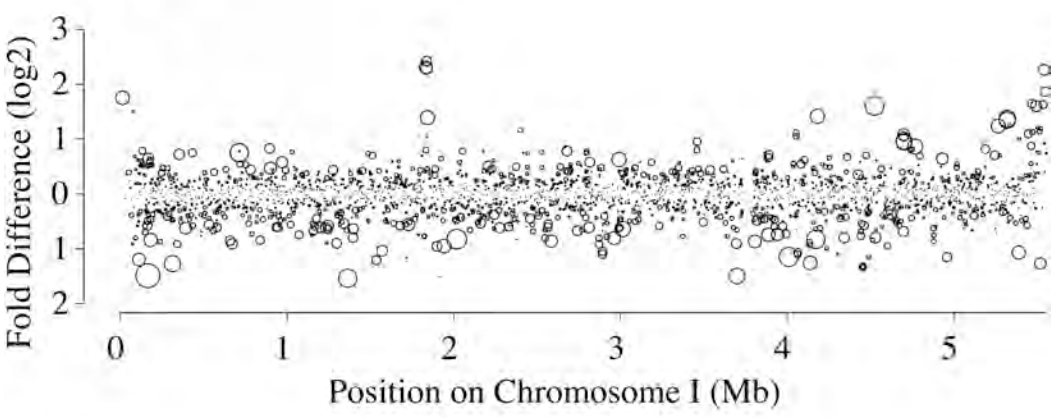
CONCLUSION

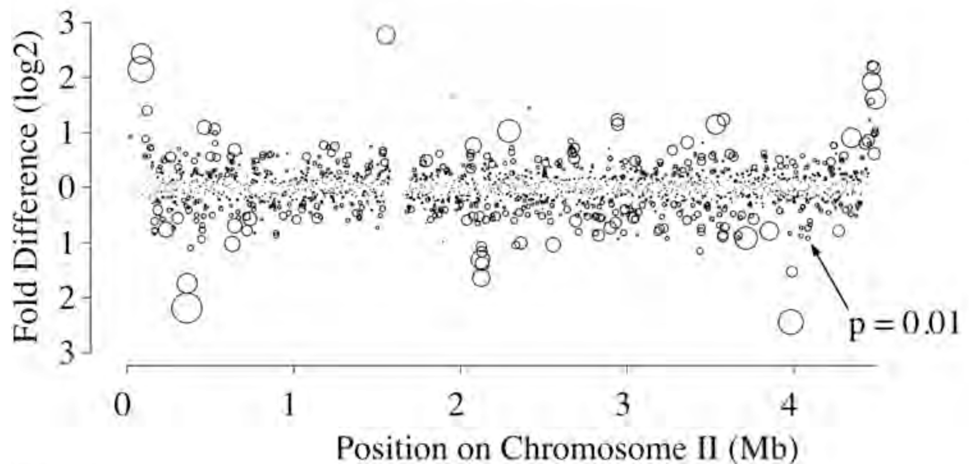
Using DNA combing and other methods we identified Hsk1-Dfp1 as the regulatory kinase responsible for firing efficiency. I used microarrays to measure a change in firing efficiency when Hsk1-Dfp1 levels are increased globally and locally. Here we show that putting Dfp1 under *adh1* promoter increases the Dfp1 protein and kinase levels 3-fold and does not activate a checkpoint response. Increasing local Dfp1 levels also does not change firing efficiency indirectly by having any effect on local chromatin state.

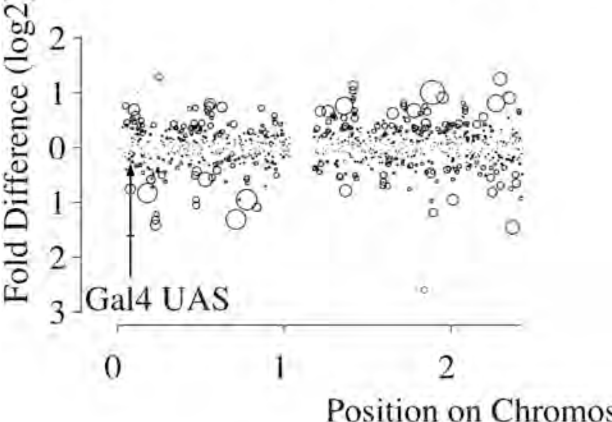
Figure V.8 Genome-wide transcript levels of in cells with Gal4-Dfp1 tethered at *AT3003*.

Relative transcript levels in wild-type and *5xGal4 UAS:AT3003 Gal4-Dfp1* (yFS459) cells were determined by competitive hybridization of labeled cDNA to an microarray containing probes for all 5004 pombe annotated ORFs as described (Oliva et al., 2005). Wild-type (yFS105) cDNA was used as a reference in both cases to control for dye bias. The figures show relative difference in transcript levels between the two strains (Log 2) versus chromosome position. Relative p-values are shown by circle size; a circle of p = 0.01 is indicated on Chromosome 2. 98% (5282/5414) probes showed less than a two-fold difference between the two strains. The location of the Gal4 UAS site on Chromosome 3 is indicated; the bar shows the 50 kb surrounding the sites. All array data will be available at ArrayExpress (www.ebi.ac.uk/arrayexpress).

Figure V.8 Genome-wide transcript levels of in cells with Gal4-Dfp1 tethered at *AT3003*.







Chapter VI

Discussion and Future Directions

Maintenance of DNA replication fidelity during S phase is essential to prevent cells from disastrous consequences. Cells must replicate in an efficient, timely and error-free manner. Cells ensure error-free manner of replication through cell cycle checkpoints, which make sure that cells do not replicate when encountering damage. To ensure that replication is completed in an efficient and timely manner, the cells regulate origin firing. Checkpoints and origin regulation are the crucial components of a successful S phase. During my thesis research, I have studied the various aspects of checkpoints and origin regulation. My studies have also indicated origin regulation during DNA damage by the cell cycle checkpoints.

Mechanism of slowing and role of Cdc25 in damage checkpoint

Studies from the metazoans have identified the downstream targets for the origin regulation by the checkpoint (Falck *et al.* 2002). Another mechanism by which cells can slow replication is by slowing the progression of forks in the presence of DNA damage. Such a mechanism has been shown in mammals when damage is induced by MMS (Merrick *et al.* 2004). Recent work from our lab shows that in the presence of damage, fork progression is slowed (Willis N, personal communication). However, it remains to be seen whether checkpoints affect replication only by slowing the fork progression or also by inhibiting origin firing. Using 2-D gels, I did preliminary experiments where DNA damage was induced using MMS. Looking at rDNA origin *ars3001*, origin firing seems to be prevented in the presence of damage. It is possible that inhibition of origin firing

might not contribute to the bulk slowing observed due to a relatively minor role in slowing replication. This result along with Nick Willis's results seems to favor slowing by inhibition of origin firing as well as slowing the replication fork progression.

When I began my thesis research, it was not clear as to what the downstream targets were for the S-phase DNA damage checkpoints in fission yeast. Work from the Huberman lab suggested that Cdc25 is the target for the checkpoint and the checkpoint functions in a similar manner to the metazoans (Kumar and Huberman 2004). It is believed that different levels of Cdc2 trigger different functions of Cdc2 and dephosphorylation by Cdc25 during S phase would lead to catastrophic mitosis (Lundgren *et al.* 1991; Stern and Nurse 1996). Cdc2 is also known to remain phosphorylated during S phase (Gould and Nurse 1989). Cdc25 levels are also known to be low during S phase (Moreno *et al.* 1990). Due to these reasons we did not believe that Cdc25 would be a downstream target of S-phase damage checkpoint. In an attempt to resolve this confusion in the field, we initiated experiments using mutant strains in which *cdc25* is over-expressed or *cdc25* is deleted. Studies from our lab showed that asynchronous as well as different synchronized cultures of these mutant strains had a slowing of S phase comparable to a wild type strain in the presence of damage. Although data from the Pyp3 studies shows a lack of slowing when *cdc25* is deleted, based on the literature and data from *cdc25* deletion in the *cdc2-Y15F* background we drew the conclusion that Cdc25 is not a target of the

S-phase DNA damage checkpoint (Kommajosyula and Rhind 2006). The difference in the results obtained in these studies is primarily due to the kind of flow cytometry being used in both the labs. We use isolated-nuclei approach as opposed to a whole-cell method used by the Huberman lab. Our protocol increases the resolution of the assay and is quantifiable allowing us to reproducibly detect checkpoint dependent slowing in situations in which is slowing is not apparent in whole-cell flow cytometry experiments. Recently, work from their lab has raised concerns over our paper and I have addressed those concerns in the discussion.

Hsk1 as an alternate target

If Cdc25 is not acting downstream of the damage checkpoint, then we need to identify the target of the checkpoint in order to understand how replication is slowed in the presence of damage. The downstream target of the S-phase damage checkpoint could be Hsk1. Previous work has suggested that Hsk1 may play a role in S-phase damage checkpoint in fission yeast (Snaith *et al.* 2000; Sommariva *et al.* 2005). Hsk1 interacts with Cds1 and is phosphorylated by it. However, Hsk1 is essential to cells and its deletion leads to lethality due to its requirement in replication initiation. To study the role of Hsk1 in S-phase damage checkpoint, we need to bypass the replication function of Hsk1. Such a bypass has been done in budding yeast where a mutation in *mcm5* (P83L) is able to bypass the Cdc7 or Dbf4 requirement for origin firing (Hardy *et al.* 1997). This mutant is known as the *mcm5-bob1.* The corresponding amino

acid is at position 85 in fission yeast. We initiated work to replace the endogenous *mcm5* with *bob-1* by transforming the *mcm5 P85L* into wild-type fission yeast. However, due to the recessive nature of *mcm5-bob1*, where bob1 will express only in the absence of the wild type *mcm5*, we were unable to get transformants. Future work in the lab will focus on making the *bob-1* mutant in fission yeast. Hsk1 will be deleted in the *bob-1* background and then the effect of DNA damage on replication can be studied. If Hsk1 is involved in the DNA damage checkpoint, then its deletion will lead to the absence of slowing when damage is induced using methyl methane sulfonate (MMS). If cells slow replication when presented with damage in the *hsk1* delete, then Hsk1 is not a target of the DNA damage checkpoint.

It is possible that the checkpoint may function by not targeting either Cdc25 or Hsk1. Nevertheless, it is important to identify the targets for the DNA damage checkpoint to give us a clear understanding of the mechanism by which this checkpoint functions. Identifying the downstream targets for the DNA damage checkpoint will advance the field of S phase checkpoints in fission yeast.

Origin efficiency studies

In the second part of my thesis, I have tried to measure the efficiency of a late replicating sequence. Previous studies and work from our lab have demonstrated that origin firing is stochastic in nature (Patel *et al.* 2006). The stochastic nature of origin firing could lead to a potential problem in replicating by

taking long time to finish replicating DNA stretches where no origin may fire (Lucas *et al.* 2000; Herrick *et al.* 2002; Jun *et al.* 2004). This potential problem has been termed as the random gap problem. Work in *Xenopus* embryos have suggested that the efficiency of a particular origin increases as the cells progress through S phase (Lucas *et al.* 2000; Herrick *et al.* 2002). We proposed that a similar mechanism in fission yeast was responsible for efficient replication in spite of inefficient origin firing. To study such a phenomenon, we have to measure the efficiency of a late replicating sequence. Two things made this project challenging: the absence of well defined late firing origins and passive replication of origins, which are capable of firing late.

Regions are capable of firing late in S phase, but if they have not fired in the early stages, the chances are that replication forks from neighboring origins would passively replicate them. We set out to solve the problem of measuring efficiency of a late firing origin by preventing the passive replication from neighboring origins. The block was established using fork terminators known as RTS1 sequences (Dalgaard and Klar 2001). The RTS1 sequences are present at the mating locus in fission yeast and it has been shown that RTS1 is capable of blocking replication when placed at other sites in the genome (Codlin and Dalgaard 2003; Lambert *et al.* 2005). We integrated RTS1 sites on either side of a late replicating sequence *AT2062*. Using a combination of methods we tried to study the efficiency of this region. Initial results from two-Dimensional gel electrophoresis showed that passive replication was occurring at *AT2062*. We

also used microarrays and deep sequencing to look at the 80 kb region flanked by RTS1 sites. Deep sequencing revealed several inefficient origins yet to be identified and we surmised that when forks from neighboring regions were prevented from passively replicating *AT2062*, these inefficient origins would be responsible for the replication of the 80 kb region. We looked to make sure that forks were getting blocked at RTS1 sites and the forks were getting arrested. However, the block seemed to be weak leading to the possibility that forks were beginning to bypass the block. Deletion of *rad51* in the strain having RTS1 flanking *AT2062* had no effect on the viability of the strain. Deleting *rad51* prevents any bypass of RTS1 sites by neighboring forks and hence replication of the 80 kb region is due to origins present within the region. We realized that *AT2062* was not a good choice for studying origin efficiency due to the large distance between the RTS1 sites. I have recently focused my attention on integrating RTS1 sites at *AT3003*, a well-defined origin which fires during the early S phase. The RTS1 sites are being integrated at *AT3003* about 7.5 kb apart on either side. To ensure that neighboring forks do not bypass the RTS1, I am integrating the fork blocks in a *rad51*Δ background.

A biochemical explanation for the increasing efficiency model would be a rate-limiting factor required at each origin for it fire. Factors essential for origin firing could be an ideal candidate. One of the prime candidates for regulating origin firing was Hsk1. We showed that Dfp1, the regulatory subunit of Hsk1, was indeed a rate-limiting factor (Patel *et al.* 2008). Increasing Dfp1 levels in the cell

led to an increase in a global increase in origin efficiency as determined by DNA combing. Reducing Hsk1 levels by using a temperature sensitive allele led to a decrease in origin efficiency. Also, when the local concentration of Dfp1 was experimentally increased near one of the origins in the *ura4* locus, an increase in efficiency was noted for that particular origin. I measured the protein levels of Dfp1 as well as kinase activity in the Dfp1 overexpression strains and there was a three-fold increase in the Dfp1 activity. Normal activation of checkpoints was observed in the dfp1 overexpression strain showing that the effects observed were not the indirect results of checkpoints.

I used microarrays and 2-D gels to show similar effects of increase in origin efficiency using a different method. Because we were looking to see an increase in efficiency by only about 10-20%, these techniques were not sensitive enough. Although I observed replication profiles using microarrays at the *ura4* locus, I was unable to note a change in efficiency due to the high background noise when measuring the change in copy number using oligonucleotide arrays.

Ways of using RTS1 sites to study origin efficiency

Another mechanism to prevent forks from bypassing RTS1 is by putting *rad51* under an *nmt1* promoter. The presence or absence of thiamine allows for shutting off the *nmt1* promoter thereby shutting off *rad51*. Experiments will be performed in the presence of thiamine to prevent *rad51* expression. Efficiency of

AT3003 will be measured using two-Dimensional gel electrophoresis as well as deep sequencing described in chapter IV.

We hope to address the random gap problem by the efficiency studies at *AT3003*. There are several lines of data that suggest that increasing origin efficiency is the way cells deal with stochastic firing. Mathematical modeling (Monte-carlo simulations) done in collaboration with Bechhoefer lab show that inefficient origins firing randomly would lead to longer replication times whereas increasing the efficiency of origins through the S phase leads to completion of replication in a finite amount of time (Rhind N, personal communication). Studies from *Xenopus* embryos have led to similar models being proposed (Lucas *et al.* 2000). Finally, identifying Hsk1-dfp1 as the rate-limiting kinase responsible for origin efficiency suggests that our model may be correct. Understanding the mechanism of how cells avoid the random gap problem will lead to a more realistic picture of how origin regulation works.

However, the direct evidence for the increasing efficiency model will come from measuring the efficiency of a potentially later firing origin, which usually gets passively replicated before it has a chance to fire. Presently, the work is focused upon measuring the efficiency at *AT3003*. In the future a cryptic origin like *ars727* can be targeted.

One of the advantages of using RTS1 to measure firing efficiencies will be to study the effect of DNA damage on late firing origins in fission yeast. Studies

in budding yeast have shown that DNA damage during S phase prevents late origins from firing (Shirahige *et al.* 1998). However, the mechanism by which S phase DNA damage checkpoint slows replication is not known. Isolating a potential late firing origin from getting passively replicated will allow us to study the effect of drugs like methane methyl sulfonate (MMS) on replication. The advantages of these studies lies in the fact that mammalian origins are similar to fission yeast origins and the results inferred from DNA damage studies in fission yeast can be extrapolated on mammals.

Identifying replication origins

We recently started genome-wide search for *Schizosaccharomyces* genus using deep sequencing. Looking at origin efficiency at *AT2062*, we realized that the 80 kb region flanked by RTS1 seemed to have some inefficient origins which have not been identified. The reason for this is the low resolution of the various genome-wide analyses for origins in fission yeast. The only known common features of origins in *S. pombe* are the presence of AT rich islands, AT asymmetry and their presence in intergenic regions. We are using sequencing in collaboration with Helicos Biosciences and the Weng lab. Preliminary studies of the replication profiles in *S. pombe* show that the origins identified by sequencing correlates very well with previously characterized origins from other studies. The peak finding algorithm in the Igor software being used for studying replication profiles has identified potential origins in both *S. octosporus* and *S. japonicus*. The identification of peaks as well as correlation with previously identified origins

makes deep sequencing a good method for identifying origins. The reason for identifying origins in three different fission yeast species is an attempt to overcome a lack of sequence specificity between the known origins of *S. pombe*. Identification of sequences determining origin activity can help in finding origins in fission yeast *de novo.* The origins are large and inefficient similar to mammals. Resolution of genome wide origin studies in mammals is only about 100 kb which is very low for identifying specific regions acting as origins. Sequencing can also be used to identify origins in mammals and should provide sufficiently higher resolution compared to the previous studies. Identifying origins will also lead to the advancement of research in *S. japonicus* and *S. octosporus* where origins are yet to be identified. Studying the origins across the three distant pombe species will also give us an idea about the evolutionary divergence of replication amongst these fission yeasts. To identify these origins I have performed the experiments using HU arrest in the early S phase in all the three fission yeast species. I have also done a time course in *S. pombe* in order to make replication kinetic profiles and see the how replication progresses through the genome. Bioinformatic analysis will be done on these datasets and common motifs identified. The origins identified in *S. octosporus* in our studies do not have AT rich islands and due to this I believe that it is not simply the presence of AT rich islands that defines an origin but a sequence motif that is yet to be identified. The results from our sequence analysis should help us in understanding the nature of the origins and identify this sequence motif. The absence of origins defined by

AT rich islands, markers for origin prediction in *S. pombe*, in *S. octosporus* and *S. japonicus,* is the proof that some other sequence specificity governs origins.

Potential uses of using sequencing to identify origins

In the future, the effect of nucleosome positioning on origin firing can also be studied using deep sequencing. The presence of nucleosomes at a particular region would alter the firing potential of origins. Genome-wide analysis of such a nucleosomal effect on origin regulation can be studied using sequencing.

Effects of DNA damage can also be studied using sequencing. Cells can be sequenced in the presence or absence of DNA damage like MMS and the corresponding replication profiles can be compared. If origin firing is inhibited upon damage, then we will see the absence of origins in the replication profiles of damage induced cells.

Deep sequencing approach will open up many avenues of research especially in identifying origins in organisms with newly completed genome sequences. This is clearly seen by finding potential origins in *S. octosporus* and *S. japonicus*. The potential impact on mammalian origin studies can also be huge.

SUMMARY

Origin regulation and checkpoints have been the focus of my thesis research. My results have shown that neither tyrosine-15 phosphorylation of Cdc2, nor Cdc25 itself, is involved in the S-phase DNA damage checkpoint in fission yeast. I have also studied the origin efficiency of a late replicating sequence. Although, my work has not proven that firing efficiency increases through S phase, my work has setup the platform for future studies pertaining to origin efficiency. This work can be pursued in a number of different ways. This work also has the potential to help in future studies pertaining to DNA damage and becoming a good model to study the effects of damage on origin firing. Finally, I have collaborated with other labs to carry out a genome-wide analysis of origins as well as finding origins in three *Schizosaccharomyces* species. Two of these *Schizosaccharomyces* species did not have previously identified origins and this work identifies the origins across the genome in these species for the first time. This will advance the field of origin studies in *Schizosaccharomyces* species and information gained from this study can potentially be used to identify origins de novo in mammals.

References

Adachi, Y., J. Usukura and M. Yanagida (1997). "A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast." Genes Cells **2**(7): 467-79.

Aparicio, O. M., D. M. Weinstein and S. P. Bell (1997). "Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase." Cell **91**(1): 59-69.

Arias, E. E. and J. C. Walter (2006). "PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication." Nat Cell Biol **8**(1): 84-90.

Bartek, J., C. Lukas and J. Lukas (2004). "Checking on DNA damage in S phase." Nat Rev Mol Cell Biol **5**(10): 792-804.

Bell, S. P. and A. Dutta (2002). "DNA REPLICATION IN EUKARYOTIC CELLS." Annual Review of Biochemistry **71**(1): 333-374.

Bell, S. P., J. Mitchell, J. Leber, R. Kobayashi and B. Stillman (1995). "The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing." Cell **83**(4): 563-8.

Bell, S. P. and B. Stillman (1992). "ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex." Nature **357**(6374): 128-34.

Bentley, N. J., D. A. Holtzman, G. Flaggs, K. S. Keegan, A. DeMaggio, J. C. Ford, M. Hoekstra and A. M. Carr (1996). "The Schizosaccharomyces pombe rad3 checkpoint gene." Embo J **15**(23): 6641-51.

Blow, J. J., P. J. Gillespie, D. Francis and D. A. Jackson (2001). "Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times." J Cell Biol **152**(1): 15-25.

Bochman, M. L. and A. Schwacha (2008). "The Mcm2-7 complex has in vitro helicase activity." Mol Cell **31**(2): 287-93.

Bousset, K. and J. F. Diffley (1998). "The Cdc7 protein kinase is required for origin firing during S phase." Genes Dev **12**(4): 480-90.

Brewer, B. J. and W. L. Fangman (1987). "The localization of replication origins on ARS plasmids in S. cerevisiae." Cell **51**(3): 463-71.

Brewer, B. J. and W. L. Fangman (1991). "Mapping replication origins in yeast chromosomes." Bioessays **13**(7): 317-22.

Burhans, W. C. and J. A. Huberman (1994). "DNA replication origins in animal cells: a question of context?" Science **263**(5147): 639-40.

Carlson, C. R., B. Grallert, R. Bernander, T. Stokke and E. Boye (1997). "Measurement of nuclear DNA content in fission yeast by flow cytometry." Yeast **13**(14): 1329-35.

Carpenter, P. B., P. R. Mueller and W. G. Dunphy (1996). "Role for a Xenopus Orc2-related protein in controlling DNA replication." Nature **379**(6563): 357-60.

Chahwan, C., T. M. Nakamura, S. Sivakumar, P. Russell and N. Rhind (2003). "The fission yeast Rad32 (Mre11)-Rad50-Nbs1 complex is required for the S-phase DNA damage checkpoint." Molecular and Cellular Biology **23**(18): 6564-6573.

Christensen, P. U., N. J. Bentley, R. G. Martinho, O. Nielsen and A. M. Carr (2000). "Mik1 levels accumulate in S phase and may mediate an intrinsic link between S phase and mitosis." Proc Natl Acad Sci U S A **97**(6): 2579-84.

Chuang, R. Y., L. Chretien, J. Dai and T. J. Kelly (2002). "Purification and characterization of the Schizosaccharomyces pombe origin recognition complex: interaction with origin DNA and Cdc18 protein." J Biol Chem **277**(19): 16920-7.

Chuang, R. Y. and T. J. Kelly (1999). "The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks." Proc Natl Acad Sci U S A **96**(6): 2656-61.

Clyne, R. K. and T. J. Kelly (1995). "Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe." Embo J **14**(24): 6348-57.

Codlin, S. and J. Z. Dalgaard (2003). "Complex mechanism of site-specific DNA replication termination in fission yeast." The EMBO Journal **22**(13): 3431-3440.

Coleman, T. R., P. B. Carpenter and W. G. Dunphy (1996). "The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts." Cell **87**(1): 53-63.

Cook, J. G., D. A. Chasse and J. R. Nevins (2004). "The regulated association of Cdt1 with minichromosome maintenance proteins and Cdc6 in mammalian cells." J Biol Chem **279**(10): 9625-33.

Costanzo, V., K. Robertson, M. Bibikova, E. Kim, D. Grieco, M. Gottesman, D. Carroll and J. Gautier (2001). "Mre11 protein complex prevents double-strand break accumulation during chromosomal DNA replication." Mol Cell **8**(1): 137-47.

Costanzo, V., K. Robertson, C. Y. Ying, E. Kim, E. Avvedimento, M. Gottesman, D. Grieco and J. Gautier (2000). "Reconstitution of an ATM-dependent checkpoint that inhibits chromosomal DNA replication following DNA damage." Mol Cell **6**(3): 649-59.

Coue, M., S. E. Kearsey and M. Mechali (1996). "Chromotin binding, nuclear localization and phosphorylation of Xenopus cdc21 are cell-cycle dependent and associated with the control of initiation of DNA replication." EMBO J **15**(5): 1085-97.

Czajkowsky, D. M., J. Liu, J. L. Hamlin and Z. Shao (2008). "DNA Combing Reveals Intrinsic Temporal Disorder in the Replication of Yeast Chromosome VI." Journal of Molecular Biology **375**(1): 12-19.

Dai, J., R. Y. Chuang and T. J. Kelly (2005). "DNA replication origins in the Schizosaccharomyces pombe genome." Proc Natl Acad Sci U S A **102**(2): 337-42.

Dalgaard, J. Z. and A. J. Klar (2001). "A DNA replication-arrest site RTS1 regulates imprinting by determining the direction of replication at mat1 in S. pombe." Genes & Development **15**(16): 2060-2068.
Delmolino, L. M., P. Saha and A. Dutta (2001). "Multiple mechanisms regulate subcellular localization of human CDC6." J Biol Chem **276**(29): 26947-54.
Diffley, J. F. (1996). "Once and only once upon a time: specifying and regulating origins of DNA replication in eukaryotic cells." Genes Dev **10**(22): 2819-30.
Dijkwel, P. A., S. Wang and J. L. Hamlin (2002). "Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies." Mol Cell Biol **22**(9): 3053-65.
Dolan, W. P., D. A. Sherman and S. L. Forsburg (2004). "Schizosaccharomyces pombe replication protein Cdc45/Sna41 requires Hsk1/Cdc7 and Rad4/Cut5 for chromatin binding." Chromosoma **113**(3): 145-156.
Donaldson, A. D., W. L. Fangman and B. J. Brewer (1998). "Cdc7 is required throughout the yeast S phase to activate replication origins." Genes Dev **12**(4): 491-501.
Donovan, S., J. Harwood, L. S. Drury and J. F. Diffley (1997). "Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast." Proc Natl Acad Sci U S A **94**(11): 5611-6.
Drury, L. S., G. Perkins and J. F. Diffley (1997). "The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast." Embo J **16**(19): 5966-76.
Dubey, D. D., S. M. Kim, I. T. Todorov and J. A. Huberman (1996). "Large, complex modular structure of a fission yeast DNA replication origin." Curr Biol **6**(4): 467-73.
Dubey, D. D., J. Zhu, D. L. Carlson, K. Sharma and J. A. Huberman (1994). "Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe." Embo J **13**(15): 3638-47.
Dutta, A. and S. P. Bell (1997). "Initiation of DNA replication in eukaryotic cells." Annual Review of Cell and Developmental Biology **13**: 293-332.
Dutta, C., P. K. Patel, A. Rosebrock, A. Oliva, J. Leatherwood and N. Rhind (2008). "The DNA replication checkpoint directly regulates MBF-dependent G1/S transcription." Molecular and Cellular Biology **28**(19): 5977-5985.
Elledge, S. J. (1996). "Cell cycle checkpoints: preventing an identity crisis." Science **274**(5293): 1664-72.
Ellison, V. and B. Stillman (2001). "Opening of the clamp: an intimate view of an ATP-driven biological machine." Cell **106**(6): 655-60.
Enoch, T. and P. Nurse (1990). "Mutation of fission yeast cell cycle control genes abolishes dependence of mitosis on DNA replication." Cell **60**(4): 665-73.
Eshaghi, M., R. K. Karuturi, J. Li, Z. Chu, E. T. Liu and J. Liu (2007). "Global profiling of DNA replication timing and efficiency reveals that efficient

replication/firing occurs late during S-phase in S. pombe." PLoS One **2**(1): e722.
Eydmann, T., E. Sommariva, T. Inagawa, S. Mian, A. J. S. Klar and J. Z. Dalgaard (2008). "Rtf1-mediated eukaryotic site-specific replication termination." Genetics **180**(1): 27-39.
Falck, J., N. Mailand, R. G. Syljuasen, J. Bartek and J. Lukas (2001). "The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis." Nature **410**(6830): 842-7.
Falck, J., J. H. Petrini, B. R. Williams, J. Lukas and J. Bartek (2002). "The DNA damage-dependent intra-S phase checkpoint is regulated by parallel pathways." Nat Genet **30**(3): 290-4.
Fangman, W. L. and B. J. Brewer (1992). "A question of time: replication origins of eukaryotic chromosomes." Cell **71**(3): 363-6.
Feng, W., D. Collingwood, M. E. Boeck, L. A. Fox, G. M. Alvino, W. L. Fangman, M. K. Raghuraman and B. J. Brewer (2006). "Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication." Nat Cell Biol **8**(2): 148-55.
Fisher, D. L. and P. Nurse (1996). "A single fission yeast mitotic cyclin B p34cdc2 kinase promotes both S-phase and mitosis in the absence of G1 cyclins." EMBO J **15**(4): 850-60.
Forsburg, S. L. (2004). "Eukaryotic MCM proteins: beyond replication initiation." Microbiol Mol Biol Rev **68**(1): 109-31, table of contents.
Forsburg, S. L. and N. Rhind (2006). "Basic methods for fission yeast." Yeast (Chichester, England) **23**(3): 173-183.
Friedman, K. L., J. D. Diller, B. M. Ferguson, S. V. Nyland, B. J. Brewer and W. L. Fangman (1996). "Multiple determinants controlling activation of yeast replication origins late in S phase." Genes Dev **10**(13): 1595-607.
Friedman, K. L., M. K. Raghuraman, W. L. Fangman and B. J. Brewer (1995). "Analysis of the temporal program of replication initiation in yeast chromosomes." J Cell Sci Suppl **19**: 51-8.
Fujita, M., Y. Hori, K. Shirahige, T. Tsurimoto, H. Yoshikawa and C. Obuse (1998). "Cell cycle dependent topological changes of chromosomal replication origins in Saccharomyces cerevisiae." Genes Cells **3**(11): 737-49.
Gilbert, D. M. (2001). "Making Sense of Eukaryotic DNA Replication Origins." Science **294**(5540): 96-100.
Goldar, A., M. C. Marsolier-Kergoat and O. Hyrien (2009). "Universal temporal profile of replication origin activation in eukaryotes." PLoS One **4**(6): e5899.
Gomez, M. and F. Antequera (1999). "Organization of DNA replication origins in the fission yeast genome." EMBO J **18**(20): 5683-90.
Gould, K. L., S. Moreno, N. K. Tonks and P. Nurse (1990). "Complementation of the mitotic activator, p80cdc25, by a human protein-tyrosine phosphatase." Science **250**(4987): 1573-6.

Gould, K. L. and P. Nurse (1989). "Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis." Nature **342**(6245): 39-45.

Gregan, J., K. Lindner, L. Brimage, R. Franklin, M. Namdar, E. A. Hart, S. J. Aves and S. E. Kearsey (2003). "Fission yeast Cdc23/Mcm10 functions after pre-replicative complex formation to promote Cdc45 chromatin binding." Mol Biol Cell **14**(9): 3876-87.

Hardy, C. F., O. Dryga, S. Seematter, P. M. Pahl and R. A. Sclafani (1997). "mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p." Proc Natl Acad Sci U S A **94**(7): 3151-5.

Hartwell, L., T. Weinert, L. Kadyk and B. Garvik (1994). "Cell cycle checkpoints, genomic integrity, and cancer." Cold Spring Harb Symp Quant Biol **59**: 259-63.

Harvey, K. J. and J. Newport (2003). "Metazoan origin selection: origin recognition complex chromatin binding is regulated by CDC6 recruitment and ATP hydrolysis." J Biol Chem **278**(49): 48524-8.

Hayashi, M., Y. Katou, T. Itoh, A. Tazumi, Y. Yamada, T. Takahashi, T. Nakagawa, K. Shirahige and H. Masukata (2007). "Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast." EMBO J **26**(5): 1327-39.

Hayashi, M. T., T. S. Takahashi, T. Nakagawa, J. Nakayama and H. Masukata (2009). "The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus." Nat Cell Biol **11**(3): 357-62.

Heichinger, C., C. J. Penkett, J. Bahler and P. Nurse (2006). "Genome-wide characterization of fission yeast DNA replication origins." EMBO J **25**(21): 5171-9.

Henry-Mowatt, J., D. Jackson, J. Y. Masson, P. A. Johnson, P. M. Clements, F. E. Benson, L. H. Thompson, S. Takeda, S. C. West and K. W. Caldecott (2003). "XRCC3 and Rad51 modulate replication fork progression on damaged vertebrate chromosomes." Mol Cell **11**(4): 1109-17.

Herrick, J., S. Jun, J. Bechhoefer and A. Bensimon (2002). "Kinetic model of DNA replication in eukaryotic organisms." J Mol Biol **320**(4): 741-50.

Herrick, J., P. Stanislawski, O. Hyrien and A. Bensimon (2000). "Replication fork density increases during DNA synthesis in X. laevis egg extracts." Journal of Molecular Biology **300**(5): 1133-1142.

Homesley, L., M. Lei, Y. Kawasaki, S. Sawyer, T. Christensen and B. K. Tye (2000). "Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins." Genes Dev **14**(8): 913-26.

Hyrien, O., K. Marheineke and A. Goldar (2003). "Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem." Bioessays **25**(2): 116-25.

Hyrien, O. and M. Mechali (1993). "Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of Xenopus early embryos." Embo J **12**(12): 4511-20.
Hyrien, O. and M. Méchali (1993). "Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of Xenopus early embryos." The EMBO Journal **12**(12): 4511-4520.
Itzhaki, J. E., C. S. Gilbert and A. C. Porter (1997). "Construction by gene targeting in human cells of a "conditional' CDC2 mutant that rereplicates its DNA." Nat Genet **15**(3): 258-65.
Izumi, M., K. Yanagi, T. Mizuno, M. Yokoi, Y. Kawasaki, K. Y. Moon, J. Hurwitz, F. Yatagai and F. Hanaoka (2000). "The human homolog of Saccharomyces cerevisiae Mcm10 interacts with replication factors and dissociates from nuclease-resistant nuclear structures in G(2) phase." Nucleic Acids Res **28**(23): 4769-77.
Izumi, M., F. Yatagai and F. Hanaoka (2001). "Cell cycle-dependent proteolysis and phosphorylation of human Mcm10." J Biol Chem **276**(51): 48526-31.
Jackson, A. L., P. M. Pahl, K. Harrison, J. Rosamond and R. A. Sclafani (1993). "Cell cycle regulation of the yeast Cdc7 protein kinase by association with the Dbf4 protein." Mol Cell Biol **13**(5): 2899-908.
Jallepalli, P. V., G. W. Brown, M. Muzi-Falconi, D. Tien and T. J. Kelly (1997). "Regulation of the replication initiator protein p65cdc18 by CDK phosphorylation." Genes Dev **11**(21): 2767-79.
Jallepalli, P. V., D. Tien and T. J. Kelly (1998). "sud1(+) targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1 proteins for degradation and stops unwanted diploidization in fission yeast." Proc Natl Acad Sci U S A **95**(14): 8159-64.
Jares, P. and J. J. Blow (2000). "Xenopus cdc7 function is dependent on licensing but not on XORC, XCdc6, or CDK activity and is required for XCdc45 loading." Genes Dev **14**(12): 1528-40.
Jares, P., A. Donaldson and J. J. Blow (2000). "The Cdc7/Dbf4 protein kinase: target of the S phase checkpoint?" EMBO Rep **1**(4): 319-22.
Jeon, Y., S. Bekiranov, N. Karnani, P. Kapranov, S. Ghosh, D. MacAlpine, C. Lee, D. S. Hwang, T. R. Gingeras and A. Dutta (2005). "Temporal profile of replication of human chromosomes." Proc Natl Acad Sci U S A **102**(18): 6419-24.
Johnston, L. H., H. Masai and A. Sugino (1999). "First the CDKs, now the DDKs." Trends Cell Biol **9**(7): 249-52.
Jonsson, Z. O. and U. Hubscher (1997). "Proliferating cell nuclear antigen: more than a clamp for DNA polymerases." Bioessays **19**(11): 967-75.
Jun, S., J. Herrick, A. Bensimon and J. Bechhoefer (2004). "Persistence length of chromatin determines origin spacing in Xenopus early-embryo DNA replication: quantitative comparisons between theory and experiment." Cell Cycle (Georgetown, Tex.) **3**(2): 223-229.

Kastan, M. B. and D. S. Lim (2000). "The many substrates and functions of ATM." Nat Rev Mol Cell Biol **1**(3): 179-86.
Kelly, T. J. and G. W. Brown (2000). "Regulation of chromosome replication." Annual Review of Biochemistry **69**: 829-880.
Kim, J. M., M. Yamada and H. Masai (2003). "Functions of mammalian Cdc7 kinase in initiation/monitoring of DNA replication and development." Mutat Res **532**(1-2): 29-40.
Kim, S. M. and J. A. Huberman (1999). "Influence of a replication enhancer on the hierarchy of origin efficiencies within a cluster of DNA replication origins." J Mol Biol **288**(5): 867-82.
Kim, S. M. and J. A. Huberman (2001). "Regulation of replication timing in fission yeast." EMBO J **20**(21): 6115-26.
Kommajosyula, N. and N. Rhind (2006). "Cdc2 tyrosine phosphorylation is not required for the S-phase DNA damage checkpoint in fission yeast." Cell Cycle (Georgetown, Tex.) **5**(21): 2495-2500.
Kong, D. and M. L. DePamphilis (2002). "Site-specific ORC binding, pre-replication complex assembly and DNA synthesis at Schizosaccharomyces pombe replication origins." Embo J **21**(20): 5567-76.
Krysan, P. J. and M. P. Calos (1991). "Replication initiates at multiple locations on an autonomously replicating plasmid in human cells." Molecular and Cellular Biology **11**(3): 1464-1472.
Kumar, S. and J. A. Huberman (2004). "On the slowing of S phase in response to DNA damage in fission yeast." J Biol Chem **279**(42): 43574-80.
Kumar, S. and J. A. Huberman (2009). "Checkpoint-dependent regulation of origin firing and replication fork movement in response to DNA damage in fission yeast." Mol Cell Biol **29**(2): 602-11.
Labib, K. and J. F. Diffley (2001). "Is the MCM2-7 complex the eukaryotic DNA replication fork helicase?" Curr Opin Genet Dev **11**(1): 64-70.
Labib, K., J. F. Diffley and S. E. Kearsey (1999). "G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus." Nat Cell Biol **1**(7): 415-22.
Labib, K. and A. Gambus (2007). "A key role for the GINS complex at DNA replication forks." Trends in Cell Biology **17**(6): 271-278.
Labib, K., J. A. Tercero and J. F. Diffley (2000). "Uninterrupted MCM2-7 function required for DNA replication fork progression." Science **288**(5471): 1643-7.
Lambert, S., A. Watson, D. M. Sheedy, B. Martin and A. M. Carr (2005). "Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier." Cell **121**(5): 689-702.
Larner, J. M., H. Lee, R. D. Little, P. A. Dijkwel, C. L. Schildkraut and J. L. Hamlin (1999). "Radiation down-regulates replication origin activity throughout the S phase in mammalian cells." Nucleic Acids Res **27**(3): 803-9.

Lee, D. G., A. M. Makhov, R. D. Klemm, J. D. Griffith and S. P. Bell (2000). "Regulation of origin recognition complex conformation and ATPase activity: differential effects of single-stranded and double-stranded DNA binding." Embo J **19**(17): 4774-82.

Lee, J. K., K. Y. Moon, Y. Jiang and J. Hurwitz (2001). "The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein." Proc Natl Acad Sci U S A **98**(24): 13589-94.

Lee, J. K., Y. S. Seo and J. Hurwitz (2003). "The Cdc23 (Mcm10) protein is required for the phosphorylation of minichromosome maintenance complex by the Dfp1-Hsk1 kinase." Proc Natl Acad Sci U S A **100**(5): 2334-9.

Legouras, I., G. Xouri, S. Dimopoulos, J. Lygeros and Z. Lygerou (2006). "DNA replication in the fission yeast: robustness in the face of uncertainty." Yeast **23**(13): 951-62.

Lei, M., Y. Kawasaki, M. R. Young, M. Kihara, A. Sugino and B. K. Tye (1997). "Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis." Genes Dev **11**(24): 3365-74.

Lei, M. and B. K. Tye (2001). "Initiating DNA synthesis: from recruiting to activating the MCM complex." J Cell Sci **114**(Pt 8): 1447-54.

Lindsay, H. D., D. J. Griffiths, R. J. Edwards, P. U. Christensen, J. M. Murray, F. Osman, N. Walworth and A. M. Carr (1998). "S-phase-specific activation of Cds1 kinase defines a subpathway of the checkpoint response in Schizosaccharomyces pombe." Genes Dev **12**(3): 382-95.

Liu, E., X. Li, F. Yan, Q. Zhao and X. Wu (2004). "Cyclin-dependent kinases phosphorylate human Cdt1 and induce its degradation." J Biol Chem **279**(17): 17283-8.

Liu, J., C. L. Smith, D. DeRyckere, K. DeAngelis, G. S. Martin and J. M. Berger (2000). "Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control." Mol Cell **6**(3): 637-48.

Lopes, M., C. Cotta-Ramusino, A. Pellicioli, G. Liberi, P. Plevani, M. Muzi-Falconi, C. S. Newlon and M. Foiani (2001). "The DNA replication checkpoint response stabilizes stalled replication forks." Nature **412**(6846): 557-61.

Lucas, I., M. Chevrier-Miller, J. M. Sogo and O. Hyrien (2000). "Mechanisms ensuring rapid and complete DNA replication despite random initiation in Xenopus early embryos." Journal of Molecular Biology **296**(3): 769-786.

Lundgren, K., N. Walworth, R. Booher, M. Dembski, M. Kirschner and D. Beach (1991). "mik1 and wee1 cooperate in the inhibitory tyrosine phosphorylation of cdc2." Cell **64**(6): 1111-22.

MacAlpine, D. and S. Bell (2005). "A genomic view of eukaryotic DNA replication." Chromosome Research **13**(3): 309-326.

MacAlpine, D. M., H. K. Rodriguez and S. P. Bell (2004). "Coordination of replication and transcription along a Drosophila chromosome." Genes Dev **18**(24): 3094-105.

Mailand, N., J. Falck, C. Lukas, R. G. Syljuasen, M. Welcker, J. Bartek and J. Lukas (2000). "Rapid destruction of human Cdc25A in response to DNA damage." Science **288**(5470): 1425-9.

Maiorano, D., W. Rul and M. Mechali (2004). "Cell cycle regulation of the licensing activity of Cdt1 in Xenopus laevis." Exp Cell Res **295**(1): 138-49.

Marchetti, M. A., S. Kumar, E. Hartsuiker, M. Maftahi, A. M. Carr, G. A. Freyer, W. C. Burhans and J. A. Huberman (2002). "A single unbranched S-phase DNA damage and replication fork blockage checkpoint pathway." Proc Natl Acad Sci U S A **99**(11): 7472-7.

Masai, H. and K. Arai (2002). "Cdc7 kinase complex: a key regulator in the initiation of DNA replication." J Cell Physiol **190**(3): 287-96.

McGarry, T. J. and M. W. Kirschner (1998). "Geminin, an inhibitor of DNA replication, is degraded during mitosis." Cell **93**(6): 1043-53.

Merrick, C. J., D. Jackson and J. F. Diffley (2004). "Visualization of altered replication dynamics after DNA damage in human cells." J Biol Chem **279**(19): 20067-75.

Mickle, K. L., S. Ramanathan, A. Rosebrock, A. Oliva, A. Chaudari, C. Yompakdee, D. Scott, J. Leatherwood and J. A. Huberman (2007). "Checkpoint independence of most DNA replication origins in fission yeast." BMC Mol Biol **8**: 112.

Millar, J. B., G. Lenaers and P. Russell (1992). "Pyp3 PTPase acts as a mitotic inducer in fission yeast." EMBO J **11**(13): 4933-41.

Mizushima, T., N. Takahashi and B. Stillman (2000). "Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro." Genes Dev **14**(13): 1631-41.

Moreno, S., P. Nurse and P. Russell (1990). "Regulation of mitosis by cyclic accumulation of p80cdc25 mitotic inducer in fission yeast." Nature **344**(6266): 549-52.

Moser, B. A., J. M. Brondello, B. Baber-Furnari and P. Russell (2000). "Mechanism of caffeine-induced checkpoint override in fission yeast." Mol Cell Biol **20**(12): 4288-94.

Moyer, S. E., P. W. Lewis and M. R. Botchan (2006). "Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase." Proc Natl Acad Sci U S A **103**(27): 10236-41.

Myung, K., A. Datta and R. D. Kolodner (2001). "Suppression of spontaneous chromosomal rearrangements by S phase checkpoint functions in Saccharomyces cerevisiae." Cell **104**(3): 397-408.

Natale, D. A., C. J. Li, W. H. Sun and M. L. DePamphilis (2000). "Selective instability of Orc1 protein accounts for the absence of functional origin recognition complexes during the M-G(1) transition in mammals." EMBO J **19**(11): 2728-38.

Newlon, C. S. and J. F. Theis (1993). "The structure and function of yeast ARS elements." Curr Opin Genet Dev **3**(5): 752-8.
Nguyen, V. Q., C. Co, K. Irie and J. J. Li (2000). "Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7." Curr Biol **10**(4): 195-205.
Nguyen, V. Q., C. Co and J. J. Li (2001). "Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms." Nature **411**(6841): 1068-73.
Nieduszynski, C. A., Y. Knox and A. D. Donaldson (2006). "Genome-wide identification of replication origins in yeast by comparative genomics." Genes Dev **20**(14): 1874-9.
Nishitani, H., Z. Lygerou, T. Nishimoto and P. Nurse (2000). "The Cdt1 protein is required to license DNA for replication in fission yeast." Nature **404**(6778): 625-8.
Noguchi, E., C. Noguchi, L. L. Du and P. Russell (2003). "Swi1 prevents replication fork collapse and controls checkpoint kinase Cds1." Mol Cell Biol **23**(21): 7861-74.
Ogawa, Y., T. Takahashi and H. Masukata (1999). "Association of fission yeast Orp1 and Mcm6 proteins with chromosomal replication origins." Molecular and Cellular Biology **19**(10): 7228-7236.
Okuno, Y., A. J. McNairn, N. den Elzen, J. Pines and D. M. Gilbert (2001). "Stability, chromatin association and functional activity of mammalian pre-replication complex proteins during the cell cycle." Embo J **20**(15): 4263-77.
Oliva, A., A. Rosebrock, F. Ferrezuelo, S. Pyne, H. Chen, S. Skiena, B. Futcher and J. Leatherwood (2005). "The cell cycle-regulated genes of Schizosaccharomyces pombe." PLoS Biol **3**(7): e225.
Orren, D. K., L. N. Petersen and V. A. Bohr (1997). "Persistent DNA damage inhibits S-phase and G2 progression, and results in apoptosis." Mol Biol Cell **8**(6): 1129-42.
Painter, R. B. and B. R. Young (1980). "Radiosensitivity in ataxia-telangiectasia: a new explanation." Proc Natl Acad Sci U S A **77**(12): 7315-7.
Pasion, S. G. and S. L. Forsburg (1999). "Nuclear localization of Schizosaccharomyces pombe Mcm2/Cdc19p requires MCM complex assembly." Mol Biol Cell **10**(12): 4043-57.
Patel, P. K., B. Arcangioli, S. P. Baker, A. Bensimon and N. Rhind (2006). "DNA replication origins fire stochastically in fission yeast." Molecular Biology of the Cell **17**(1): 308-316.
Patel, P. K., N. Kommajosyula, A. Rosebrock, A. Bensimon, J. Leatherwood, J. Bechhoefer and N. Rhind (2008). "The Hsk1(Cdc7) replication kinase regulates origin efficiency." Molecular Biology of the Cell **19**(12): 5550-5558.

Perkins, G. and J. F. Diffley (1998). "Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders." Mol Cell **2**(1): 23-32.

Perkins, G., L. S. Drury and J. F. Diffley (2001). "Separate SCF(CDC4) recognition elements target Cdc6 for proteolysis in S phase and mitosis." Embo J **20**(17): 4836-45.

Petersen, B. O., C. Wagener, F. Marinoni, E. R. Kramer, M. Melixetian, E. Lazzerini Denchi, C. Gieffers, C. Matteucci, J. M. Peters and K. Helin (2000). "Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1." Genes Dev **14**(18): 2330-43.

Petrini, J. H. (2000). "The Mre11 complex and ATM: collaborating to navigate S phase." Curr Opin Cell Biol **12**(3): 293-6.

Raghuraman, M. K., E. A. Winzeler, D. Collingwood, S. Hunt, L. Wodicka, A. Conway, D. J. Lockhart, R. W. Davis, B. J. Brewer and W. L. Fangman (2001). "Replication dynamics of the yeast genome." Science **294**(5540): 115-21.

Randell, J. C., J. L. Bowers, H. K. Rodriguez and S. P. Bell (2006). "Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase." Mol Cell **21**(1): 29-39.

Reynolds, A. E., R. M. McCarroll, C. S. Newlon and W. L. Fangman (1989). "Time of replication of ARS elements along yeast chromosome III." Mol Cell Biol **9**(10): 4488-94.

Rhind, N., B. Furnari and P. Russell (1997). "Cdc2 tyrosine phosphorylation is required for the DNA damage checkpoint in fission yeast." Genes Dev **11**(4): 504-11.

Rhind, N. and P. Russell (1998). "The Schizosaccharomyces pombe S-phase checkpoint differentiates between different types of DNA damage." Genetics **149**(4): 1729-37.

Rhind, N. and P. Russell (1998). "Tyrosine phosphorylation of cdc2 is required for the replication checkpoint in Schizosaccharomyces pombe." Mol Cell Biol **18**(7): 3782-7.

Rhind, N. and P. Russell (2000). "Checkpoints: it takes more than time to heal some wounds." Current Biology: CB **10**(24): R908-911-R908-911.

Rhind, N. and P. Russell (2000). "Chk1 and Cds1: linchpins of the DNA damage and replication checkpoint pathways." Journal of Cell Science **113 (Pt 22)**: 3889-3896.

Rhind, N. and P. Russell (2001). "Roles of the mitotic inhibitors Wee1 and Mik1 in the G(2) DNA damage and replication checkpoints." Molecular and Cellular Biology **21**(5): 1499-1508.

Romanowski, P., M. A. Madine, A. Rowles, J. J. Blow and R. A. Laskey (1996). "The Xenopus origin recognition complex is essential for DNA replication and MCM binding to chromatin." Curr Biol **6**(11): 1416-25.

Rowles, A., S. Tada and J. J. Blow (1999). "Changes in association of the Xenopus origin recognition complex with chromatin on licensing of replication origins." J Cell Sci **112 (Pt 12)**: 2011-8.

Rowley, R., E. N. Phillips and A. L. Schroeder (1999). "The effects of ionizing radiation on DNA synthesis in eukaryotic cells." Int J Radiat Biol **75**(3): 267-83.

Russell, P. and P. Nurse (1986). "cdc25+ functions as an inducer in the mitotic control of fission yeast." Cell **45**(1): 145-53.

Saha, P., J. Chen, K. C. Thome, S. J. Lawlis, Z. H. Hou, M. Hendricks, J. D. Parvin and A. Dutta (1998). "Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase." Mol Cell Biol **18**(5): 2758-67.

Santocanale, C. and J. F. Diffley (1996). "ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae." Embo J **15**(23): 6671-9.

Santocanale, C. and J. F. Diffley (1998). "A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication." Nature **395**(6702): 615-8.

Santocanale, C., K. Sharma and J. F. Diffley (1999). "Activation of dormant origins of DNA replication in budding yeast." Genes Dev **13**(18): 2360-4.

Sato, M., T. Gotow, Z. You, Y. Komamura-Kohno, Y. Uchiyama, N. Yabuta, H. Nojima and Y. Ishimi (2000). "Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex." J Mol Biol **300**(3): 421-31.

Savitsky, K., A. Bar-Shira, S. Gilad, G. Rotman, Y. Ziv, L. Vanagaite, D. A. Tagle, S. Smith, T. Uziel, S. Sfez, M. Ashkenazi, I. Pecker, M. Frydman, R. Harnik, S. R. Patanjali, A. Simmons, G. A. Clines, A. Sartiel, R. A. Gatti, L. Chessa, O. Sanal, M. F. Lavin, N. G. Jaspers, A. M. Taylor, C. F. Arlett, T. Miki, S. M. Weissman, M. Lovett, F. S. Collins and Y. Shiloh (1995). "A single ataxia telangiectasia gene with a product similar to PI-3 kinase." Science **268**(5218): 1749-53.

Schubeler, D., D. Scalzo, C. Kooperberg, B. van Steensel, J. Delrow and M. Groudine (2002). "Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing." Nat Genet **32**(3): 438-42.

Schwacha, A. and S. P. Bell (2001). "Interactions between two catalytically distinct MCM subgroups are essential for coordinated ATP hydrolysis and DNA replication." Mol Cell **8**(5): 1093-104.

Segurado, M., A. de Luis and F. Antequera (2003). "Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe." EMBO Rep **4**(11): 1048-53.

Segurado, M., M. Gomez and F. Antequera (2002). "Increased recombination intermediates and homologous integration hot spots at DNA replication origins." Mol Cell **10**(4): 907-16.

Shendure, J. and H. Ji (2008). "Next-generation DNA sequencing." Nat Biotechnol **26**(10): 1135-45.

Sheu, Y. J. and B. Stillman (2006). "Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression." Mol Cell **24**(1): 101-13.

Shinomiya, T. and S. Ina (1994). "Mapping an initiation region of DNA replication at a single-copy chromosomal locus in Drosophila melanogaster cells by two-dimensional gel methods and PCR-mediated nascent-strand analysis: multiple replication origins in a broad zone." Molecular and Cellular Biology **14**(11): 7394-7403.

Shirahige, K., Y. Hori, K. Shiraishi, M. Yamashita, K. Takahashi, C. Obuse, T. Tsurimoto and H. Yoshikawa (1998). "Regulation of DNA-replication origins during cell-cycle progression." Nature **395**(6702): 618-21.

Snaith, H. A., G. W. Brown and S. L. Forsburg (2000). "Schizosaccharomyces pombe Hsk1p is a potential cds1p target required for genome integrity." Mol Cell Biol **20**(21): 7922-32.

Sommariva, E., T. K. Pellny, N. Karahan, S. Kumar, J. A. Huberman and J. Z. Dalgaard (2005). "Schizosaccharomyces pombe Swi1, Swi3, and Hsk1 are components of a novel S-phase response pathway to alkylation damage." Mol Cell Biol **25**(7): 2770-84.

Sorensen, C. S., R. G. Syljuasen, J. Falck, T. Schroeder, L. Ronnstrand, K. K. Khanna, B. B. Zhou, J. Bartek and J. Lukas (2003). "Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A." Cancer Cell **3**(3): 247-58.

Stern, B. and P. Nurse (1996). "A quantitative model for the cdc2 control of S phase and mitosis in fission yeast." Trends Genet **12**(9): 345-50.

Stevenson, J. B. and D. E. Gottschling (1999). "Telomeric chromatin modulates replication timing near chromosome ends." Genes Dev **13**(2): 146-51.

Takahashi, T., E. Ohara, H. Nishitani and H. Masukata (2003). "Multiple ORC-binding sites are required for efficient MCM loading and origin firing in fission yeast." EMBO J **22**(4): 964-74.

Tanaka, S. and J. F. Diffley (2002). "Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation." Genes Dev **16**(20): 2639-49.

Tanaka, S., T. Umemori, K. Hirai, S. Muramatsu, Y. Kamimura and H. Araki (2007). "CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast." Nature **445**(7125): 328-32.

Tanaka, T., D. Knapp and K. Nasmyth (1997). "Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs." Cell **90**(4): 649-60.

Tatsumi, Y., T. Tsurimoto, K. Shirahige, H. Yoshikawa and C. Obuse (2000). "Association of human origin recognition complex 1 with chromatin DNA and nuclease-resistant nuclear structures." J Biol Chem **275**(8): 5904-10.

Tercero, J. A. and J. F. Diffley (2001). "Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint." Nature **412**(6846): 553-7.

Vas, A., W. Mok and J. Leatherwood (2001). "Control of DNA rereplication via Cdc2 phosphorylation sites in the origin recognition complex." Mol Cell Biol **21**(17): 5767-77.

Vogelauer, M., L. Rubbi, I. Lucas, B. J. Brewer and M. Grunstein (2002). "Histone acetylation regulates the time of replication origin firing." Mol Cell **10**(5): 1223-33.

Walter, J. and J. Newport (2000). "Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha." Mol Cell **5**(4): 617-27.

Walter, J. and J. W. Newport (1997). "Regulation of replicon size in Xenopus egg extracts." Science **275**(5302): 993-5.

Willis, N. and N. Rhind (2009). "Mus81, Rhp51(Rad51), and Rqh1 form an epistatic pathway required for the S-phase DNA damage checkpoint." Molecular Biology of the Cell **20**(3): 819-833.

Wohlschlegel, J. A., S. K. Dhar, T. A. Prokhorova, A. Dutta and J. C. Walter (2002). "Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45." Mol Cell **9**(2): 233-40.

Wolf, D. A., F. McKeon and P. K. Jackson (1999). "Budding yeast Cdc6p induces re-replication in fission yeast by inhibition of SCF(Pop)-mediated proteolysis." Mol Gen Genet **262**(3): 473-80.

Woodfine, K., H. Fiegler, D. M. Beare, J. E. Collins, O. T. McCann, B. D. Young, S. Debernardi, R. Mott, I. Dunham and N. P. Carter (2004). "Replication timing of the human genome." Human Molecular Genetics **13**(2): 191-202.

Wu, J. R. and D. M. Gilbert (1995). "Rapid DNA preparation for 2D gel analysis of replication intermediates." Nucleic Acids Res **23**(19): 3997-8.

Wu, P.-Y. J. and P. Nurse (2009). "Establishing the program of origin firing during S phase in fission Yeast." Cell **136**(5): 852-864.

Wuarin, J., V. Buck, P. Nurse and J. B. Millar (2002). "Stable association of mitotic cyclin B/Cdc2 to replication origins prevents endoreduplication." Cell **111**(3): 419-31.

Wyrick, J. J., J. G. Aparicio, T. Chen, J. D. Barnett, E. G. Jennings, R. A. Young, S. P. Bell and O. M. Aparicio (2001). "Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins." Science **294**(5550): 2357-60.

Xiao, Z., Z. Chen, A. H. Gunasekera, T. J. Sowin, S. H. Rosenberg, S. Fesik and H. Zhang (2003). "Chk1 mediates S and G2 arrests through Cdc25A degradation in response to DNA-damaging agents." J Biol Chem **278**(24): 21767-73.

Yabuki, N., H. Terashima and K. Kitada (2002). "Mapping of early firing origins on a replication profile of budding yeast." Genes Cells **7**(8): 781-9.

You, Z., Y. Ishimi, H. Masai and F. Hanaoka (2002). "Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex." J Biol Chem **277**(45): 42471-9.

Zappulla, D. C., R. Sternglanz and J. Leatherwood (2002). "Control of replication timing by a transcriptional silencer." Curr Biol **12**(11): 869-75.

Zegerman, P. and J. F. Diffley (2007). "Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast." Nature **445**(7125): 281-5.

Zhou, B. B. and S. J. Elledge (2000). "The DNA damage response: putting checkpoints in perspective." Nature **408**(6811): 433-9.

Zou, L. and B. Stillman (1998). "Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin." Science **280**(5363): 593-6.

Zou, L. and B. Stillman (2000). "Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase." Mol Cell Biol **20**(9): 3086-96.

Appendix Table List of oligonucleotides for *AT2062* microarray

Sequence Name	Bases	Sequence
SPB1801420.63	63	GTA AAT GTT TTA ATG ATA CGG TGA GTG ACG GAA AAT TTG AAT ACT GGT CAG CTA CGC AGT TTA
SPB1803820.59	59	TAA ATC ATA CCT TAT TGG CAA TTT ACA ACG AGA CTG TAG AAC TCC ACA ACA GGT TTG GA
SPB1804370.61	61	ATT AAC TAA AAT CCT TAC CCG ATT AGG TGG GTG TTG TTC ATC CAT TAT CTG TGC TAC ATT T
SPB1804970.60	60	TAG CAA GTC CAA CCT ATA ATA AGG AAA ATA AGC TCA CTC TGG AAG TTC TAC TGG TTT GGC
SPB1806370.58	58	GAA AAT TTG TGG AAA CGA GAT GTG CAG TTA AAA CGT GCA ATT TCC GAA GGA ACT ATT C
SPB1806920.63	63	TTT ACC TTA TGT TTC CCT TGT GTT GAC TAA CAT ATT ATT TGT AGC CTG TTT TTG CAG TTT GTC
SPB1807420.63	63	GAA GCT GAG TAT GTG GAA CTT GCC AAT TCT TTA AAA ACA AAA GTA GAG ACC AAT ATA GAG ATT
SPB1808020.58	58	TTA ATT GCA AAA CTA CGC ATA ACT TGT CTA TCG TTT AAT TGT ATC TTT TCA CCC TAG T
SPB1808670.63	63	GTA GAT AGA AAC ATC TGA TTA TCA ATG AAC TTC CAT CCC GAT TGT AAA TCT TGT GTA CTC AAA
SPB1809220.59	59	CCT CTT TTG AAC CCG CAC AAG CAG AAC AAT GAG AGA TTG CAG GAT GAA GAA CGG TGA TC
SPB1809670.63	63	TGA ACT TAA AAA ATG TGA AAA GGA AAA ACC CAC CAG GAA ATT AAA GAG AAG AAA GAA TTA ACG
SPB1810120.58	58	GGC TAG AGA AAA TGA GAG GTT CGA TGG GCA GTT TAT ATT CGC AAG ATT CGC AAT AGA T
SPB1822220.62	62	AAT CAT GGT CAA ACT TTT CGA GCG TAA TAT TTA GAA TCT ACG TAC CAC GAA CAA AGA GCA TC

SPB1822820.63	63	TGA GAC GCT ATA GAA TTT TAC GTA TTC GTG TGT GTC CGA ATC TTT TGT AAT AGT AAA CAG CAG
SPB1823320.63	63	ATT ACT ATG AAA GAC CTC CTC GAT AGT AGA TGC TGT AGA GGT ATT AAA CAT GGA GTT GAA CTT
SPB1823770.57	57	AAA CTT TGT TAT GTC TAT TGA TGA TTC CGT GCA GCC TCC TTT ACT TTC TTG GAT GGG
SPB1824320.63	63	ATG ACT AAG AAA ACC GGC ATT TGT AAT AGA AGT GTA AGC GAA CCA TAG TAG TAC ATT AAC CTT
SPB1824920.59	59	TTT TGG GAA ACT ATG GGT CGT GAA TGG ACG GGA ATT GAT ATT TTA CGA ACA GAC AAA TT
SPB1825470.63	63	TTC TGA ATA TAC AAA GGA TTC ACA ATA ATT AAT TGA TTC GGC ACG TTT ATT AGG TTT TGG TGA
SPB1826120.63	63	TGA GTA AAA CAG TCA GAT TGC AGT AAG ACC TCG ATA TCA TAA TTA ACC TCT AGA AAT ACA CCA
SPB1827120.57	57	ACT TAG ATG CAG ATA CTA CCA CTT TCA ACC TTT ACC AAC CAC CAT GTC TCA ATT TCC
SPB1827620.63	63	GAG ATA AAG TCA TTG CCA AGT CGA GAA AAG AAC AAG AAA AAC TGT TGG AAG AAA ATG AAA GCA
SPB1828220.63	63	CGT TGT AGC AAA TAC TGA ACG TCC GTA CTG TAG TTA TTT CAT TTA ATG TGA TTT GTA TCT AGC
SPB1828720.59	59	AGA CAA TCG TTC AAA AGA CCC TTG CTC CAA AAC TTT CCC GTC TCC AAC AAC AAT TAT TT
SPB1835770.63	63	AAT GTC AAT GGC AAA TCA TAT ACT GTT TTG AGT TAA AGA GGC TTT GGC AGA TCA TAA ATC AGC
SPB1836370.63	63	CAC CTC CTC CCG AAG AGG CAT CCC TTT ATG CTA GTA AAT GTG ACT ATT ATG AAA GCG ATA TTT
SPB1836820.62	62	ATC GAT TAA ATG AAC AAC AAA GCT TCT ACC AAA GGC ATT TAG AGT CTA CTA GGC ATG AAT TT

SPB1837370.61	61	AGG AGG AGA AAC GAG CCC AGA AAA TTC AAA AGT CAA AAT CTA CTC ATC TTC TAT AGC GAA T
SPB1837820.58	58	TAA ATT TGG ACA ATC CCA ATG GCA CCA GAT ACT GCT TGA AGC CCC AAC CAA GGA TCA A
SPB1838270.63	63	AGC TGA AGG ATT AGG ATA CGA CAT CCC GAA GTA TTT GTT TGC TAA TTG AAT TTT AAG TCA GAA
SPB1838720.59	59	ACT AAC CGT TAT TAA TCA ACT TGC TAT ACG CTG TGA CAT CTA ACA CTC TTT GAC CAG TT
SPB1839370.57	57	CCA GTG TAG CCT TCT ATA TAA ACA AAG GTC CTA TTC TTT GAC CTC CAC CAC AAA CGG
SPB1839970.60	60	ATA TAA AAT GCT CCA AAG ATT ACC CTT GTT TGT GCT TCT TGT TGT TTT GAT TTA CTA ATT
SPB1840470.57	57	GCT TTC AAA ACC CGC ATT TTA TAG CGT CTG ATT CTC CAT CCT CCG TAT TAT CTG GTC
SPB1840970.57	57	ACA ACA ACA TCA CCA TCT AAA TAT GGA TAT GTG TCT GGC AGA ATA GCT CTA TCT CCT
SPB1841570.58	58	CGG AAA TCT TTT CAA CCT GAA CTG TCT ACC TTA CTT TTG AGG CGA GAC GCT TTG TCT A
SPB1848620.63	63	TCG TAG AAA GAA CAT CAG CAC CAG GAG CTC TTG CAA CTA ATT TTT GAT ATA AAG AAT CTC TGT
SPB1849320.62	62	AAG TAA ACA TGG CAT TAG TCT TTT CCT CAT CCT TTC AGA GCC TTC AGT AGA ATC TAA TTC CC
SPB1849820.58	58	TTT TTG CTA TAT CGC TTA GAT AAC CAG GAG AAT GAA GAA ACA GAA GTC GCC AAG CAT C
SPB1850470.61	61	AGT CCT TTA ATA ACA TCT TTA GCT TCA GCG TAT TGA ACA GTA AAA TTG TCG AAA TAT TGC A
SPB1851070.63	63	GTC AAG TCA GTT CTA AAT TGA ACA CTT CTG ATG CCT AAC CCT AAG GAT AAA AAC AAC AGA TTT
SPB1851520.57	57	TAT GTA GCA ACG GAG CGA ATA TTT TCA GAA ACA AAA GCA GAG GTC ATC AAA ACT TTG
SPB1852070.63	63	TAT TGT AAA TGG CAT CAC TAA TAC TTT GCA TTA

		TGA TTG GAG TAA CAG ATC CAG AAG ATA CAC
SPB1852570.61	61	GTT TTA TAA GAC TTA GAT GTC TTT TGA GAT GCT GCT CCA GTG TTC TCT GAA ACA AGT GCA T
SPB1853070.60	60	GTT TAA CCA AAG GGA AAC AGA ATA TCC ATT AAT TGG CAT TAC AGG AGA GCT CAG CGT AAA
SPB1853570.63	63	ATA ACT CAA CAA GAG ATG CAT TGG TAT TAT CTT GCA AAG AAG GTA CAA GCT TCA AAA GCG AAA
SPB1854020.63	63	AGC AGC ATT CAA AAA TAT CAA GAC TTA GTA ACG ATT GAA AGT TTA GCT TAG TGA CAG CCT TCT
SPB1854470.60	60	GAT GCT TGC TTA AAT TTA TCG TTA TTG AAA CTG GCA ATA ATA GTT CGT AGT GGT TGG CGG
SPB1862420.61	61	AGC TCA GAG CTA AAT TAG AAG GAA GAG GAG GTG AAA ATA AAT GAT TAC GAA CGG TCT TGT G
SPB1862920.57	57	AAT GCT TCC TTT AAC CCT TCC GAC ATA GAT TCC CAA AAT TCC GGT AGC ACA TTA GTC
SPB1863420.57	57	CGC TTT CGG TTG TCA ACG GTG CGA TTA ACT CGA TCT AAT GTT ACT AAG CCA AAC CCG
SPB1863970.60	60	GAG GAT AGA CGA GGT TTC CAA CAC CGC ATC CGA CTT CAA GGA TAG TCA ATG GCA ACT TAT
SPB1864570.61	61	AGT GCA TTG GAG CAG AGA ATG GGT TCC TGT TGT TGT AGA TAC ATA TTC TAA TGA AGA TGA T
SPB1865120.63	63	CTT TGT AAA ATA TCA ATA TCT TCT ATG ATG ATC GTG GAG TTT TCA AAT TCA AAA TCG TTT TCC
SPB1865520.59	59	AAT GTG ACA GGT GTT TCT CTT GTT CCC AGT AGC AAA GTT AAA GAA AAA GGC TCA GGA AT
SPB1866070.63	63	AAT AAT TCC AGA GAT TGG TGA GAA CAA GGA ATA GTT TTA TAA ACA TTT GTT TCT GTA CAT TGT
SPB1866570.58	58	GGT AAC AGA AAA GGA TGC CAG GAA GTA TTG CTC CAA GTG CCC CAT AAA TGC CAA GGA A
SPB1867220.63	63	CGA CCA CAA CAT AAA TAG ATA TTG TTT CCT GGT

		AAA TTC TTA TAT CGG CTT TCA ATA TGC GGT
SPB1867920.60	60	TAC AAC AAT TCA GGG CTC AAA GTT AGA AAA ATC AAG TTT AAT AAA TTA TAC TAT CCG TGC
SPB1868470.63	63	ATT CAC TGT ATT TAG TGG AAA TGT AGT TCA CAA GAT TAT CGA ATT AGA TTG TTA TTC CTC AAT
SPB1875820.57	57	CAT ACA AAA CGT AGC CGG GAG AGG AGG GGA GAG AAG TAG AAA AAG ATT GAA AAT TGT
SPB1876370.57	57	CTG GGT CAC TAG CGC TTT GAA CCA TTA AAT CTG AAG GCA CAC GAT ACG GAG TTT GAT
SPB1876820.60	60	AGT AAA ATA TCG GTA TGA AAG TTC TGC ATT CCT TGT TGC GAT ATA TAC CAT TAT ACC ACT
SPB1877420.63	63	AAC CAT ACG CTA ACG ATT AAT TAG GAA TTG ATA CAG GTT CAT AAC ACC GTT CAT GGA TTG GAA
SPB1877920.60	60	GAT TGT TTC ACT ATC CAC TAG TTT TGG ATT ACA TGC CAT TAG TGA TCA GGG CTA TAA GCG
SPB1878520.59	59	CGT GCA AGA CAA GTG AAG AAT AGT AAC ATC TGG CTA GTG AGA AAT AGC TGA GAA CGA AT
SPB1879020.57	57	CAC CAA ACG TTC TTG AGT TTT ACT GGT CTT GCA CAA CTT GGA ACA AAC AAA CAA CCT
SPB1879470.57	57	CTC TCT CTT TTC CAA GGC ATT GGG TTC AAC TCG AAG TGT CGT TGT TGT GTA ATA ATG
SPB1879970.63	63	CAT GCA TTC CGA TTT GTT CCT ACT TAC ATC CTT TGG AGT ATA CAT ATT CTC CTT AGC ACC TTC
SPB1880520.58	58	TTT ATC TAT CTG GTT CGT TTG CAT TTA CTT TCT CAG TCG TCG GTT GGT TAC TTG CAT T
SPB1881020.58	58	TCC ACC TTT GCT GGA CTA TGC TGA TTC TTA TAC TGC TTT GGA CTC TTC TCT CAT TAA T
SPB1881620.57	57	TAT TAG TAA CAT TGT TAT TGG CAT GGC TCA TCG AGG TCG TCT GAA TCT TCT TCA CAA
SPB1889170.58	58	ATG ATA AGA CAA GTT ATG TGG CTA CCT TCG TAA CTG CCG TAG TGT GTC ATC AAC GGT A

SPB1889720.63	63	TCT GAT AAC CAT AAT GGC TTA ACG AAT TTC ATA TGC TAA TCA CTG ACC GTT ATG AAT ACA ATT
SPB1890270.59	59	TAA AAG TAC AGG ATC GCC AAA ACA AAA GGG AAC ATA AGA GAC AAT CGA AGT ATG ACT CC
SPB1890820.58	58	TTG CAC CAA TTC AAA AGT GTT CCA ATC ATA AAT TTC GAC GCG ATC ATC AAC TGC AAT T
SPB1891320.62	62	CGA TTT CTC CAT GTT CTA TAA ATT TTG TAT GAC CTT CAC CAA CTT CAG GTT CAT AGT CTT CA
SPB1891870.57	57	CGA GAC TTA AGC AAA TAC TGT AAG AAT ATT AGA TCC CCG TAT CAG AAT CTA TCG AAT
SPB1892420.63	63	ACG AAA TTG CCG CGA GAC TAC ATT GTC CAT GAA ATT AAT GTT TGT GTT TAT TCT ACT AAA GGG
SPB1892970.57	57	CTC CAC ATC CAG GGT ATC TGA TGT CCT CGA CAA AGT CTC AGC GAA CGA GCG ATT TGG
SPB1893570.63	63	ATA GTC CAA AAA TTA TTT CGT ATG TTT AAG TCG CCG AGT CCA TTA AGC TAA CAT TGT TTT TCT
SPB1894170.60	60	CGA AAA GCG GAA GAA GAA TCA ACT AAT AAC AAC AAA TGG GCT AAA AAG TTC TCA ACC AGC
SPB1894770.63	63	CCA ATA AAT TCT TTA ATG TTG GGG GTA AAG CGA AAA GGA ACA ATT TCT CCA TTA TGG TAA ACT
SPB1895320.62	62	CAG GAA TGT AAA ATT TCA AAC AAC GTC TTC GAG TCT CAC AAT TTA AAG CAA GTG CAT CAT TT
SPB1902670.57	57	ATT TGG AGT GCT CTA AGC TAA AAT TCT GAG AGG AGC CTT GGA AAG AAA ACT GGA TTG
SPB1903120.62	62	TCA ATC TGT TGA AAC ATT CTA GCA TTA CTT GGA GAA GTG GTA GTA TTT CCT TGT AAA GAG AA
SPB1903520.57	57	ACA CTC AAA TAT ATT CGG AAA CTC GGA TTC CAA AAT TTC GAC GAA AAC AGA TGG GTC
SPB1904020.63	63	CAG CAT TTT CCT GTA AAG AAT CTC TCA GCT TCA TAA ACT CGA AAT TCA GAG AAG TCA ATT TTC
SPB1904520.62	62	TCA AAC TGT TTA AGA GCA GCA CCA AAA CTT

		CTC AAT AAA TAG GCG AGA AAA GAG AAT GAT TT
SPB1905070.57	57	ACA CTG CCA TCT CTT CAT TAT CAT TAA CGA CTA TAC GAA GAA GCA CGG CAA AAG TCC
SPB1905570.63	63	TTC GCT ATT TTA ATT AAC GGC CTT ATT ATT TTG TGC GGT AGA AGT TGC TAA AAT GAT GAA CGA
SPB1906170.62	62	TCA TTT AAG AGA AGT GCT TTT ATT CAT TAG TAA TTG ATG AAT CCT TCG ATT TAT TGA TGA GT
SPB1906720.63	63	GGC TCG TAA AAC CTT GGG ATT ATA ACC ATG AAT TGT ATA TTT GAT TAG ATT GCC ATG AAT AAC
SPB1907220.63	63	CAA ATA AAA CGA TCA TAT TCC ATC TTC CAA TTT TTC TTC ATC ACC AGT GTT TTC TTC ACT ACC
SPB1907870.63	63	GCA GGT TAA TTT TTG ATT TAT AAG AAG GAG AAA AAG AAG TTA CAG GAG TGT TTA TTT CCG AAA
SPB1908470.57	57	CAG CAA AGT TGA TCT GGA AGA CCC AGT CTT TCA TGG AGA AAT AAT TGT TCC CGG AAA
SPB1810820.57	57	TGG TGC TAA ATC TGA TCT GAT CTC ACT GTC ATT GTT TTG TTT CCT CCC ATA AGC TCA
SPB1811320.58	58	TAA GGC ACA ATC CTA AGG CAG CAA CGT AGT AAG GAG ATT TGG AAT TGT AAC TTA TTG C
SPB1811920.63	63	AAT TAT TCC CGT AAG TTT CCC GCT TTA GAA AAC AAA GAA TCA TGG TAA CAA AGA ACG TAT GGA
SPB1817420.63	63	ATC AAA CGA TGC GTT TTC GAT CCT ACC ATT AAC CAA CTA TAT TGC CAC AAA TTC TCA TTG ATT
SPB1818170.63	63	TTC ATA CTT AAC TAC ATA TCC TAA CCT ATC CTA CCC TAT TCT TTT ATT TTC TCT CTT CTC TTC
SPB1818620.57	57	CTT ATC AAG CAA TGT CTT ATT CAG TTA GTC CTA CCT CTA CCA ATT CTT CCG TAG CTA
SPB1819170.58	58	CAA ATC TTT TCG CCA TGG AGT TAC TTA AAG TAA TTT CAC TAT TGT TTT CCC CTT AGA T
SPB1819670.57	57	CCC CCT TTT AAG TTT TAA TAT CTT TCT CTT AGT CTC ACT ATG TCT TTC GAC CCG CCG

SPB1820270.60	60	GGT ACG ACG ACC TCC TTC CTC TTT TGT CAA AAT ATA GAA TGA TGC CTT GAA TTT CTG ATG
SPB1820720.57	57	GCC GCT AAC AAT TGG AGT ATT GTC ACC ATC AAA TCC GTA TTC GGA GAG TAG CTC ACG
SPB1821120.58	58	AGA TCA GAA AGG CAT TTA GTA ATA GCA GCC GTC AAC GTC GTT TTA CCG TGG TCA ACA T
SPB1821720.57	57	TCT GAT CTC TAA AGG AAT GTC AAT CTG ATT TTG CTA GAT CGC AAG AGC AAC CAC CAT
SPB1829320.59	59	AAT GTC GAA AGA TAA GTT ATC AAA TAT TGA AGC TGA AGG ACG AGT TGG GTA TGC AAA CC
SPB1829820.60	60	TTA AAT CAC CGA GAG CAT CTT GTG TAG TCC TGG AAA GTT TCC GAA CAT ACT CTC CAT AAA
SPB1830370.57	57	ATT CCA TCC CTG ACC CCT TGC TAG TGT GAA GAG ACG AAA AAC ATT GAC TTT CTT GTT
SPB1831020.62	62	TTC CAC GAA GTA CTC TAC ATT CGC ATA AGG TAG GTG CTT TGG TAC ATA CTA AGA TTT CAT TT
SPB1831620.63	63	ACA TTT AGT GAC ATT GAA GGT TTG TTA CGG GTT GTA TTC ATA TTA AAT TAA CCA CAT GCA GTT
SPB1832170.58	58	AGA AAA AGG AGC CTG AAA AAC CTA AAA AGA CTA AAA CCA AAA AGG AGA CCA TCC AGG T
SPB1832720.57	57	AGT AGA ACA AAT CCA GCT CTG TGC AGT TGA TCC TAA TAC GAA ACT TTC CAT CAC CAT
SPB1833170.58	58	AAC AAT TTG TAC ATT TGC CAT CAA AGC AAC GTG GTC GCT GGG ATA TCT TCC TTC TAA A
SPB1833670.62	62	CAA ATA CGT AAT TCA TAC TAC GTT CTA TAA TGT CAG TAG CCC TTT GGC AGA TGG ATA GAG GG
SPB1834270.58	58	GAC GAA TAT TTG TCT GAG CAA GTA CGT TGT AAG TCA TAA TAG TAA TAT CCA ATG CAG A
SPB1834870.57	57	CGT ACA ACC TCA GGT ATA TCA TCT AAA ATA CCG CAT ACT AGC GTG CTC TCC TTG ATG
SPB1835320.57	57	GGA GAC TTA TGG TGC TTT TAT ACG GTT GGT

		GTT ATT GGA GAA GTC GGT GGT AAG TGT
SPB1842220.59	59	AAA CAG GCT AAC ACT CTT GAG AGT CCC GAC GCT TAT CGT TTC TTG TAT AAA TTT TAT CG
SPB1842620.57	57	AAT CGA ATG ACA TCC AAG AAC CTT GCA GTA GTA TTT TCC CCA ACA TTA ATA CGC GAC
SPB1843170.62	62	TTT TTC CCA AGA TCC CTA ACT TAG GAA ATG CTA GTG CAA CGT ATT GTG AGT AAG TGA TTT TG
SPB1843670.60	60	TCA AAA GCC CAA ACA TAT AGA ATG CCT ACT TGT GGA CTC TAA AAT CAC CAA AAC AAG GAG
SPB1844270.63	63	TAA TAT TGG TCC TCA AGT TTA TTT CGA TCC TAG GGA AAG ATT GCT TCA TGT TAA GGA GTT ACT
SPB1844770.58	58	CCC CCA TGA TGT TTC ATT CTT AGT CAC TCC TTT GTT TAT CTT CGA GGA TTC ACT CAT C
SPB1845520.60	60	GTT CAG GAG TTA TAT ATC TTT CTC GTA TCC CTC CTT ACA TGG CAC CAA ACA AGC TTA GAC
SPB1846020.57	57	GAA ACA AAA GAC TTA TCA GAA GAT ACT AAG CCA AGC AAG CAA GCC CGC CGC TTC TTT
SPB1846670.63	63	GTC CTT TGT CAT CTC CTG CAA AAA TAC CTC TAG CTT GGC CGT TTG GAG TAA TGG TAA TAA TAT
SPB1847220.58	58	AAC ACA TGT AAT TCT ATA ACG GTG ACC ATA AAG AAC CCG CTT CAA GCT AGC CTT AAT G
SPB1847720.57	57	CTA CTG CGG CAG CAA GCT CAA ATT CAT TTT CAA TAT TTT CAA GAT CTA TAG CAC CTT
SPB1848170.62	62	CAA AAG GTC GTA AAC GAA TTA AGT AAG AAC ATA CAA TCA TTG CAG AGG AGT AAT GAG TAC CA
SPB1855070.57	57	CCC CAA GGC TCA TCA CAC ATA AAA AAG TAA CCA ATT TGC TTA TGG CGT CGT AAA AAA
SPB1855720.57	57	GCG GCG GAT TAA TAG CTT GAG GAG TTG CAG ACA GTT GTT GCT GTA ATT GCT GCT GAT
SPB1856670.57	57	TAG CAC GAG TTG ATT CTA GCT TGA TAG CGA CCT CTT CAC CAG AAA CGA CAT TAG TCC

SPB1857320.63	63	GAT CTA AAA AGA TAG AGA AAA TTC GGG TTC TCA GAC GAA AAT AAA AGC TGT AAG ATA CAC TGT
SPB1857870.63	63	TCG TAA TTT TGA AAG ACG AAA ATA GAT CAG CCG TTG GAG ATA GCG ATA TTT GAG AAA TGG TTT
SPB1858420.59	59	GCC CAA TAG ACG CAA TTG TTC AGT AAC GAT TTC CAT TGA AAA CAT ACC TTC TCC GAT TT
SPB1859020.62	62	ATG TTC TAA TTG CAA CTG TAA ATA CTA GGC ACA TAC TAT TAA GGC TGA CAA CCA TTA GGA GT
SPB1859670.62	62	AAC TAT TAA TAT GCT GGC ACT ATG GCA ACA TTA GCA CAT CCC AGG CTT CTA TAT ATC TCT TC
SPB1860170.63	63	AAC AAG AAA TTT AAC AGT TTA GTG TCA CCG CCA ATA ACC CAT ATA AAC AAT GAT ACC AGC AAC
SPB1860720.63	63	ACA ATT CTC ACT TGG GCA GAT ACC TTT TGA TGC ACT ATT TCA TCT TCA TTG TAT ATA TCA ACA
SPB1861170.63	63	ATG TCG AGA GCG AAA TTG TCC TTT TTA ACC GTT GCG ATC AAA ATA TAC CAC AAA AAT CAA ATG
SPB1861770.63	63	TAT AGT TCA TGT ACA GTT GTT TCT TTT AAA TGC ATC ATA TAT AGG TCG GAA AAT TAG GAC AGA
SPB1868920.58	58	ACA AGG GCG GTT CGA TCC CGA TTG CAA ACG CTG GAA AAT CTA GAA ATA TTG GAA AGT T
SPB1869370.58	58	GCT GAT CAG TTT GCT GTT AGA TAA CTT GGT GTG TCA TGA CTT ATA CCT TTG GAG TTG C
SPB1869920.57	57	TTT TCA CTT TAC AAA TAT CTC ACT CTC TAC TTT ATG GTA GGC TCT CAT TTT TGG AAT
SPB1870470.60	60	TTG AAG ATG CTA AAG ACG CTT GTA ACC GCA TTT CTG CTA CTG AGA ATG TAG ACT GTA ATT
SPB1870970.59	59	GAT GGA CAA CGA AAA GCT TCC ATT TTG CGT GAA AAG AGG GAA GAG TTG GGT TTA AAT AA

SPB1871620.63	63	CTC TTA AAT CTG CTT CGT TAT GCT GAT AGA CTT TGA AAT TGA TTC GTG CAA TGA CTA CAC CTT
SPB1872170.57	57	GAC TGG GAA GGA CGT TCG TTC TAC TAT TCA GGT TCA AGT TGG GTA AGC GCA TAG ATA
SPB1872720.57	57	TTT AGA CCT GGA CCA AGC TAC ACA ACA CAC GCT AGG AAG AGA ATT GAA GAG GTT CGC
SPB1873370.63	63	TAG AGA AGG ATT TTA TTT TGT CCG CTA AGT AAG TGA AAT AAA TAA ATA TCT CGG CCA CCT CGG
SPB1873970.62	62	GGC AAT TGA AGA TTA CGG GAT GCT ATT TTC CTT TTC ATG CAA TTA GCA AGT ATT ACT GAT TT
SPB1874620.57	57	TAT ACT GGC GTA TCG CAT GTG TGC TGT CTA TCA CCA TGT TGG TTA TAA GGA TGT CTC
SPB1875220.58	58	GCC CTA TTT CAG AGT GAT TGC TTG GGA TCC AAC TAT TGC CCT CAC TGG AAG TAA GTT T
SPB1882320.62	62	CCT ACG TTT AAA ATA TAC ACA CAA CAA CTT TTA CAA GAA AAG ACG GTT TCA AAA GCA GAA GT
SPB1882870.57	57	CTC TGT CTG AAT ATG GTG TTT TAG GTT TTG AGT ATG GTT ACT CTT TGT CAT CAC CTA
SPB1883520.59	59	TGG AAA CAG ATG GCA GCC AAT ATT TCG CAG TAT CCT AAT TTG AAG GAA ATT ATT TGG TG
SPB1884120.63	63	ATA ACG GTT ATA CGA TAT AAG AAA ATG AAT TAA AAC ATA ACA GAA CAG AAC TTT ACG GCG ACA
SPB1884670.57	57	CCT CGT CAA ATG GCC TAA CAC GAA CCG AAC TAG GCA TGA CTA TTC GAG AGT TAC CAG
SPB1885220.57	57	TGG GTT CAT GCC CTT TTG ATT TAA AGG AAC GTA AAT ATC ACG AAC GAC ATC AAT GCT
SPB1885770.61	61	AAA GAT TTA GCA ATA GAA ATA TGT TCG AGC AAA AGA ACG ATG ACA GAG ACA GGA AGT TCG C
SPB1886270.57	57	GAA CTC AAT GAT AAA ACC AAG TCG TAA AAG ACC TAA ACC GCA AGT AAT GGC ACC TGC
SPB1886870.57	57	TGG TTC GGG ATA AAC AGG GAC GAA ATT AGA

		TGG TGG TGG AAC AAG ACC AGG CAC ACT
SPB1887620.57	57	AAA GCG GTG TAA GAA CTT TTA CGA CTC CGT TCA GCT ATG AAA CAG ACA GAG TGG TTG
SPB1888120.60	60	GCG GAC GGA CGT TAT AGA GAA GAA GAA GAG AAT TGT GAG AGA AAT GCA AGG AAA AGA AAA
SPB1888670.58	58	CAA AGG TAC ACA TAA ACA CCT GTT CAT TCA CAT TAT GTC TCT TGA AAA CCT CCC TCC C
SPB1895820.58	58	AAG TTC GAT TCG GTT GAA GAA GGA AGT TTC ATT TCA TCC GTT ACG ATC CCC AAA CGA A
SPB1896420.57	57	AAT AAC CAC AAT ACC CTT GCC AGC ATC TTA CGA GCT TTA GAA CTT TGG TAA GTA CCA
SPB1897020.63	63	ATC ATA GAG AAT ACT ATT TGT CTC CAA GAA ATG AGA TCA CGC CAA AGA TTA ATA TCG TCC CAG
SPB1897570.57	57	ACA AGT TGT ATT CTG ATT CTG AGA AGG GAA TGG CTT CAC TGC GAG CCT TGG TCT GTG
SPB1898170.61	61	TTT TTA GGA ACA AAT CCA GTT GTT TGA AAA AGG TAA ATC TGC CAA GAG GAA TAG CTG AAC C
SPB1898770.63	63	GCA CCA CAA CAC TTA AAA ACC ATC TAC GTT GAT CTC CTA AGC TTG CCA TGC GTA CTC TAA TAA
SPB1899420.58	58	AAT AAT GAT AAC CCC TTC TTT GAA AGT GGA TTT TCT TCA ACT GGT TCG GGA AAA AGC G
SPB1899970.57	57	CAA TCG TAG CGT AAG CGG CTT GTT TTA CCA TAG GAT CTT CGA GTT TAA TGT AGT TCC
SPB1900470.63	63	TTT AAA CTA GAT AAA AGC ATT TTT GTA ATG GAC GAT ATA CAC CTG CGA ATA GGA AGA AAT AGA
SPB1901020.63	63	TCA ATA ACG AAA GTG AGG CAG TCT CTA GGA CTT TAG AAT CAG ATA GAA CAT TAA GGT GTT GAA
SPB1901670.62	62	AGA AGA TCA ACC ACA CTA AGT TTA GAA TTG TTG GAA ATG TTC TCA ATA GCT TGT TGC ACA GT
SPB1902170.60	60	AGA CCT GGT ATT TTC ATT TTC AGA ACA TAA AGA

		TTC GAA AAT TCC GTT GAT GCA CGT AGT
SPB1908870.62	62	CTT AAT AAC ACA ACT GGG GAG ATT GAA GTC ACA AAA GAA TTA AAC GAA GAG CAA TTG GAT GC
SPB1909420.57	57	CAG GTA ATG AAG ATT ATA CGA GCA GAG AAG CTC TTG AAT CCT GGT GGT CGG AAA ATC
SPB1910020.61	61	TAC TCA ATA TAA GAA GAA TTT CGT ACT ATA TGT AAG AAA CAG AAA TAG TGA ATT ACG CTG T
SPB1910620.61	61	AGG AAG AGT CGA TTC GAG ATA ACT TTG TTC TTG TTT ATG AAT TGT TAG ACG AGA TCA TGG A
SPB1911120.63	63	TCC TAT CGC ATG AGT TCA AAT GTA AGT TAT TAG TTT GTG CCT GTT CTT TTT ATT CCC TCT CTC
SPB1911720.63	63	TCT ACC TTT AGA TAC CTT TCC CTT CAT CTA TAT ATT TGT TTT CTT AGT GTA CTG ATC CCA TTT
SPB1912320.57	57	TAC ATA TTA TTT GCT CCA CTC AGC GAA AAC ATT AAT TCG TCT TGC CAT TTC AGC TGC
SPB1912820.58	58	ACC ACC AAC TAC AAG TAC TGA TTC ACC CAC GAA TAA TTC TGG TTC ACG GAA TAG CGA T
SPB1913570.63	63	ATA TCC AAT TGT GAA GGA GTG TAA TTG AAT GAA AAC TAC ACT TTA TGT GCG TTT TAT ATC TTT
SPB1914070.60	60	AGC GTT TAA GAA TTA GTT ACT TAT AAA GAC CGA AGC GAT CTT CCA GAT AAT GAA TAG CAA
SPB1914720.63	63	TGG TTA AAA TTG GCA GCA CAA TAT AGT TTT TGA TGC AAT TAC AGT TAA CAT GCA TAA TAC TTG
SPB1915470.62	62	AAA AGG AAG CAA TGT AAT GAA CAA AGC GAT CTA GAC TAG AAT ATG TCA AGG AAA CAT ACT TT
SPB1801770.60	60	TGG GAA ATC GCA TTA TTT AAA ATA GGT TTT CGC TCT ACA CCA AAT TGA TTT TTC ATT TTT
SPB1804120.63	63	GTT TTA AGT CTG CAC GTA CTC ATT GCC GAT ATT TAT TTC AAC TCA ATA GGT ATC ATT ACG GTT
SPB1804620.57	57	CTA ACT CGT AAA GCA AAT GTT CAT TCT GCC TCG TTT GAC ATA TCA GAC GGA AGT CAA

SPB1805520.63	63	AAA TTC ATT CAA CCA TTA CGG AAG ACA CTA AAT TTC CTT AAA GAA TGA GTT CTG AGC TAA AGG
SPB1806620.60	60	TCA AAT CCT TGA AAC AAA GCC CTC AAT TGA TAG TTT ATA CCT TGT ATC TCT TTT ACC CAT
SPB1807170.57	57	TGA AAT GCA CAT TGC TAA ACA TGA AAT TCG AGC TCA TTT TCA GGA GGA GGA GTT TCA
SPB1807770.58	58	TAT CTT TCC CCT TTA TCC AGC AAC TCG ACC AAT TGA GCT TTT TGA ACA TGC CAT TTA T
SPB1808420.58	58	AGC GGG TCC TAC ACT TTC AAA AAC TGG CGG GCA TTC TTC ATC CCA ACT AAT ACC TAA A
SPB1808970.58	58	TGG ATG GTA TCA GGC GAA GGA ATG GCA GCT CGC TTT AGT TGG AGA GGA AAC TTA AGA A
SPB1809420.63	63	ATA TCT AAC GTA CAG CGA GCA GGG TCT TTT GGC GTG GTT GAA CGA GGT AGT GCA AGT TTT AAT
SPB1809920.57	57	AGC TGA ACA AAG GAA CTT TAT GAG TTT GCT TTA TGT CTC CCT TAG TGA TGA TGT TTT
SPB1810470.57	57	ACG GTG ATT AGG TTT ACG CTA AAC GAA TTA ATG CAG AGA ACC GTG TTT CCA TAA GTG
SPB1822570.57	57	GAT TTC ATC CAA AGA AAG CCT CAC CAA ATC AAG AGG TTC CGC GAC CTA CAA ACA TTA
SPB1823120.57	57	CCT CTG GAC GTG GTC ATA AAG GAT CTG GTC AAA GGC GTG GAC GGA GAA TTA AAC CTG
SPB1823570.58	58	TTT CAA AGA AAG ACA TAC GCT ACT ACA CAA ATC CAC ATT TTG CAG GGT ATC TTG CCA A
SPB1824020.61	61	TAG TCA GTT CAT TCT CTG TCA GAT AAA CCC TTT CTC ACA TTT ATT CAT TTG CTT CAT TAG T
SPB1824620.63	63	ATA AAT GAT TAA TGT ACT AAC ACT GTA AGA TCG AAA AAC TCG TGA TAA GGC TCT TGA ATC GTT
SPB1825270.63	63	CTA ACT GCT TTT AGT ACC GAA AAT CTT GCT TTT CCA AAT GGG ATT TTA TTC CAT TTA GCT GAT

SPB1825920.61	61	TGT TGC TTG ACT TGT GTT TTA AAC ATT CCT AAT GAA CAT TCG TAA TCA ATT ACC CGA CGT G
SPB1826770.58	58	GTG TTT GCT TTA ATG TCT TGG CGT TCC TAT TTG ACA AAT CTC AAA CTC CCG ACT TGA T
SPB1827320.60	60	TGA ATT TTG AAA ATG ATT CGG AAG CTG AAC AAA CTA GAT TTG AAC AGA ACT TTC CTC CCA
SPB1827870.63	63	CCT TTT TGC TTT CAA TGC TTG CAC ACT ACC CTT CCT TTT GAC ACA ATT AGC AGT GTT TAA AAT
SPB1828420.63	63	GTT TTA GAT TTT ACA GTT TCT TCT CAG TTG ATA TTG CTT AAA GGA ACA ATT TCC GCA GCC CTT
SPB1829020.57	57	CTC CGG GAA AGA AAG CAC AAA ACT ACA ATT CGC TCG TTT GGC AGC ATC TTC GAT TTC
SPB1836120.62	62	TTT CAC GGA AGC AAC AAC ACA ACC AAC AAC CAA CCC TTC AGA AGT AAC TTT CGG AAT AAT TT
SPB1836620.57	57	TCC CAA AAT TCT AAT CAG CAA ACT TGT GCT CGA AGA ACC CCT CGA AAA CCT CCT TTT
SPB1837070.63	63	AAT GAA GGC TAC TGA ATC AAT TAA TAT TTA TAC ATT ACC GTT TTG CAG CGA CTC AAT CCC TCG
SPB1837570.59	59	TTT CAA GCT CTG AAT TGT CCT CTA TTG TAT CCT CGA CTG GCT CCA TAG TGG AAA CAA AT
SPB1838020.63	63	GAG TAA TTA AAG GAA TGA ACA CAA TGT TGC TTG AAT TTC CAG CTG AGA TGC ACT TTT ACG TTC
SPB1838470.63	63	ATA CAC ATC TTA AGT GCA CAA TAT AAT TTA ATG AAC GTG ATC CAC GCA GAA TTT TGA AGC ACA
SPB1839120.63	63	AGG GGG AAT GTT GTA TAG TTC ATT TTT AGT CGT ACA ATA AAT CTC GAT AAA TTT GTA TAA GCA
SPB1839620.57	57	CTT ATG AAT TCG GGT ACG ACA CTT CAT CTG GAC AGA CTT AGT TTG GAG ACG CCA TCC
SPB1840220.60	60	AAT TGT CGA GCT TTC ATT AAT TCA AAT GCT TTT GGC AAG TTC AAA GGT CAA CAT TAC TGC
SPB1840720.58	58	TTC AGC TAA TGC CTC TCC ATT TGA TTC GTG TGA

		CTC TTT TCA TTC TGG AAG TTC GAT A
SPB1841320.62	62	ATG TGA ACC TGA ATA CAA TCG AAG TTC TTT GGT TAG AGC AAG TGA TGT GTT TAC TTC AAA CG
SPB1841870.62	62	GTT GCA TCT AGC CTT GGA TTT AGA CCC AAA GAC AAT AAA GAC AAG GAG AGT GGA GGG TAT AA
SPB1849020.58	58	CCT AGA GAC ATT ATA AAC ATC TTG GAT TGA ATC ACC CGA CTC AAT AGA CCG CAT AAC C
SPB1849570.63	63	TTT TCA TTA AAT GTG TTC TGT GCC AAT TGC TTT CTT AAA TGT TCC AAT CGA GAA AAT TGA AGA
SPB1850170.58	58	AAA CTA CCC AAA TCA TCA ACA ATT TCA AGC CAG CTG CCC TTC AAA TAT CCT AAG AGA T
SPB1850770.63	63	TCA GCA GTA ATA GTA TCC ACT GGA GAT AAA AAC GGA AAC ACA ACC TTA GTC AAA AGT TCA CTA
SPB1851320.58	58	GAG TTA TCA TGG AAG AAA ATC GCC TCA AAA AGG ACG GTT CAC CAT TTG ACT TTA TGA G
SPB1851770.63	63	AAT CAA AAT CTA AAA TTA AAT ATC GAA ACA ACA GTG GAT TCA TTA CAA GAG GAT GCT TAA AGC
SPB1852320.62	62	AGA ACG AGA GAA TGC TTG AAT ACT ATG CAG TAA ATC TTC GGT ATT TGA TTT AGA ATC TTG GA
SPB1852820.59	59	GGA CTA AGG TAA TCA AAT GCC ACG AAG GTG GAT TTA GAC TAG GAT TAT CTT TGT AAA TT
SPB1853270.57	57	CAT TCG ACA CTC CAA AAG TGG CAC AAT GTA ATG TGA TCT CAT TGA GCA AAT CTT TCG
SPB1853820.63	63	AGG AGA AGC AGC TTG AAT AGA AGA ATC TGT CTT GGA AGA AGT ACC TAT TGC ATA AAT GGA AAT
SPB1854270.62	62	TTA CTT GAA TAA AAG ATA GAA GAA AGA ACC TAA CAA AAA TAT TCG TTC CTA GCC CCT GGT GC
SPB1854820.57	57	TCA TGA AAT GAA CGC TTC AAT AAA ACA GTC AAG CAC GAG ATC ATG CAC CCA GAA CAC

SPB1862670.63	63	TTC TGT ATT ATG TCG CAA TCG GGG CAA GCA ATA TAA AGC GTA ATG CAT ATA ATT ACT CAA GTT
SPB1863120.62	62	ACG CTG AAG GTG GGT AGT AAG TAA AGT TAT ACT AGT CAC TCG TTG ATA ACA ATT AAA TAG CA
SPB1863720.62	62	GCA GTT AAT GTA TCA ATG CAA GCA GAA CCC AAA ACT TCA AGC AAA GAA TCT TCG GTT ATA TC
SPB1864170.62	62	TTT GAT ATA CCA CTC ATA TCT CCT AAC AAT TTC GTA GCA CGG TCT TCT ATA TGT ATG AAA GA
SPB1864870.57	57	CTC GCA TAT TCC ACT ACG AAA GTG ATT TGG CAT TCG AGA AAG ACG CAG GAA TCC GGT
SPB1865320.61	61	AAT ATC CAT TCC CTA ATA TTG GTT CAG TAG GTT AGT CAT ACG AAT AGT TTA TCA TTC AAC C
SPB1865720.63	63	AAC CAG AGT AAG AAT GAA CAA GCC GGT CAG TAA TGC AAA GCT AGT TCT AAT ATT ATA GAA GTT
SPB1866370.57	57	TTA TTT TCT TTG TGA AGC CAA GGC ATA TTA CAC GTA AGA CTT CCT AGT GGG ACG AAC
SPB1866870.58	58	GTG GAG GAC GAT ATA AGT GGC ACG TAT GGC AAT AAA CTG TAT TAA CAA AGA CAG AGT C
SPB1867520.58	58	TTT TGT TGT CGG TAA GGA TAT GAA CCC CTT TTG CGA TGG AGG GTG AAG TAA ATA TTA G
SPB1868170.61	61	AAA ATA ATG GGA GGG TGT AAT TGA GAA AAT ATT AGA AAT GAA GGT ACG GAA GCA ACG AAA A
SPB1868670.58	58	GCG GTC AGC AAC TAC TAT TAA GTT ACT AAC CTC ATG GCA GCT AAA TTT TGA AAA ATG A
SPB1876070.57	57	TAG GAA ATC CTG TTC CCC AAA GAA AGG CAT TTG ACG CAT TGT ACG TGC ATA ATA AGG
SPB1876620.58	58	TGC GGG TTT AAA TTT TGT AAT CCT AAG AAA TTG GCA AAG ACG ACC AAA CTG TAT AAT T
SPB1877070.63	63	CTG TAT GAC TCA TTT TCA ATC AAA TTT ACT TCG TCT TTT ATA AGA AGA TTC ATT GAC TTG AGA
SPB1877670.61	61	CAT ACA AAT GTT TTT GAA GAA TAC CTA CCA TTG

		GGG CTG AGC TAG AAG AAA TAT GGG ATT T
SPB1878220.58	58	ACT CAG CAT GTC TGT TTT GTA ATT ACC AAT CCG TTT AAT TTA TTT GAT GTG CGC TCA A
SPB1878770.57	57	GAA GGT GGT AAC CCC TTT TAT TGG CTA CGT GTA ACT TTA GCG AGT CAA GAT ATT GGC
SPB1879270.63	63	TGA AGA GGT GTT TCG ATG TTT GCC ATT TAT CTC AGA TTC ATT TGA TAA GTA TAC AGT TGA AGG
SPB1879670.57	57	TTC GCC TTG TTC ATT CGC TCT CTT TAT TTG ACA TTT ACG TGA AAT TTG GAT GGT GTC
SPB1880270.61	61	TTG TAC ATC TCG ATT TCC TCT TCC TTA ATC TCT CTT TGC CAT TGT TAT AAT ATT AGT GTT T
SPB1880820.58	58	ACG TGT CTT TCT TCT CTT CAA CAA AAT CGT ACC TTT GCT ACT CAA CCT ACT GAT GAT T
SPB1881220.57	57	CAC TTG AAC ATT ATG GTT TTA CGG AAT CTG ACC TTA ATC GCA CAA TTC ATC TGG GTC
SPB1881970.57	57	GTG TTG TAT ATG AAA CAT TTG GTT TAC ATG CTT TAC CTG GCT ACA GCA CTG GTG GTA
SPB1889470.63	63	TGA AAG ATA GAG GAT TGT GCT TTG ATT ACT TTT AAT TTC TAC TTA CTA GCC GTT CTC TTT TGA
SPB1890020.61	61	GTA AAT GAA ACC ACC CAA AAC ATA AAT CTT CTA TCC AAA CAA AAG AAC TCT AAA AGT CTG C
SPB1890520.63	63	AAT ATT TAA GTA AAC ACA TAC CGT TAA GTC TAA AGA CAT GAA ACG AAT GCT GCA ATC TGC TCC
SPB1891120.63	63	CTG AGA GCG TTC TTT TCC TCA AAT AGT TTA TCC TTT CTT CCA TAT AAA CGT AAA TGG TCA ATT
SPB1891570.58	58	AGT TCC TCC TCC ACC ACC AAC CGC CAT TTG ATG GTT ATT AAT CCA ACA AAG AGA ATA A
SPB1892220.63	63	CAA CTT TGC TAC AGA CAC TAT TTA ATA TCT CGG TCT CCT GTA CAT TTT ACC AGG TTT ACG TTG
SPB1892620.62	62	ATA ATA TGT TGT TTC ATT TCC TTA TTT AAC TAC CAA TTC TTG TTC ATT TGC AAT TCA CTC AA

SPB1893320.63	63	AGT TTT GCC TCC AAC TCT TAC CGT TGA AGA AGA ACT TGA ACT TGT AAA TTA TTA TAG TTT CCA
SPB1893820.60	60	TTT TAG ATT GTC AAA CAG TGC TTC CGA AAG TTG CTG TCG AAA AGT TTT ACG AAT GTC ATG
SPB1894520.60	60	GAC TGG TAA GTT ACC ATA AGC AAC CTG ACT TAA TGA TGC AAC CCT TCG CAT TAT AAG ATC
SPB1895120.61	61	TTT TTC AGC ATC AGT AGA TTC CTG ACT AAT TTG GGA GTT TAA TTC CAT CAT ATG TTT TGA T
SPB1895520.60	60	AAT TCC ACT TCA GGA AGA AAC CTT TCT ATG CAA GAA AAA CTA TTG TTG TTA TCC TTG TGC
SPB1902920.58	58	AAA TTC TTG AGT AAG ATT ATC CAA GCA AAG TTC AAA TAT TCG AAG CCC TTG ACT AGC T
SPB1903320.63	63	TTT TCG GGA AAA ACA TTC ACA GTC ATA AAG GCC AAC TTG AAT AAT CGC AAT AAA ACA GAA ACC
SPB1903770.57	57	AAA ACT TCA TTA GTT GAA CTA AGA TGT AAC TTT TCA GAC CAT TCC TGA GCT GTA AAA
SPB1904320.60	60	GTC GTA ATG AAT GCT GAG AAG TGA TTC CAG AAC CAA CTA GAA TTT TGG TAT CAA GTA GTT
SPB1904770.60	60	TTT GAA ATA GGA GAA AGA CGA TAA TGG GAA ATT CAG CAG TGA CCT TGA AAC TCG AAG AAG
SPB1905270.59	59	GGA ATG AAA GTT TCA AAA TTC GTC TTT TCA GTT ACC AAC ACA TCA TCC AGT GCA TCT CG
SPB1905820.57	57	CGA CGG TTA ATG AAG GAG TAT AAG GAG TTA ACG GAG AAT GGG CCT GAT GGT ATT ACT
SPB1906470.57	57	CGA GAC TAA CCT TCA TTT CAG TTT ACA AAG ATG GGA CGG TTT GCA TTT CCA TCC TAC
SPB1906970.63	63	TAA TCC ATC TTT TAA TAA ACA TCA CTT GAA CTT TCT ACA CAG ATA CAT GGC AAC GTA CAT CGC
SPB1907570.62	62	TTT CAT CAC GAC GAC TTA ATT CTT GTG GAC ATT TCT CTA TAA TCT CAT CTA GTA ACT CCT CA

SPB1908170.58	58	ATC CTC TAT TGA ATC TTC AGA TTT GGA ATC CAG ATT TTC TCG TTT TGG AAG TGT TTC A
SPB1908670.57	57	ATC CGC TGG GTA AAA ACT GCT TAC AAA AAC TAT GGG GCC GAG TGG AAA CGA ATA TGG
SPB1811020.57	57	TAC TCG TAC CAG TCT ATT CAT AAG GTC AGG AGC ATC CCG TCG ATC TGT ATC CGC CTG
SPB1811570.57	57	TCT TAA TAA GCC CAT TGC AAT TTG GCG ATT CTT ATG CTA CAA GGC CTC TTT GTT TGC
SPB1812220.63	63	AAC TTA TAC AGT ACA GAA CAG TGT ACA TGC TCA AAC ACT CTT AGT CCA TTC ATT TAC CTA TGT
SPB1817820.59	59	TTT TAC ACA ATA AGT GAT ATT TCA CTC TCC AGC ACA TCC GAA CCC TTA CTA AAC GGA AG
SPB1818420.57	57	TTC TGT TTC TGG CTC CCC TAA CTC TTC TAG CAA TAG TAC TCC CGC TAA TCA AGG TTC
SPB1818870.58	58	GGA TGT CAA GTT TTG TCT TGC TGG CTC TTT ATA CCG TTC TTG CGT CTC TTG CTT TTA T
SPB1819470.63	63	AAC TAT GCT TTA CTT AAA CCG AAT TCA TAA AGG TAT TTA CAA TAT GTG AGT AAT GGT CGT CCA
SPB1820020.59	59	CAA TAA GTT CAT AAA TGC ACT AAT CCA AAA GTT CAG TAA CCA AAG CTG TGC CTA CAG TG
SPB1820470.60	60	TAG TCT TTA AAT GGC TAC CAT AAC CGA CGA TTT CGA TTT CAG CAC CCT TCT TTA AAG TAC
SPB1820920.60	60	CAA CAA TGA TAG CGC CAT CCA TTG TAG CAG CAC CAG TAA TCA TAT TCT TAA TGT AAT CGG
SPB1821370.57	57	AGC CTT GAA AGA GGA GAG AGG TAG CTT TAG CTG AAT TCA TAA CGA AGG GAG CGT GTA
SPB1821920.63	63	ATA ACA AGA AGT GGT GCG TGA CTA AAG AAA GAT AAA ACT ATC CAT CCA ACG GAA TAA AGG AAA
SPB1829570.57	57	AAT GAA CTT AGC TGT CCT ACA GTT ATA TCA CCT GCA GCA ACC ATT CTT CCG CCT AAC

SPB1830170.63	63	ATG ATT CTA CCA AAG TTA CAA GCT GAA CCA AGG AAA AAT AAA CCC AAC AAA CCA ATA TAA AAC
SPB1830670.58	58	ACT ATA ATA TTC TGC ATT TGG CGT GGG AAG TAA TAG AGT ATT CGT ACT TTT GTA AGC C
SPB1831270.61	61	AGG GAT AAG CTA GAA CTT AGT GAT GAC AGG TAT GAA ATA TAC GGT TTG GAG AGG AAA TCG G
SPB1831920.62	62	ATT CCT GGA GGC ATG TCG TAT CAT CAT ATG CTT ATG TCT TTA TTA AAA GTA ATT AAA GAT GC
SPB1832470.59	59	TTC TTT TGG AAG CCT TTA ATG CTG AAC TAG AAG GGA AGC CGT CCT TAG CAA AAC AAT AT
SPB1832970.63	63	GAA GGA ATC ATC GAC ACA ACG AAA GGT GAG ACT ATT CCT CAA TTG TCT TAA ATT TTT CTG CTT
SPB1833420.57	57	TCT GGG TGA ACC AGT TTA TAG CCC ACT GAA TAA AGA CTG AAA AGT ACG GGA TTT TGC
SPB1833970.61	61	TGG TGT TCA TTC TAC CTG GTA ATT CAT CAT GAT CAT CGT AAT AAA TAG TCA AAT CTT GCA C
SPB1834670.61	61	GGG TAT CGA ATC TGC CAT CTT GTT CTT TTT CTG ATC AAA GAA TGA TAT AGC AAG TTT GGA T
SPB1835120.60	60	ATT AAG ATT AGA TTT GCA CTA GAA TGT TGA ATA GTC ATC TCA TCT TCT GTA ACA GGT GGT
SPB1835520.60	60	TTT GCT ACT ATT CTT CTC AAA ACC TTC GTC TCT TAA TCT ATT TAT TCC TAT TGG CAT GTT
SPB1842420.60	60	AGC AAT ATC TTC GCA ACC TTC CTA ATC CAA TTA TTA CCT ATG ACC AAT ACT TTC CGT TCA
SPB1842920.63	63	GTC ATT TTT AGA TGT CAT CGT GAA ATC AGT GGC AGA TGA GTG TTT TAT GGA TTT GGC GTA TAG
SPB1843420.63	63	TAT TTT ATT ATC TAT GTG AAC AAA ATT ACG GAA TTT TAC TAA TGG CAG ATG TTA TTA AAC GAT
SPB1843970.57	57	GGT AGC TAG AGC CTG TAA ACA GTT TCC GTT AAA CCA AAG GTC CCC TTT ACT TGA CTT

SPB1844470.62	62	AGC GTC TTG CTC ATC TTG AAA AGG AAT ATT TTG ATA AGG ATG TAT CCA AAT TAG TGC TTT GA
SPB1845170.57	57	AGC TGA ACA ATT ATT TGG GGC CGA AGA AGA AGA GAA GTA CGT TGA CCA AGA TTT GAG
SPB1845770.63	63	AAA GAG CAG TTG GTA TCA CGA TGA TAT ATG GAA CAT GAA ATA TCT TCC TAA GTT TAA ATG GCA
SPB1846420.63	63	TAT TAA GAC ATA AAT TAA TTC TTA TAA CAT AAA GGC ATT TGT CGG TAA CAC TCA TAA CAG GTT
SPB1846970.63	63	CAT CCG AGT TTT CAT TGT AAA TCC GAG ATA ACT TAT CTT TAA CAA GTA TTT TCC CAT TGA CGT
SPB1847520.63	63	AGA AGG ATC ATA TTG TAT GTG AGC AAC CTT TTT ACA AAT CCT AAT AAT GGG AGA GCA GCT TTG
SPB1847920.58	58	ATA GCT GGA TCG CCT TTA GCC CAA GGT GGT AAT ATA ACA TCG TTG ATA GGG GTG GAT T
SPB1848420.57	57	AGT TAA ATC ATT ATA GGT ACG CCC AGC TAA TGT GTT CAC AAT TTG AAG ATA GTT AAA
SPB1855370.58	58	TAG AAT TTA AGT CAA CAC AAG ATC TTC AGC CAA ATA AGC TAT GAA CCC TCA ATA AAC T
SPB1856070.57	57	GGA ATC CCC GAC ATA AGA CCT CTG TAG GCG TAG AGA TCT TCT TCT CCA TAA TCT TTT
SPB1856920.62	62	GAG CAA GTT AGG TAA CGT AGA AGT AAA GAA CTT AAA GTA AAC ATA ACG AGA GTT AGC ATT GC
SPB1857620.63	63	AAT AAA ATA AAA TAA ACT GAC AAT GGG CAG GAT AAA AGA AAG GGG AAA AAC ACG CAA AAC ACT
SPB1858170.63	63	GAT TCA GCA ACT CCG AGT CAT CAG TCT GTA AAC ATT TAG TTT GCA GTA TAT CTT GAG GAT CAA
SPB1858720.57	57	TAT TTG ACT GGG GTC TTC TTC ACT TTT GGG ATT AGT GGG ATT ACT TTT TGT TTT GGA
SPB1859320.63	63	ATT TTA GGA TCG CAG TAA TAT AAT TCA TAA GTC AAA TCT AGC ATT TGT TTT GCA CCT TCT TCT

SPB1859920.57	57	GAC TAA TGT CGT TTT ATG CAA ATC TGT AAG GTT TTC ATT GAA TGG ACG CTC CTC ATG
SPB1860470.58	58	CAA AGA GGA AAT GTG CGG AGA GCT TTT GCG GGA ATA ATG ATG ATT ATT ACG ATT ATC G
SPB1860970.60	60	GAT TAG AGG GAC TAG GTG AAG TTG ACT TGT TGG TGT GTC GTC CAA TTT TAT AAA TCG TTG
SPB1861420.63	63	TGT ACA TAA AAG CAC TTT AAA TTA ACG ATT AAT GCA GCA TAG TAA CCT ACT TTT CGG TTT CAC
SPB1862070.63	63	AAC ATA AAA CCA GGG AGA TTC CCA ATA TCA AAC CAT TTA TAA CTG CTT AGG CAA TTG AAC AGA
SPB1869120.63	63	AAT GTT GCG AAT GTT CGC CTA CGT ATT TCA TTT TGA CAT TTA CAA TCA TAA AAG GAT TTT AGA
SPB1869670.58	58	GAA GAC ACC AAA AGC AGT AGC GGA AAA AGT GAT GAG AGT CTT GAT TAT AGC TAC AAG T
SPB1870270.57	57	TTG CCG AAA ACA ACG CTT TTC AAC TTA GAA AAT TTA GCA AAC CCA CTC ATA CCT TTT
SPB1870720.62	62	TCT ACT ACA AGG GTT GAG CGT CGA TGT CAT AAA TAA AGT CAT TGG CAA AAT TAC TTT CAC TC
SPB1871320.58	58	TTT TCA ATC GCA AGG CGC TGT CAA GAA CTC TGA GCT ACT GTG GCT TGG TTT TTA TAA A
SPB1871820.60	60	TAT ACT ACT ACC CAT GCG AGA TTC ATT ATG GCT AAA TCA CCA TTC CTG GAG CTT TCT TAC
SPB1872470.63	63	TGA CAG ATG AAC AAA ATA TGC TAT CAG CCT ATC GCG GAA GAA TTG CTA ATT TTA AAT CAG GTT
SPB1873020.58	58	TGG GTG TTT AGG TTT GGG TTC TCT TCA TTT GCA AGA ACA GAC ATC TGA TCC CAG CAT T
SPB1873720.57	57	AAA ATT TGT TGG GCT CTT GCT CTC TAT TTG TCG TTT GCG TTG GTG GTT TGT AAT ATT
SPB1874270.57	57	GCC TTG TCT TGT CTT GTC TCG TCT AGT GTT GCT TGT ATT GTC TTA CCC TAT CTG TTC
SPB1875020.63	63	TGG TAT ATT CAA GTC GCA TTT TGC GTA ACC TCT

		GGA TTC TTT TCA TAT TTG TTT ATT AAT GCC
SPB1875520.63	63	CTC GCA AAT TTA ATT TAC TGT TTA ATG TCT ATG TTT GAC CTT CCG AAA CCT CTA AAG TAT CAG
SPB1882620.59	59	AAA AGG ATC TTG AAC AAT CGT AAC AAG TCA ATT AGC TCA GGT GAA GGT ATC GAT ATG CC
SPB1883170.62	62	ATC AAG ATT GTA ATA TTC AAG CCA TTT ATG TTA CTA AGC CAT CGC AGT ATT TCC ATG CTC TT
SPB1883820.61	61	GAT TTA CAA TTA TTG ATT GTG TTC CAA TGC CAT GAA TTC CTA TCG TAG TTA ACG AAT TTG T
SPB1884370.57	57	CGG CAC CGT CTG AGT GCT CGA TAT TTT CAT TTT CAA CTT CAT TCG TGA CAA TTT TGG
SPB1884920.58	58	ATC AAC GAG AGC TTG AAC ACC AGT AGT ATC AAT ATG ATT GAC AGC ACT GAA ATC TAA A
SPB1885470.57	57	AAA AGG AGA GTT TGT CGC CAA GGA ATG ATG AGA TCA AAA ACA GAA TGG ATG ATG ACG
SPB1885970.63	63	AAT AAT GAT GAC GGC ACT ACG AAG CAC ATT GGT GAG GAA GAA AAC ACG TTG AAA TTT AGT ATA
SPB1886520.57	57	GTA GCG ACT TTG GCG TAA GAC ATA CCT TGA GGT ACG ACA ACA CAA CCG ACA GTG ATA
SPB1887270.58	58	ACT CGA ATA ATT CAA TAG TAA ACA CGT CAA AAC TAA GTC GTT TCG GGG AAA ATT AAA A
SPB1887870.63	63	TTG TAC CCA CCA ATC CCA ACT TTA CTT TAA ATA AAC TTC AAT TTT AGG TAG CAA ACA AGG TCC
SPB1888470.57	57	TGC TTA GTA AAC ACT CCT CAT AGC ATC TCT CCA ATA CCT TCT TTC AGC ATT ACC AAA
SPB1888870.63	63	ATT TTA ATT ATT GAA GTG TGC TAT TAT ACC AAT TTT TAA GTG CGA TGC CTC ACG TCC TTT GTC
SPB1896070.60	60	CTT TAT TTG TTG CTG GAG CAT TCG TGC TTT GAT TGG TAG AAA GAA AGT TAG ATT GTA TGC
SPB1896720.63	63	TTT GTA ACA ACC TCC AAA CCT GTT TTC AGA TCT

		CTT GGG TTT TGT AGA TAA CAT AAA ACT TGC
SPB1897320.63	63	TCT TTA ATG GCA AGA GGT TGA GAT ACT GAC TTT TGA AGT GTC ATA AAG CAT TGA AAG GTA AGT
SPB1897920.58	58	GCA GCA CAC TTA ACG AAA GAT GAA ACC AAA GTA GAG ATC ACA TTC GGT CGA TTG TTT A
SPB1898470.58	58	CCA TAA AAT CCA TTC GTA TTT TGG TAT CTT TAG ATG CCG TAG CAA GAA GGA AGG CTT G
SPB1899070.63	63	GAA TAA CTG AAA GTA GTA GTT GTT TAA AAT CAT CAG CCT CAT TCG TGT CTT CTT CTT CTT CTG
SPB1899720.62	62	ATG AAT TGG TAA ACT GTC AAA GGC TGT GAT ATA CCC TGA ACA TCT TCG GAG ATA ACT AAT CT
SPB1900170.63	63	CAA ACA TGA GTG TAA ACA GAA TGA ACA ACA TCC TTA TCA ATA AAG TTA TTT TCT CCG CAT GAC
SPB1900770.63	63	GCA CCA ACG AAT CGA GTT ATT AAA TCG CCA TTA CGT GCA TTA TTC AAA ATA TAC ATC CAG GTA
SPB1901370.63	63	TTA TGA GCC ACC AAA GAA GCA AAT AGT AAA TGC AAG CAA GTA GCC CTT AAT TTT CGT AAA GAA
SPB1901970.63	63	AGC TTA TGA TGT GAA TCA TAC GAA AGT ATT GCT TTT AAG CCG AGA AAT CCA GCA TTC TTT TGG
SPB1902420.63	63	ATA GAT CTT TGC AGA TAA TAA GAT TCA TCT TTG TCA GGG CAC AAC CTA CCA TTA GAG ACT TGC
SPB1909120.57	57	ACA TAG ATG TTC CTG CTG AAT TTG AAG CAT TTG ATG AGC GTA CTG CTG TTG TTA ACG
SPB1909770.60	60	TCG TTG CAC TTT CTT GGG ATT TGT GGC TTT TAA TAT TTA CAA TGT TTG AGA TCA AAT TTT
SPB1910320.63	63	TGA TAT CCC CAT GTC AGT AGT TGA AAA GTT TTT ACC TTT AAA ATC AGA AGT TGA AGA AGA ACA
SPB1910870.57	57	TGC TGC CGA TGG TAC GGT CAT TCA AAG TGA AAT CTT AGG AAA AGT TCG TCT CAA ATG
SPB1911470.57	57	GCG TGC TGA AAT GGG TTT ACC AAG TGT TAA

		AAA TGA AGA TAT TCA AGT TCA GAA AAA
SPB1912020.63	63	AAC ATA ATA AAT TAA AAA CCA AAA GCG AGA GAA TGA GAG GGA GAC CGG ATA GTA AGA AAA AGA
SPB1912570.63	63	ATT TCA ACT TAG CCA GGG AAG GAA ATG GTA CGG AAT ACA AGT ATC CGG TAC AAT AAA TAA CTC
SPB1913170.59	59	AGC TTT GAT CAC AAT AAC CCA TCA ACT CAA TTG GGG TAT TAG TTT GCA GAT CAC GAT AC
SPB1913770.60	60	ATG GAT TCA AAA ACG CAT CCG CTT TTC GCA ACT TTA AAC AAT CTT TTC ATC ATA CTC AAT
SPB1914470.62	62	AAA CGT TCG TTT TTA TAA TCA GAG ACA AGA AGA TTT TAA GAT TTG CAG GGA GGT ATA CTC AT
SPB1915070.57	57	CAA CTC TTT CTA CTT GAA TGC GAA AGT GGA TTG GTC TCA CCA AAA GAT TAC CGT ACA
SPB1915770.57	57	AGA TCA CTT ACA AAC AGA AAA GTC GCT GCT GAG ATG CAA TAA CAA GAG GAG ATG AGA
SPB1916020.57	57	GGA CAC GTT TAA CAG ATT TAG GGC TTC CAG AGT ACA AAC ATC CTT CTA CAC TAT TGC
SPB1916570.57	57	TTT AAT AGG TGC CTT TCT TCT TCC CAG AGC ATC TAA CGC AGG TTT CCT TCC TGG TTT
SPB1917170.57	57	ATA CTA CAA TCA TTC GGC ACA CCA ACT ATC ATA GTA ATG CTT ATC CAT GTT ACT TTT
SPB1917770.57	57	AGA ATA AAG GTT GCT ATG CAC GCA AGG ATG GAG TGA ATG TAT GGA TAC TTA GCT ACT
SPB1918320.57	57	CGC AGT TTC TTT GCT ATC GGT CTT TTC CTC CTT GTC ATC GAC CTC TAT ACT ATC AGT
SPB1919120.57	57	ATT CTA AGC GAA GTC TCT CAA AAT CTT TCT CAT TCA GCA TGT GTG GAT CAG ATT CAT
SPB1919620.58	58	GGA TCG TCA TGT GCG GCT GTG TCA TAA GCA TAA GAA GTT TCA GCT TGT TCC ATT TCT T
SPB1920070.60	60	CAT ACG TCG GTA TTA GGA AAT GCT TCT TTA AGA

		GTC GCA CGC GCT TTC TCA GCC ATA TAA
SPB1920670.60	60	GGA AAG AAA ACA GCC ATC ATT AGT TTA AAA GTT TGG ACG TAA ACA AAG CGA TTA TAA CAT
SPB1921120.57	57	AAA AGC AGC AAA AAG GAT CGG CAG TCT CCT TGG TGA GGA AGA TAA TTC AAG CAA ATC
SPB1921720.57	57	TAA ATT TGA TAC TTT GTC AAA AGG TCG GGT GCC TCG ATG TAA TGG AAT AGG AAG GAC
SPB1922170.58	58	TTA ACG GTC TGA ATG TGA ACA ATC GGT AAT GTT TGT CTC AAT TGG TTT GAT GCT CTT G
SPB1929770.61	61	AAA TGT CTA TAG CGA ACT TAA TAA ACA CAC AAG TTG AAG TGA CAA ACA CCA TTG AAG AGC T
SPB1930420.63	63	AAT ATC AAG TTC ATT ATT TAA AGA GTG CAA AGT AAA ATG CTT GAC AGT TTT GAC TAA GAA TGA
SPB1930970.63	63	CAT TTG TAG TAG TAA AGC GCA CAG CTT CAT TAG AGC GAA GAT CAT TTT TAT AAA TAA GAT GGT
SPB1931520.58	58	AGC AAA TGT CGG GGT CAA AGT TCG ACT TGC ACA AGT ATT GAC ACC TTA CAA AGA AAA T
SPB1932170.57	57	TCT CGA CCA AGT ACA TCC AAA TTT ACC TCT TCG TGT AAG GGC GTC TGT TTT GTC AAC
SPB1932570.63	63	TGG ATA TAA AAT TTG CTT TAC TTT CAT TTG CAA CCG CTT TAG CCA ACA GAG TTT TAC CAC ATC
SPB1933120.62	62	TAA TTG ACG GAA GTC AAA ATC TCC ACT AAG TTT CAA ACC CTT TGC CAT AGT ACG TAA AAT TT
SPB1933720.63	63	CAA GCA ATT CGT TAA TAC AAT CAT CTA GTC CTC CTA TAT CAC TCA AAG AAA TGT CAC TTG GTG
SPB1922620.57	57	AAC GCA AAT TCA TTT GGA AGG TCT TTC TAC AGC GGT GCT AAC GTC GCA ATC TTA GAG
SPB1923220.63	63	GAT TTA CTA TAA TGG TAT CAT TAA GGT TGA AGT AAT AAC AAA CCA CCT CAC AAA CGC CGA TCT
SPB1923920.62	62	TTT TCC TTA AAA ATA GCA CAT TTA CGA TTA TTT CAT AGA CGA AAG CGT TTG TAA CAA TCC CG

SPB1924620.57	57	ATG GTG AAA TAT TCC GTG TTT TAA AGC CTG GTG GTG TCT TCG GTG TAT ATG AAT GGG
SPB1925120.63	63	TAG TTT TCT TAC AGT TAC ATT TGT TTG TAG TGA AAC CAA TGT TTG AGT ATT AAC TTT TCA GCG
SPB1925720.58	58	ATT TGG AGA GCA TGA AGG CCG CCA TCA AAT TCA TCT TTT AGT AAT CCA TAA ACA AGT C
SPB1926370.63	63	TCG GAT TAC CAA ACG CTG TAG TAC ACC AAA AGG TTT ACT ACT ATA TGC CAG TAT CTT CCT ATA
SPB1926870.63	63	CAC AAT GAC TGG TAA TTT TCA TAT GGC TTT ATG AGA ATA GAT TGA TTG GAA TTT AGT ATC AAG
SPB1927520.57	57	CAT GAG CCC AAA ATT GCA AAA GGT TAT GCT CGC ACA GAA GTC CTT GAT ATT AAA CAA
SPB1928170.63	63	TAA ACC TGT ATT TTC AAC TCT TCT TTA CCT CTT AAC CGT ACT TTG AAC ACA TGG AAT CTT CAT
SPB1928770.61	61	AAT CTG CAT TCT GGG ATT GAT TAC TCA TTA TAA ATG GAG CCC TGT GAA GAC AAA TGT AAG C
SPB1929220.57	57	CGT CGC GGA ATT AAA GCA ATT TTT AAG ACA ACA GTG CAA TTC TTG CCT ACA AGA CTG
SPB1916270.62	62	AAT CTC TGT ATA AAC GGC ATT CTT AGC ACA GGG CGG CTC GAC CGG ATA ATT GTA CTG ATT TA
SPB1916920.57	57	ATA TCT GAT CTA AGT GCC GTG TGG TTC GAT GAT TGT ACT GTT AAC CCG GTA TGC AAC
SPB1917470.59	59	AGA ACC GTA TTA ATG GTT GCA GCT TTT TCC TCT TCA TGG AAT CGG TAT AGA TTA TGT CG
SPB1918070.58	58	TAC TCG TAC CCT GAA TAC TCA AGG CAT ATA TAG TAC CAT TAG CTG ATT ACA ACT CTT G
SPB1918570.63	63	CTA GCT TTG TTA TTT TCG TCC TTA CTC TCC TTT TCT TCC CCA ACG TCA GAT TTC TCT TCT AAA
SPB1919370.63	63	CAC CCA TTC CAT ACT TCT TTG CTT AAA TGA TTG GAA TCC ACG CAT AAA TGT ATC TTT GAT GAC
SPB1919870.59	59	CAC AGT ACA TCC CCA AAG ACA CGA AAT ATA

		TGT GTA GCC TAA TAC TTA CCA AGC AAT GC
SPB1920370.59	59	GTG ATG AAT TTG AAT GAA CAA CGA GGC GTA CTG CTT CTT ACT GAG GGT ATA CGT CTT AT
SPB1920870.60	60	GCT GAC TTT GGT AGT GAA CAG TTC CAG AAG CAA GAG ATC CTG AAA CCT CGA AAT TGA TAT
SPB1921470.63	63	AAG GAA GGA ACC TTG TGT AAT CCA AGC TTT TAG AAG ATA TCA TGT TTC CTA AGA TAA AGC AGA
SPB1921970.57	57	GAA GAC GAC ATT AAA TCA CAT TGG GCA CAA ATT ACA GCA GGC CAG GTT GGT TCT TTA
SPB1922420.57	57	GGT CTT TCC TTT CTT AGA TGT ATG GAG CTC AAC AAA CGA TGC GCT TAG TGT GCG ATC
SPB1930170.63	63	GGC AGT AAT CTA ATA CTC GCC AGT TGA GTT ATT GAG CCA TTT AAA AAT TCT AAG TAA AGT TGA
SPB1930720.57	57	CAC CAA GGA TGG ATT CTA AGC GTG TAT TCA AAG CAG CAA CTC TGT CGG TGA GAT CTA
SPB1931220.61	61	ATC TTT CAG GAA ACA TTT TGG TTT TGA TCC ATC AAA GTA AAA ATG AAC TGC CTG TGA CTG G
SPB1931870.59	59	ACC CGA AAT CCG TAA ACA ATA ATT TCA ATC TAG TCA GCA TCA TTT GTA CTT GCC GAA GA
SPB1932370.57	57	AAG ACC GTC CAG TTC AGT AAG AAG GGT ATT TAC AAC TCT TGA CGA GGC TTC TGA CAG
SPB1932820.57	57	TTT GGC GAG TGC TTC AAT AAA ATC TTG TGG ACA AAT AGC GAG TGG TTC AAG TTC CTC
SPB1933420.59	59	TTT GAG CAC TTT CAC GTT TGG GAG TAA CGG CAT CAA TTT CAT CTA TAA ACA TCA AAC AG
SPB1934070.57	57	AGA GCA GCA AAG CAT ACC TTA ACT TCC ATC AAG TTT GAG TCA GAT TGT TCT TCC ATG
SPB1922920.63	63	TAA TGT CAT AAG GAA AGG AAT AAA TCG AGC AAT TCA TGG GAA TTT CGG TTA GCG CCA TAA ATG
SPB1923470.63	63	GAC AAT GAT TAA TAG ACG ACG AAA GCT CAT TTT

		ACC TTT TAC ACC TTC CTG TTT TGG CTT AGT
SPB1924220.57	57	TAC GAT TTA GCT ACC GAT TTG TAT GAA TAT GGC TGG TCG CAA AGT TTC CAT TTC TCC
SPB1924870.57	57	ATT GAC TGG TGA CAT TAC CAA GTG TCA AAA CAT TTG GGA TGT CTT CAC TGT TTT CCG
SPB1925420.63	63	TGT TTT GTA GAA TCG ACG GTA CAT ATT GGC AAC GTG TAT ACT AAT TTA AGT AAA TAA TCG CTG
SPB1926170.61	61	TGG AAT TCT TTT CCA CGA GTG TTA GAG GAC AAA CAT ACT TTT ACC AAA CCG AAT GAT ATG T
SPB1926520.63	63	GGA TTT AGT AGA GAA GCA TTG TTT TTC TTT AGT TTG TGA AAT TTG ATA AGG AGG CAG TTG TAC
SPB1927220.59	59	TTT GCT TGG AAA GAT TGA AAG TTA CTT GAA CAA TGG CTC GTC CAT GAA ACA CGA GGA AT
SPB1927820.59	59	TAA TGA AAA TTC TTG TGA TTT GAA TCT CTC AAG GAC CGC AAC CAT CAA GCT AGA ATC GG
SPB1928470.58	58	TCT TTG GGA GGG AAT GAT ATT TAC ACC CAT ATC AAT ACT GGA TTT CAC GAA AGG TTC G
SPB1929020.60	60	ATA TAC GTG AAT CCG AGA GAG AGC CAC CGC AAT CCC ATA ATT CAG TCT TTC AAC TGT TAT
SPB1929470.61	61	TTC AAC AAA ATC GAC TAC TTG ATA GGA ATT TGA AAG GAT GGA TTC AAG GGG AAT CAG AAA A

Appendix Table List of oligonucleotides for *ura4* microarray

Sequence Name	Bases	Sequence
SPC45000.63	63	AAA AAC TCG ATT TCT GTT CAA ATT TCC GAA AGT GAA AAG GTG AGT TAC GAT ACA GTG GAT TCT
SPC45550.57	57	TTT CAT TGG GTG ATT CAT TCA TTG CTT GAC TAT TTG ATT AGT AAA CTG TCG AAG AAA
SPC46200.62	62	AAT TGA AAA TCA GCA ATT AGC TTT GAA TAC GAG ATA TCA TCG AAG CCG AGA TAG AAC TTT TC
SPC46750.62	62	CTA TGG AAT GAT GCT TGA TTT TTC TAG GGT ATT CAT CTT TCA ACG AAC GAG AAA TAC AGT GT
SPC47350.58	58	CTA TAA ACT GTT AAC TTT ATG GGA AGG TTG CTT GGA AGG TTG CTT GGT TGG ATG GTT A
SPC47900.57	57	GCT GTT CAG ATC TTG CAT GTT CCT TTC TCT CAA AGC ACT TGA TAC ACC CAT TAC AAG
SPC48450.58	58	AGA AGT ACT TTT CCT TCT GAT GCC ATG GTG TAT AAT GAA TGG GTG ATC TTT GCA GCT T
SPC48950.61	61	ACA TAC CAA GCT AGA ATG TTG CTC GTT TCA AAA GAC AAC TGT AAA TTT GAT CTA TTC AAT T
SPC49700.58	58	CTA CTA AAC AAT CGT GGC GAC TGT GTG GAT AAC GTA TAT GCG GAA ATA TAC TAC AAT T
SPC50250.61	61	GAT ACT TAG CAA GTC GAT ATT TAT TTA GTG AAG TTT AGG AAT GTC GGT TAC GGC GAA GAT T
SPC50750.61	61	GAC AAT AGG AAG AAA AGG CAC TCT CTT AGT TAT CTT CAT TCA TTG GAA AGT CAG TTG GCT C
SPC51300.63	63	AAC TTT TAG AAT TTG GAT TAG ATC GCT CCC ATT TAA CTT CCG TTC AAT GTC TGT TAT GTC TGG
SPC58200.57	57	CCA ACC TTA GAA AAC ATG GTA TGT TTG GTA TAC TTG GTA ATG TCG TCG GTG CAT TCG

SPC58700.63	63	TAA ATC ACT CAA GAC ACA ACC ACA ATT TCT ATA ATT CGA TAT TCG AAC AAG GGA ATT ACA ATC
SPC59200.61	61	GAA AGA TGC TTA ATA AAA TTA AAA GCT AGA CTT TTC AGT CGC AAA CTT GTG CTG GAC ATG C
SPC59700.57	57	TTC GTC CAA TAT ACT TAG GAG TGC AAG TCT AAG GTC GAA GAA AGG CTG GAA GTG AAA
SPC60300.63	63	AGC ATC AAC TTA GTA AAC AAA ATA GTT ACT TAG ATT TCG CTT TAT TCT TTA TTT CCC TCT CCA
SPC60800.61	61	TGT TAA TGA TTA TTA GCA CTG ATG GAA TGT GCT CTT AAA AGG TCA CCT TTA GTG GCA AAG C
SPC61250.63	63	CTA CAT AGG ATT CAT AAA CTG CAA CTT TTA TTT ATC TTT TTC TAA GCA CAT TAG CTT TAA ACG
SPC61750.57	57	TGT ACT AAG TAG CTG AGC ATT TCG GTA AGT CAA GCA AGA GCA AAA CTG AAA TTG ACG
SPC62250.57	57	TTA TCA AGA TAA CCG AAA CAA CCT TAG TAG CTG CGT CAA TAC ACT CAC TAC TCA GTG
SPC62750.63	63	GTT AAC GCA ACA GAT TCG ACG CAA AGT ATA GTA GAT AAT TTA AGC AAA CTT GTA ATA TTG GAG
SPC63300.60	60	TCA CTC CTT TGG CAT ATT CTG CAT TTG AAG CTG GAA GGT AAG AAA TGA TCT GTA TTG TAC
SPC63800.63	63	TTG TGC AAC GGA AGT ATG ACT ATT ACA CTC AAC AAA AAT CGA GAA GTC TGT CAA ATA TAC AAA
SPC70600.57	57	GAC TCA CAA GAA TCA AGG CCA CAC TTT TGG CTT TAT CTG CGA GTG CAT ACA TGA TAT
SPC71100.57	57	AAA ACT GGC CTG TGG AAT AAG GTA TAG AAG GAA CGA GAA AAC AAA GAG AAA TTC TTT
SPC71600.61	61	GGT AGA CAT GAT GGT AGA AGG AGA ATA GAA TTA TGA TAG GAT ATG ATG ATA AGT ATC GAA T
SPC72150.63	63	GCT TTA CAA TCT CTC TTA CAT TGT ACT ATG TAT GAT GAA AAT CCA AGT TTT GCA TAT TTC GCC
SPC72850.63	63	GAT CCT CGC TAC ACT ACT GAA CCT CTA CAC ATC

		TAA CAT TTT GCA TAT TTT TCT AAT ATC CTC
SPC73450.60	60	ATT TCA ATC ACT TCC AGA AAA ATT AGC AGA AGT ATG GGG TTT TCT AAA ATA GCA CTT TTT
SPC73950.61	61	CTC GAA TCT TGT GCC TTT TAA TGA CTC ATC CTA CTT CCA TCC TTA CTG TTG GAT TGA CTA T
SPC74500.60	60	ATT GTT TTC CTC CTT CTG CAT ACT GGC ATT CCC ATT ATT TAT TAT GGA GAA GAA CAA CGC
SPC75050.57	57	ATA CAT CAT CTG CCA CAT CAA GCT CCA AGT CTA GCT CTA GCT CTA GCT CCA GGT CCG
SPC75600.59	59	TTG TAA ACA TAT TTC GCT TTT ATG TAA GGA TAC TGA GAT TCT ATC ATA TGC GTT CAA AA
SPC76050.58	58	AGT CTT TAA GAG CAA AAG CAG TTA AAC GGT TTT GTA GAA CAG GTA GTG AAA CGT CAA T
SPC76700.63	63	TGT TTT TCA AAA ATA AAG TCT CTT GAA GGA ATT TAT AAA CGG ACT CAC AAA CAC AGA ATC AAC
SPC83650.63	63	TTT TCG CTT GGT GCA TAT ACG GTA TCC TAT TGT CTA TCT TTT TTA AAG AAA TTC GAG CCG ATG
SPC84200.58	58	TAT ATC TGC TGC AGA TAC AGT TGG TGC TAT CTG TGG AGG CGC TAT ATT TGA TAT AAC C
SPC84900.62	62	TCA TGA GTA AAA ATT GGT TGA AGA AGT AGC TCC ATT TGT AAA AGC GTA AAA TTA GTA GGA CT
SPC85400.59	59	TTT TAT TCA AAA GTG CGT TTG CAG GAT GGA TGG ATA CAC ACA TGA AAG ATG AAC ATC AC
SPC85900.57	57	ATC CCT GAC AAG TAT CGT CGT GTG AGC ATT TAA ATT TTC CCA TTG ATA GAT TTA CTT
SPC86450.63	63	TAT AAG TGC GTA ATG TTT CAA CAG CCT TTG GTG TTT TTA TGA ACT TGT TAA ATA TTT CTT GAT
SPC87000.62	62	CTT TGA CAT AGG ATG AAT TGC TCG GTT CGA TGG AAA CGC CTT AGT AAA TAT ACT ATG AGG AG
SPC87650.63	63	TCA TTG AAA TGA GTA TAG CAA CTC TTG AAG TGT GAT TAT CTT TGA TAA ATT CGC TAC ATT CTT

SPC88200.58	58	GCC AAC AGT TGA AAG AAA GAA AAG TTT TAA TGG AAA GTC CGT CGG ACA TAT TGA AGG C
SPC88750.62	62	GCA GCT AAG ACG TAA CCC GAC TCT ATG TCA AAT TTG GGT AGT AAC TAC TGT TTA ACA CTT AC
SPC89200.57	57	TAA TAA AAA CCG ATG CAA AGA ATT GGT TGG CGA CAT CTA AAA AGG TGC TAC AAG CTG
SPC89750.57	57	GTA CGT TAG AAA AAA TCA AAC AAC AGC AAC ATG GGC CGC GAT ATT TAC AAG GAC GAG
SPC97100.57	57	AAT TCC TTC CTT TCC AGT TAA CAG AAG AAG TTT TAC CGA CCG CTC GCA TTG AGT CTG
SPC97700.62	62	CTG CCT CTG AAT CTT TCA AGG CCC TTA AAT CAT TAT CTT CTA ATA TCC CCA ATG ATG TTG TT
SPC98150.63	63	TCT CCT TTA GGC AAC AAT GAT TAA TGT TAT TTT CCA ATG AAT TAA TAC CAA TAA ATA ATT TGG
SPC98600.62	62	TGA ACT CTT GAA CGT TTC ACT CCC AAT AAT TAT ATA CAT TTC CAA TAT AAT AAT ACA CGT GA
SPC99050.57	57	GCA TGG TGT TCA TGG TCC CTT GCA TAG TGT TCA TGG TCC CTT GCA TAG TAT TTA TGG
SPC99500.57	57	AAA TTA ATA TTG TTT AGT AGT AAC GCG AAT AGC AAG CTG GGT GCA GCC AAG TGT AGT
SPC100100.63	63	CGA AAA GGA ATT CGA TTA AAT ATT CAT ACT AGC TGA GCA TTA TTT CTT TTA TAC CTT CTC TGT
SPC100650.57	57	TAC CAC TAG AAG TTA CCA AGA TTC CAG AGT TAC TTA TGG TGG CAT GGT TAA TGT CTT
SPC101200.58	58	TTT CAC AGC ATT GTC GAA TGC ACC TGT ATA AAG AAA CAG ATT ACC CTT GGA AAA ACA C
SPC101750.63	63	GGT CCA AAT TTT AGT TTA GCT TCT TGT AAA ATA AAG CAT TCC TCT GGT GTA AAA TCA ACA TGA
SPC102250.63	63	CAG AAG TTA TGG GTT GCT GAG TGA GTT TAA GGC ACT GAT TGA TAT AAA GAA TTT ACT AAG TCC
SPC102750.63	63	CTA TGG TTG CAT GTA TCA TCT GAG CTT TCT TTT

		CTT GTA GTT CAA TAA TTT TGC TTT CAA TGG
SPC109850.59	59	CGT TGG CAA TTG ACT TCT TGT TTA CAC CAA GTA ACC AAC AAA AGA GGT CTT TAA CAA CT
SPC110400.61	61	TCT TAC GTT CTT CTT TCA ACC CTT TTG TAT CCT TAG AAT TCG CTT CCT GAG TAG TTC CAT C
SPC110850.58	58	AAT CAT TAA ACG ACA AGG GCC TTC CGT GCT ATA GTG TTA TGC TTT GCC GTA ATT TAA A
SPC111350.63	63	TAA TTA GTG TAA AAC AAT TAA AGG TTA TCA AGC GTT GTT CAT ATA ATT GAT CGT GGC CTG TCT
SPC112050.57	57	TCG TGA AGC TAG TTG TTT CTA AGT ATA AAC GAT TGT GTA AAT CTC TTA TCG CAT AAA
SPC112550.60	60	GAG CAA AAC TTC TCA TAA TCA TGT TGT TAT TTC GAG AGA GCC TCA TGT CCA TAC GAT CAA
SPC113050.57	57	GGC TCA TCC AGA CTA TCT AAT GTC CTT TCA GAT CGG CCC CAT TCA TTT ACA ATA CCG
SPC113550.63	63	CCT TAT GTG ACT CAC TGA CTA TTC ACA TTT GCT GAG CAT ATC TCA TCC ATC TGA TCA TGA ATG
SPC114050.60	60	ACA AAT AAA ATC ACC ATA ATC TTG CCG CTG TCA TTA GGA ATT TCA CTT ACA AAA GAC GAA
SPC114800.57	57	AGT CTA GCT TTA CAG CTT GGC ATT GTT CAT ACA AAC GTC TTC AGC ATA TCT TTC CAC
SPC115350.63	63	TGT GAT ATT GAC GAA ACT TTT TGA CAT CTA ATT TAT TCT GTT CCA ACA CCA ATG TTT ATA ACC
SPC115900.57	57	TAC CGC AGT TTA CAA TCA CTT CTT CAG GAG TAC GAT ATT GCT GTC CCA GCC CGT CTC
SPC122800.57	57	TTG GTT TGT CTG CAA TGT TGG ATA TCA AAC TTG AAA ATA ATT TAG ACT ATA GCA TGC
SPC123550.63	63	ATA AAA TTA TCA GTA GGA TCA TTA GCA AAC AAA AGC AAT TGA GTT CGA CAA GGA GGG CAA AAG
SPC124200.57	57	TCA TTC CGT TCG TTG ATT ATT TGC CAA ACT CAC TAC CAT TCA ACT CAT ACT GTA CCT

SPC124800.58	58	TCT GTC CAA ACG GAG TTT CCA AAT GTT ATT TCT GTT GTG TAT ACT GGC GAC AAT GTT A
SPC125250.57	57	GCG GAC TAT GTG ATG ACT AAC TCT TCA TGG ACG AGA AAC CAC ATT GCA TCT CTT TGG
SPC125800.57	57	GCA CCA TTA GCG AAT ATG CTG AAG CCT ACC ATA AAG CTC TTA CAT TGA GTC CTC AAG
SPC126450.61	61	TGA AAA GCG AAA CTT AAA CAA GTA CTA TAA TGC CAT TGT AAA ACT AAT CTC CAT ACT AAC T
SPC127050.63	63	GCA ATG ATG AGA CCT CAG TTC TTT ATA AAA TAA TCA ATC TCG ATT GCA TTT AAT TGT TAA CCA
SPC127600.61	61	ATT CTA TCA GAT GAC TGG ACC AAA AGT GTT CAT TTA CAA ACT GAT CGA ACG GTC GAT TTT C
SPC128100.57	57	AAG TCC TTA CAT GAG CAA AGA CCA GGG TTA CAG CAT GCC CAT TAA ATC TCT TCA CTG
SPC128550.60	60	TCT GTA TGA AAA GGC AAG ACT GAT TGC TAA TCC CTT CTC TTA TGA GGA ACA TCG TCA AAA
SPC129050.59	59	CTC CCT CTG GTA TGG AAA TGA CAT TCA AGG TTG AAA AGA AAA AGA AGT CTA AAC CTG TT
SPC135900.57	57	GAG AAA CCG TGA CAA GCA ATT ACA AAA CGA GCA AAG ATA AGG TTT GTT TCC ATT CAC
SPC136350.57	57	ACG AGA ATT AGC CTC CGT CTG GGT CAA ATG TTC GGT TGC AAT AGA GCG TCT CCA AAC
SPC136900.61	61	TTG GCA AAG GAT TTC ATT TAC TAG GTA AAA ATT TGT CGC CTC AAT TTA ACA GTA CTT TTT G
SPC137450.60	60	AAA CAT GTA AAT GAA TAT ACA CAA AAG TTA AAG GCT GCT GCT GGC CCT TCA AAT TCT TCG
SPC137950.60	60	AAT ACT GTA CCA CTT TAA ACA CTG CCT GCT ACA ATA GTG CAT AAC TTC TAT TTC ACT TGG
SPC138550.57	57	CCT AGT AGT TCG TGG TCG ATT CCT CTA AAC TTT CGT GTA TCG TGT TGT GTC GTA TTT
SPC138950.57	57	TTC GGT GGG ATT TAA TGA ATT GTG AAA CGT AGC

		GCT TGC TGT TTG TTA TTT ATT GGT
SPC139550.57	57	CCA TCG TTT CCA GCT TCA GGT TCA ACT CCA TAA ACT TCG CAG TTT GGA GCA AAA TGG
SPC140050.60	60	TTC ACA TTT AAA AAA CAC TTC TGC CAC AAA CTC TTT ATT TAC TGT GCT CGA CGT GAG TAC
SPC140500.60	60	AAA CAG TGA CAT GTA GCA TAG CGA GTA TAA TGC TAA TCA CTA TTT GTA GTA TTC CTC TGT
SPC141100.63	63	ATT GTT AAT TCA CTT TTC GGT CCT TTT CTT TTG GTG TTA CTC TAC TTA CTG AAT TCT TTT CTT
SPC141750.63	63	TAC GAC TTG CAG TAT CAA ACC AAC GAC AAG TAA ATA ATG TTA GTC TAG CAA ATG GGA AGG AAA
SPC51800.62	62	AAT TGC GTT TTG AGT AAT CTT GGA AAA TAT AAT TTA GAG CTA ATG AAT TGG CAC TAT AAA CT
SPC52300.59	59	ATC TAC GTC TGA ATT TGT ATC TCC AAT TAG CGA TAC TGA AAA CGG CTC TTC ATC TCA GC
SPC52850.62	62	ATC GGG CAT CAC CAA GGT ATT GTA CAC ATA CAA AGA ATG ATG TTA TGT AAT TGA AAC ACA TT
SPC53400.58	58	GAA TGA CAA GCT TTG CCG TGA GCA TAA ACG TCA AAC CTA AAG TTA CCC TTG TGT CCT A
SPC53850.63	63	TAG GAT TTT AAG AAA TCT GAC ACA TTT ACT TCA TTC CCA CCA AGA GAT TCT ATT TCA ACC AAA
SPC54400.59	59	ATG ATT TCG TTA ATC AGC GTT AAG TCA GCG TCA TAC TGG AAG TCA GCT CAA TGG AAA TA
SPC55000.58	58	ACT CAA GAA TAA GTC ATG GAC TAC GGA CTA AAA TTA TCA AAT TTC TCA GTA ACT TAA G
SPC55600.61	61	TCA AGA GTG GTG TGT TTT CAT GGC TCA GGA AGT CCA TAA TTA AGT TTC AAG TTG ATA TTT T
SPC56100.61	61	GCC ACC AAA TTC ACT ATG TCG GCT TTC TAT GAT AAT TCC TCT AAT TCT TCA CAA TGA TGT T
SPC56650.62	62	TGG TAC AGA TAT TAA GAA ATT TAC GGC AGA AGA CTT GGA AAA CAC AAT AAG CCT TAG ACC TT

SPC57150.58	58	TTG CCC AAA ACA TTT TGC CAA AAC GCT CTA GGT TGC TCA AAA TCA ACG TCG GTA ATT T
SPC57600.57	57	CAG ACC TTA GTA AGA TCG AAA ATG TTG TAG CGA TAT TTT TCA GCC TCC TGA GGG GTC
SPC64350.57	57	CAT AAA CAG TGA ATG CCA CAA GAC TAT TTG CTC TTT TAT GTC CTG AAA TAT TCC AGA
SPC64850.62	62	AGA CGA AAT GTA TTG ATC CCG AGT TAA ATC ACC AAA ATA CGT ATC TAC GAT GCG AGA CAA TC
SPC65350.57	57	ATT GAT GAA CGG GAA AGG GGA AAA GCA AAA GGA TAA ACT GAA GGA GAA TGA GCA AAG
SPC66000.57	57	ATT TCA ACT GTA CGC TCA GTA GAC TGT GGA AGC ATG TCT TTG GAA GCA ATC ATA CGG
SPC66500.60	60	TTG AAT AGT AGA ACC GTT AGT AAG AGT CGA TAG AGG TGT AGA GGG CTT TGG AGT AGG TAT
SPC67000.63	63	GGT GGG GAA ATA ACG AAG TTT GAA GAA CGT ATA GTA AAG TCG CAC AAT TTA CTG ATA GGA AGT
SPC67550.63	63	ATT CGA ATA CCT ATT TGT TCA CCC GTA CCT CTT TAG GTT TAT ATT TGT TCT TGC AAA CTG TTC
SPC68050.57	57	CAC AGT AGC GTG TAC CAT TCT TCT ATA CCT CAC TGA ACG AGG AAG TCT ACA ATT CTG
SPC68550.63	63	AAA AAC AGC AGA CAG ACT TAC GGT GTA ATA AAT TAA CTG TCG ATG GAT TTG AGC TGG TAA AAA
SPC69100.63	63	CGT ACG AAC TTC AAA GAA TCA TAC TCG ACA AAT AAT ATA TAC ATG GAC AAC AAA TGA CCC TGT
SPC69650.58	58	GCG CAA CTT GAC CTT AAG TTT CTT CCA TAT AAT CAA TAG TGC AAG CAC AGA CAT AAC G
SPC70200.62	62	ACA ACG AAA ATA ACA AGA CAC CCC ACA AAT CCT AGT TTT GAA TTG GTA TCC TGT TTG ACA TC
SPC77250.58	58	CAA ACC GGG GAA TGT ACG CGA AGT AGA AAA GAT ACG AAA AAT GTC TGA TCA CAA GAT T
SPC77750.62	62	TAT AAA TGA AGC TCA ATT CCT TTC ATT TTG AAA

		CCA GAT TAG GAT TTG GTG TTT TCA TTG CT
SPC78300.59	59	TGA TTC AAT ATT TGG ACA AAA CAG CCT TGT CTT ATG CCG CTC TTT ATG GTA TGA AGA CG
SPC78800.60	60	GTT TTT ATT CCT AGT AAT CCC AGT AAA GCC CGT TTT CTG TCT TCT CGT GAG AAG CGT ATT
SPC79350.57	57	TTT ATT CAG TTG GTA ACA TTG TGT CGC CTC AGT TAT TTA AAT CGG GTC AAA CTC CTG
SPC79950.58	58	TGC AAG TGG TAC ATC TGT TCA CTA CTC TTT CTA ATA ATG TAT CAG GTG ATA TAG TCG T
SPC80550.57	57	AGC ACT ACG ACC GAA ACA GTG ACA TGG AGT CAG TAT AAA CCT CAG GAA ACT CAA AGA
SPC81000.57	57	ACG CTC TAC TTT CGA CCA CTG ATT GTT TCC TGC ATT TCT CAA ATA ATA GGC CAT TTA
SPC81500.63	63	TTT CAT TTC TAT CTT TAA TGT TGC CGG TTA TCA AAC CAG TGT ACC TAT TTA CGC TAA AGC ACT
SPC81950.60	60	TGG CTT TAT TGC TAT GAA TTT GGG TAT TGC CTT ACT GGT TTT TAT TCA GCT CCT CTA CTT
SPC82650.63	63	TTG CAT TTA CCG AGT CTT TAA TAT GTG TAT GTT CTA TAC TTC AAC TTC ATG GCC TAT GTT TGC
SPC83100.59	59	CCA CCC AAG AGA TCT ATT GCA TTT GTC CTG CTT AAC AGT ATT CTC TCA GAT ATG TCC AT
SPC90650.57	57	TAC TCA TGT AGA GCT TGG TTG AGA AGA TGT AAA CAA ATT GCA ATA TTC AGC CCA ACG
SPC91150.63	63	ATT CCT GCT TCC TTT TCA TAA ATG TAG TCT TGA TGT TCA AGT TTT CCA TCA TGT TTT ATT GTG
SPC91750.57	57	CAA GAT GCG ATT CAC GAT TTT GTT GTT TGG AAT GTA ATT TAG AGC ATA AAC GGT TAA
SPC92250.58	58	AAA GCC AAC ATA CAC GCA CCT GAT GAT ATG CAG TTT ATG ATA AAA TCT AGA AAG TCG C
SPC92800.57	57	AAT ATG GAT TTG GGA TAT GGT ATC TAG GTC ATA GAG CGT CCA TAC TGG CAA AGG ATT

SPC93300.63	63	ACC CTC AAC CCA TTG ACT TTC ATG CCT TTA AGA ACT TGT TAA AGC AAT TAT TAG GTT ATG ACC
SPC93850.63	63	TTA CTC ACT ATT TGC TGC TTT CAC GGA CTT ATA CTG AAA TTG AAT CAA AGT TAA TGG ATG ATG
SPC94300.58	58	TGT ACA GTG ACT TTA GAC ACA GGC GTT TTA GAT TTA ATG GTG TCG AAC ATT TGC TTT T
SPC94900.59	59	AAA TAA CAG AGA CCA AGG GAC ATA CAC TAT CCC AAA ACG TGG ACA AAA AGA CTT TGA AC
SPC95450.59	59	TCG AAA ATT ATG AAT TTC AAG GAT TTA TTG TTT TTA GGA TTG GGA AAG GCT TCA CAA AT
SPC96100.58	58	GCT AAT CTT TGA TGT CTT GGT CGT GGT TCA CTA GTC TTT GAA TCT AAC CAT TTC TGG G
SPC96550.58	58	CTA GTT TCT CCA TTA TAA CGC AAC TGG CCA ATT CAA ACT TTG TCA TAT CCT CTT GCA G
SPC103300.57	57	TAT GCT TAA GGG AAT GAA ACA ACT TGG ACA GTT GAC ATT TTC TCC ATC ACC AGC TGC
SPC103800.60	60	TTT GTA TTG GTT TTA GCA TTT CAG CGT TAA AAT AGC AAG TAT GAC TCA TTG GTT TAT GGC
SPC104300.57	57	GTA TGT GTA TCA ATC TCT TCT TTC CAT TGC ATA ATT GCT ACA ACC GGA GCA ACA ACC
SPC104850.59	59	CAT CAT CAA TAA ATT CGG ATT CAT TAG ATT CTT CTG GAG ATG ACA ATG GCA CAA ACT TT
SPC105400.57	57	AGT AGA AAA TGG AGC CTG CCC ATT ACT TTG CTC ATC ATT ATT GTC TGC TTG AAG AGA
SPC106000.57	57	AAC GGA TGA AGA AGC TGA AGA CAA CGA GGA CAC TTT TTC AAT GAA TTC AAG AGC TGG
SPC106600.62	62	TTT GAC TGA ATT AAC AGA TGT ATC TTT TCA AGG TGC CTT TTT GGT TTC TTC CTC AGA ATG GA
SPC107100.59	59	AAA CAT TGT CCT TAA GAA GGT GTC TGT CAC TTG GAG ATA AAA CTG TTC GTG CAG TAT TG
SPC107650.63	63	TTA TTT TCA AAT TAG ATG ATA TGG TCT GCG AAC

		AAA TAG CTA ACC ATT TCT TTG GGT AAT GAA
SPC108200.60	60	GAA AAT ATT GAT TCT AAT GCT TAA TAA TTC GAG TTT ACG ACT ACG GTC ATG AGC CTC ACT
SPC108850.63	63	CTG CGT AGC TAA CAG ATA ACA TAC ATT TAT TTA CTC ACA TAT ATT CTG GTT AAT TAT CAG TTT
SPC109350.63	63	TTT GGT TAT CAA AAT TCC AAG CTT ACT CGC TCG AAT TTC TCA ACT ATT AAT CTA TGT GTC TCA
SPC116400.62	62	CAG ACA TAG GGT CCA ATT TTA TCT ACC AAT TCT AAG ATT TCG GAT TTC TTC GTC AAA TCG AC
SPC116900.60	60	ACC ATG CCA AAA ATT ACA CAA GAT AGA ATG GAT GTT TGA AAT TAA ACG TGA GTA TAC AAA
SPC117400.62	62	ATC ATT TAT AAA ACA ACT TCT TCC ATT AAA AAT TCC TTG GGC AAA ACA AAA GTT CCA ATC AT
SPC118000.61	61	TAC CCA AAT TCA TCT GAC ATT GAT TAT GAT ACA TTG AAG GTG TGC TTA CAT CTT TCT AGT C
SPC118650.57	57	ACT TCT TCC AAG CCT TGG TGA CGA AAA CGT CAT TTA CGG CAA CGA CAT AGA TTC CGC
SPC119200.57	57	ATG GTT CGT GCG ATA TTC GGT ATT TGA CTG TTG TAA AAA GGC AAG AAA CGC TGA GAC
SPC119750.58	58	GTG GTA TTG CGA CAC AAT TAG CCA GTG CTA TGA ATA GAG CGA GAA AGG TGG AAA AAT T
SPC120100.61	61	AGC TAA GTG TCT ATG TTT ATA CGA TTT ACC GAA CTA GAA GGT AGG ATG AAT TTG GGT TTC C
SPC120650.63	63	ATG GGA TTG GAA AGC CCA ACA TTA AAT GGA TAA AAA TCA TAT GTC AAA TAC ATT CCG ATG GAA
SPC121200.58	58	TGG GTA GCT TGT GAA CTG CAT CTA ATA CTA GAT GAA TGG CTA CTG CCA AAA TTC CCA A
SPC121650.57	57	AAA CAG CGT TCT TTG AAT GAA TTG TCG CTT TTG CTA ATA TAC CTC CAA GGA AGG GAC
SPC122200.58	58	TGG AAC AAG ACT TGA ACG ATT ACT CTG AAA TGT TTC AGT AAC CGA CTC GAC TTC TGA T

SPC129600.63	63	CAA CAA ATT ATA TAT ATC TTA GCT GTG GTA ATG TAA AGA ATC TGG TAA TCC ACT TCG TAT TGC
SPC130200.62	62	GTT TAA CAA CTG GAT CAA ATC TGT TCT CAT ACA AAA ATT CGC ACC CCA TGC TTC TGA TTA TC
SPC130700.61	61	CAC CTC CAC GCT CTT TTC GAC CTC CGT ATG GAA TTC AAT ATT ATT TTT ATT TAG AGG ATG C
SPC131250.58	58	ATA AAT AAA GTA TGC TCC CAG GCA TAT TTC ATC AAA TCA GAT TGC CCC AAA AAT TGC C
SPC131700.62	62	GAG AAT AAA AAT TCC TTT GTT AGA TCT CTT TTG CAG CAT AAA GGC ATT GAT GTT TTT GTG CA
SPC132250.63	63	TCA TTC ATA AAG TTA TCC AGA TTT CTG TTA GTC ATA CCC ATT CAC TTG CTC TTA CAA AGT TCG
SPC132800.60	60	CGA AGT CTA TGA ATT TGA TAT GAA GCT TTT ACT TGA TCG TGA TTC TAC CTC TTC CAA GAA
SPC133300.63	63	ATG AGA TGC AGC TAC TTT TTG AAG GAT CTA TAC CGA TTC TTA CTA GTT ATG AAA ATT ATA AGC
SPC133900.58	58	ATA GTT TTG ACA TCC CAA TAA CCG TCA ATT TGT CTC ATT TGA CTG TTG AGC ATA TGT C
SPC134400.61	61	CTT AAA GTT CTC AGA GAA TAC TTG TTT TCC CAG GAA AGT TCT CAA TTA TGG GAC GAT TCC C
SPC134850.63	63	ACG GTT TTG TCT GAT TCT AGA TTT ATG AAA GCT CCA ACT AAA AAA TCC CAA CGT GAG AAA AAA
SPC135350.63	63	TTT GAG AGT ACA GAA AGA ACT TGG AAT CCT TAA AAC TGA AAG AGA TAC CTC TAC GAA TAG AAA
SPC142250.58	58	GCT AAC GCC TTA TCC TAC TTG CAT AAG AAG CAT GTA ATA CAC CGT GAT ATC AAA CCT G
SPC142900.57	57	CTC ATT TAA TAG AGG ATC ACC ACC ATA AAC TAA TCG TTT TCC AGA TTT TCT TTT CGA
SPC143400.57	57	AGT AGC TCG TCG TTT GAC GCA TTC TTG ATG ATG CTT CCG AAA AGT TTT AAT GTT ACC
SPC143900.63	63	GAC AAC AAA CGA TAT CTC TAA TAA ATG TCA ATT

		AAT AAA TTT TAC CTG TAC AAC ATT AGG TGG
SPC144450.63	63	CTA TTT CAT TGT GTT TGG AAT ATT TCA TTC CAT CGA GTT TCT TAT AAA GTG TTG GGA TTA CTT
SPC145150.62	62	ATT GCG ACG CCT TGA GCC TAC AGC TAT TGT TGC TTT CCA TCA AGC AGA TGT TGG TGA ATA TC
SPC145750.62	62	AAA TAG CCT AGC ATC TTT CGA CTT TTG CTT CGG ATG TAT AAG AGA AGT TGA CTC GGT ATA AC
SPC146300.58	58	AGT AGC CTT AGG ATT GAA GCT TTT AGG GGT GGG TTT TCT TCT CTT CCG AGT CTT TTT A
SPC146750.57	57	AAT CTT ATA GTG TAA TGC ATC GCC TAA TCC AAC CAA GCC CGG AGA ATA TAA AAC CGA
SPC147300.63	63	ATC TTT GGA AAT TTT GAG AGA CTC TTT GGA GGC TTT TTC ATT ATC ACC AGC TTG TAA ATA GGC
SPC147800.60	60	TAA TCA TCG CTT TCA TCA CCA AAA TCG AAG TCT TTT TAC ATG GCT ATT CAA ATA CAT TTG
SPC148250.57	57	TTT AAT CCC TAA CCG GAG CGA TAT TGT TCA AGC TAT ACA GAT GCT TCT ATC CTA CAA
SPC45200.60	60	ATT TGA ATA AGC AGA CGA GCC ACA TAC ATT CGT TTG AGG TCA TAT AGG TTA AAC AGT TGC
SPC45850.57	57	CCC AAG TCC CAA CTT CTA AAG CAT GAG CAT TGA GTT TTT CTC TTT CTG CGA TTG TCC
SPC46450.57	57	TTG GAG TTC CAC TAA ATG TTT CAG GTA CTC GTA TGT CTG TCT CAA TCA TTG GCA AAC
SPC47100.60	60	TCA CAA TAG CTA GTG AAC ACA TGT CGG TGT GAT AAT AAC AGT ATG ATG TTG CAA GTA TGT
SPC47650.60	60	AAG CCA TCA ATT ATA CGC TTT CAT CAA TGT CGA ATT CAT CAA GGC ATT TAG CAG GTA GAG
SPC48200.57	57	GGC AGG CTT AAT ATC CTT CAG CGC ATA CCA AAA CTC ATT ATC CTT GTT GTC AAA AGC
SPC48650.63	63	ATG TTA AAT TCA TCT TAA AAT CAG ATG GAT ACG TTT GAA TCT AAA GTG GTT TCG GCA CAA TTC

SPC49150.63	63	TCG AAT AAT TTT ATA CCA ACA ACG CCT CTA TAT TCG AAA TCA TTA AAA AGA TAA GAT GTC AAC
SPC49900.61	61	TGC GTG AGC TAT TCA GGA CAT TAT TGT GTT TCC ATA TAT CCC ATT TGA TAT TAA AAG AAT T
SPC50550.63	63	GTC GTA AGT AGA TTT ACA TGT TTT CGT TTT TCG GTC TTT TCT AAC TAA ACA GAT GTT TAG CTT
SPC51000.63	63	CAA GTG CCT ATG ATC TAT CTT TGC CAG ATA TCA CAA AAG ACA ACA AGA CTA CTT GGA ATT TTC
SPC51550.61	61	TAA ATT TAT CGG CTT CAT TAT GGG GAG ACC CAC GAT GTT AAA ACG GTC GGA CGC CAG TAT T
SPC58450.59	59	AGT CTT TCG AGT TCA TTT CGC TGA TGT TTG AAT CCT TAG ACT GAG AAG ATG CAA ATC TG
SPC58950.61	61	GTG AGC CAA AGA GCA TCA TTA TTA GCG AAG CTA CAA TGA CAA ATT AAC TTC TTC TAC AGC C
SPC59450.57	57	ATG TGC TAA GAT CGG GAA ATG GGT GAG TGA GAA AAG GAA GTT ATG GTA ATT TTA TGA
SPC60050.59	59	GTT GTT TCC CAT CGC TCA CTC TAC ATT CGC TTA AAT AAA CTA CTA CCA TTT GAC TCG TC
SPC60550.57	57	GTT GAA ACT CTT GCA ATG CTG CTA AAC GCC GTT GCC TAT GAG AAT AAT AAG ATT GGC
SPC61000.57	57	TCA TTC GTC TAG TCA TCA CTG CTC GGC TCA CTA ATA TCG TTG GTA AGT AAT AAA GCG
SPC61550.63	63	AAA GTT TCT ATA CTA TTG AAA TTG AAG AAA AGA AAG CAG GCA AAT TTT ACG TGT TGA AGT ATG
SPC62000.63	63	AAT TCT CGA ATT GCT TCT TTT TGA ATG TTC TTC TAG GAT TTA CTC TTG CTT GTG GTA AGA TTG
SPC62500.61	61	ATC GTT TAA CAT CAC GAT TCT AAC CCT ATA CAC CTA ATA CAG TGG GTA AAA CGC AAA GTC T
SPC63100.63	63	ATC TTT GGA AAT TCA ATT TGG AAA GGA ATA TGG ACA AGT AGA TGT TTC ATT TGG TCA TAC AAG
SPC63550.58	58	TCG TGC TTC GGT CCA TTT TAT TAA TCA TGA TGG

		AAA TTT AGT TGA TGC CGC TTG TAT C
SPC64150.63	63	AAA ACA ATG CTT TGA TGG AAT GGG ATG ATA TTT CTC TTC ATT TGA TTA ACA CTT CGC TAT CCT
SPC70850.57	57	AAC GAC AAT GGT GAA TGA AAT GCA AAG AAT TGT GTT ATT TAA CAG CAT CAA AGC CGC
SPC71350.60	60	GCG GAG TAT TTT ACC ATA TAT CAG TCA TTA GTT GTA CGG CTA CTG TAC TAG CTC ACA TGG
SPC71850.57	57	CTG TTC GGG TCG ATC ATT TTG GCT ACT ATC CGT CTG TCT ATG ATA ACC AGT GTA TGC
SPC72500.57	57	CAT CTT CCT TTC CTG AAC ATC GCA ATG TTT CCT CAT CTC TTC TTC CCT TTC CCA TGT
SPC73200.57	57	AGA ATA TAT CGC CCG GCG TTT GAT TTT GTT TCC CAT ATT TTC CTC TTC GAC ACG AGC
SPC73650.57	57	TTC TTA TTG CGG TGG TAC ATG GAG CGG TAT TCG TTC TAA ACT CGA CTA CAT CCA AGG
SPC74250.58	58	TTA CGA TCC CAA TGT CTC TTG TTC AGT GCG TAA TTA TCT CGA CAG CAT TAC CAG CTA T
SPC74750.62	62	GAA AAG GTG ATG TTT TGG GAG TTT ACA CAA ACT ACG AGT CAA GTT CTG ACA ACG TTA CTT AT
SPC75300.58	58	TGC TGT TCG TGT ATC TAT CCT TGG CGT TGC AGC ATT CAT TGC TAT CGT TCT ATT CAT T
SPC75850.59	59	GGA AAC TAT CAA AAG ACC CAA ATG GAG CAT TCG TTT ATT AAA AGC TTA ATT TTA GGC GA
SPC76300.60	60	ATC CAC GTA TCT CGA TAG CTG TGT ACT AAA ATA TCA AAG CAA AAG TAA ACA TCT TTC AGT
SPC77050.63	63	CAG AAA CCA GTA TTT ACA CAC CCT CCT GCA CAA TAG CGT AAG ATA GTA ATA AGA AAG TAT GGC
SPC83850.63	63	AGC GTT CCC ATT TAT GCT AAA GAG CTT TAT CAT TAC AAT GCC TTT CAA TCT GGT AAC TTT CTT
SPC84550.63	63	AAA AAT CAT TTA GAT TTG TGA GAC CTT AAA TCA GAT ATA TTT GTT CTA CGA GGA TTT GGA ATG

SPC85100.63	63	AAT TCT TTA TCA CTC TCG CTA ATG ATA GCT TGT GAC GAT TTA GTT TTT CTA AAA CGG TAA TGA
SPC85650.60	60	TGA TGA ATT TCT TGC ATT TCG TAA AAA TGG AGG AGA GTT AAC GGA TTT TGG ATT TCC AGG
SPC86150.63	63	GAT ATT ATC ATA TTC TTA CCA ATT ATA CAT ACA CTC ACA ATT TGT TTT ACA TTT CCG CTC CCT
SPC86700.58	58	GGT ATC TGA GAG TGC TAG TAC AAT GTG TAA ACG AGT ATT TTC AAA ATC CGA GTC CTT G
SPC87350.57	57	CAG CCA TGT TAC GAA GTT GGT TCA ATT CTC TTT CAT GCT CCG ATT TTG TGA TTT TGA
SPC87950.59	59	CAA GGT TGG ATG GAC CAT CTT GGA AGA TAA TCG CAC TAA GAA GCC TGT AAT TAG GAT TA
SPC88500.59	59	AGT AAC GAT TTG CTT CTT TTG GTA ATC AGA AAA TAC GTT TGA CTC CAC TGA ATT TTT CA
SPC89000.58	58	GAC GGA TAG CAC GAG TCT TCT TTT GAC GGA GAT CAA GAG GAA TGT ACT TCT TGT TCT T
SPC89500.58	58	ATT GCT CAG CCA AGT TCT CTT GCG ATT GCT TAC GAA GTT CAA AGG TCT TTA ATG CCA T
SPC90350.61	61	ATG TAT CAC ATG CAA AAT GAA TGA AAA TGG TTT GTT GGC GAT TTC AGA CAA GCT AAA AGG T
SPC97400.57	57	CTT TGA AGT CAG CAC CCA CAA AGA TTT CTT CTG GGG AAT CAC GTG TTT ATT CCA AAG
SPC97950.57	57	TAC AGT TGC TGT GGG TAA CCT TGA CGT GCT CCC TTA CTA CGA TGA GCT TTA GGA TAC
SPC98350.63	63	AGA GCG GTT AAC TTA TAT TAA CTT CAG ATT TCT CAT AAA ACG GAT TTT CGA AGA CAG TAT GTT
SPC98850.57	57	TCG AGA ACA TTG ACT CCT AGA CAA CCG ATC TAC ATC TTG AAC TGA AGT AAT TGG AGG
SPC99250.57	57	AAC CTT TTT ACG CGC TCA TTT GTA CGA ATC TCC TCT TCA TTC AGG TGA GTT TTG TTC
SPC99850.62	62	TTG TTT GTT ACT ATG TTT TGC TTT TGA GTT ATC

		AAA CCA GTC AGT TGA TCA TTA AAG AAA CA
SPC100350.58	58	AAC ATA CGG TTA GCT TTG AAG CAT CCT CAC TGA ACC CTC CGT AAG TTA TGG CTG AAT T
SPC100900.63	63	ACA TCC TCT GTT ATT AGA AGT ATG AAA ACT ATC ATA TCC AAT AAT GAC TTT GTC AGA TGA CCT
SPC101500.59	59	CCC GTC CAT AAG TTG ATG AAC TTG TTC TCT TTG GAA GTA AGA TGA GCC TGC ATT GTA AA
SPC102000.57	57	ACG GTT TCG GGC AAG GTA GCT CTG CGC GAT AAA AAG GAT CAT AAA TCT CAA ATT CTT
SPC102550.60	60	TTC CAA AAC GCA TAT ACT GAC TTA GGG AAT TAC GGT CCA AAA AAC TAG GAT AAC AAA AAA
SPC103050.57	57	CAC CAT CTA GCT TTA CAC AAT TAA ACC CGG CTT TTC GAA GTC GCC AAT GAA TGA GAT
SPC110100.57	57	CAT CCA ACA ATC ATC GGG GCT ACA GTG CTT AGC AAG TTC TTC AGC AGT AAC GGA AAT
SPC110650.63	63	AAG TTA CAA GCC ACA AAC GTT TAC TTT GTT CTT TAA ACA TCG CTA GTT GTA GTA TAA TTT CGG
SPC111100.63	63	ACA TTT AAA TTA TCC ACG CAA GTT GGT TTC ATT ATA CTC AAG TGT GTT AAA CGT GAA TCA GGG
SPC111700.60	60	AAA TGC GCG AGT GAT AAT ATA ATG TAT TTG TGT TCT TAA CAA CCA ACC TCT TTA AGA CGA
SPC112300.57	57	TCT GTT TAT TCG CTA GAG GAA AAT GGG CTG AAA GTG TAG CTT TCG TCG TAA GCC GCC
SPC112800.57	57	CCA CTC ATC CAT CGA GTT CCT CCA CTC TTT ATC CTC TTC TCT GGC TTT CTT CCA TTC
SPC113300.58	58	TTA ACG ACC GTT CGC AGG GAG GTT GAA CGC CCA AGT CAT TCG CAA ATA GTT CAT TTA A
SPC113850.57	57	TTA AAT TTT GTT ATT CCA ATA TGA GAT TTA TTT TGA GAA ATA CGG GCT AAC AGT GGA
SPC114250.63	63	TGC TTA TCG ATC ACT GAA TAT TTC TCT TCT CGA CAG TTA TAA ATA TTA ATT CAC ATG TTT ACT

SPC115100.59	59	TCC TCG TAA CAT TGC CAG TAA GTA AGA ATT GAT CCT ATT GTT AGC AAC TTT GGC TTG TG
SPC115650.63	63	AAA ATT TTA CCA AAG AGT ACT TGT ATA CTA ATT CTA AAT GCC TTC TGA CAT AAA ACG CCT AGG
SPC116150.57	57	CAA AGG TAA ACC AAC TTC TTT GAG GCC TTG TAT AAT ACC CTC GCC TGG CAC TGT ATG
SPC123200.63	63	AAG AAA TAA ATC ATA TAA AAT AGG AAT AAT TTT TAG CGA ACT CTT TGC ATT AGA ACG AAT GAA
SPC123800.58	58	ATG AAG TGC GCT GCT TTG CCT TTT ACA AAC CAC AAA ATT GGT TGA TTA TAC TTA TTT T
SPC124450.57	57	GAC AGT TAA AAC ATC AGC CTA ATC CTT CTA CCT CCC TTC TTT TAC TTT TCT CAT TTT
SPC125050.59	59	TTT TCA TTG ACA CCA TGG GAT ATG CTT TCA CTT TTT GTG TCG TTA AAA GCT TCC AAA AT
SPC125500.57	57	TAC TGG TCG GTA GCG TTC GAG GTG AAG AAG ACA TGT GTT TTG TGA ATC ATT TGA AGA
SPC126150.63	63	TCT CAC ATA ATA AAT AAC AGT TTG GGG CAT TAA CGT TGT CAA TTA CAG GCT CTT AAT CAC CTC
SPC126700.58	58	TAC TAG TTG TTT AAG GAA AAG CTC ACA TTC CTA TTA CCT ATT CAT TCC AAA GTT CGC G
SPC127300.59	59	CTT AAA GTT CAA AAT CCG AAT AAT GTA CGG GTT TAC ACA GTT TCT GGT GAA GGC GTA AC
SPC127850.62	62	GTG TTA ATG TAA TAG ATA TAA ATC CTA TGC ACC AAT TGT TGG CTT TTG GGA CAG ATG CTG GC
SPC128300.61	61	CTG GTT TAA TTC TTA CAG CTA ATG AAG GTT CTC CTA TGC ACG CAT TCT ATA TTC CAT CGT T
SPC128800.63	63	TTT GAA GTG GAT GAA GAT ACT TTG GAA TAT AAG CAA TTA CAT CCT TCT AGG TCT GAA GCA AGG
SPC129300.63	63	TTT CTG TGT TGT ATA CAT CTT CAT TTA GCT TTA TTT TTG GTT TAG CAA GTT GGT ATT TGG GTT
SPC136100.61	61	GAT AAG AAA TAA CAA ACT TAG CAG GAA CTT TGA

		TCC AGG TGG GAA CTT GCT GAA TAA GAC C
SPC136600.61	61	TTT CAA CTG ATT AAT TAT ATG TGT ACT GCA ACA AAT GGT ATT GGT AGC TAG AGG AAT TGC G
SPC137150.57	57	AAT CAA TTT CCC CTT CGT CCT AAA AAC CGC ACA TAC AGT GAT GCA AAT GGT GAA GCA
SPC137700.61	61	ATC AAC CAA ACG AGT GCG CTG CTG TTG ATG AAG ATT ACT CCT TTA TTT AGT TTA ATC AAT G
SPC138050.63	63	CTC CAA TGA TTA AAC CAA AGC ATC GAA ACA TCT CAT TTA TTT TTA ACT AAC GAC GAG CAG CAA
SPC138700.63	63	GCC TTA ATT TGG TAA AAG GGC TAA ATG CGT TAG TGA CTC GGT GGA GTA CAA TGC GTT TAT TTT
SPC139200.63	63	ATC AAA AAG ATG ATA ATT ATG ATT ACT GTG ATA AGA AAT GGG CAT AAC GCT CAA TGT CTA CGT
SPC139800.57	57	CAT CTT TGT ATC TAT CAT ACA TGA TTA CTT GAC CAC CAT AGC CTT TCG TGG CAG CGA
SPC140300.62	62	CAA ATC AAG CAG AGC TCT CGA TCA TTT ATA GAG AGA ACC GTC AAC TGT AAC GCT ATC TTA TC
SPC140750.63	63	CCA AGC CTC GTT AAT TTC TGT AAA CTG TAT CTG CTC AGC AAT GGA CTT GAC TGC AGC TTA AAT
SPC141400.61	61	TAA CAT TTG TAG AAT CGA AGT GGG TTC GTA TGC TAA ATA AAT TCG TTC ATA ATT TCC ATG T
SPC142000.60	60	TAT TGT GGC TCT CAA AAC ATT GCA TAA GTC TGA ACT TGT CCA GTC AAA GAT AGA AAA GCA
SPC52050.57	57	CTG TTA TAC GAG CAC ACC GTA CTG CAA ATG GCC TTC GAT ACT CGA CAC TAT ACA TAG
SPC52550.61	61	TTA TTG TTA GAA TTC CCT CTT TAT TAT GTG CGA ACC ATA TGT TAG CAT GGA CAG ATT ACT T
SPC53100.62	62	GGT ATA TCA TAA TCG AGG TAT TGA GGA GGT GAA TAT TGG ATG ACT TTT ATC TCA GAA TGC TC
SPC53650.58	58	TTT TTG CGT CAC AGG AAC CTC TAC CAT GAA TTT TAT CGC CTT CAA TGT AGT AGG GTA A

SPC54150.60	60	AGA CGG TTA TTG AAC TCG AGC AGT AAT GGG AGT GTT CAT GAT TTT ATG GTG TAG TAG TAG
SPC54750.59	59	CAA TCG TAT TGA AGC AAA GTT GGA ACC CTT CTG CGA TCT GAA ACT TGC AAT GAA GTA AA
SPC55350.60	60	ACT AGA ATC TTG TAT TTG TAA GTT TGC TAT TGG TTT GGC TAT GTT GAG TGA AGT GCC TGC
SPC55800.61	61	AAT TAA TTT CGG CAA AAT CAA AGA TGT CAA ATG CCC TCA ACG GAA CTA GTC GTC CCA TTT T
SPC56350.61	61	TTA ATA GTT ATT CGA AAA TTT CAT TGA TAA ACC GGC CAA ACC TCA ACT GGC TTT GAA TAC T
SPC56900.57	57	TTT ACA AAC TAA TTT CGC ATA AGC GCA TTA TGA GCT TCC TTG GAA CAT ATG GGG TTC
SPC57350.57	57	AGA ATA GTT AAA GAC AGG ACA CTT AGG AGA ATT GAC GGG AAT TTG CTC GAA ATT AGC
SPC57900.57	57	CCC AAG TCA GAG AAA AGA ATC ATA ACT TGA TGG ATG GAC TCT GCG TTT TGA GAA AGG
SPC64550.63	63	GAA ACC ATC AAA CCA ACC AAA TCC GTA AAG ACT AGT TTA AAA GCA AAT TTG TAA TAT AAT GAA
SPC65050.60	60	CTG CTG TCA AAA GAG TCG AGT AAA CTA AAA TTA AAA CCC ACT AAG TGA ATA GTG CCA AAA
SPC65650.58	58	TTT CCG AAA TCT TAG ATC CTC CAC GAC CTA TTA TAC AAC CAA CCA TAT CTG CGG GAA T
SPC66250.57	57	GTC CAC TAA TAG TTA AAA CAC GAT CAT GAA CAT TAG GAA CAG CCT TGG TAA CGC CGG
SPC66700.58	58	TAG TAT TGG TTA ATT GAA TAG GAC ACA TAT GCG TTT CTA GGA GGA GAG TAG AAC CTT T
SPC67250.57	57	TCG TTG GCC TTT GCT TCG GTA TAT AGG ATA GCG TGG TTC AAG CGA CTT TGT AAG ATC
SPC67750.60	60	AAG TAT CTC AGT ATT CAA CCC AAA GTC ATA AAA ATT CTT GAA CCT TTA GCT TTC CAA GCG
SPC68350.63	63	CAT TTA GAA AAC TCA ATA GAA ATT TAG CAA ATA

		TCG TCA ACC AAC CTA GAT TGA TCC TGT TTC
SPC68900.58	58	GTA GGA ACG GTG ACC GCT GTT TGG TAC TGC ACT TCA TCT TTA TCT ATT GAA AAG GTA T
SPC69300.63	63	TTT TAT AGG TTG TTT GTC TAA TTG ATC ATA ATG GCA TAC CTC TCT ATG TGT TAA TTG AAG GGC
SPC69900.63	63	AAA ACA AGG GTG TAA ATA ATA AAG TAA GGT TGT ACG AAT TCC TTG CCT GTA TAG TGA ATA GTT
SPC70400.58	58	ATT CCA GAT ATC GCT TTG AGA CTT CAT TGA AGA CTT GGC CAT TAA TCC TCC GAT GAA T
SPC77450.59	59	GCA CAC AAT GTT CGT TGT CAT TCT ATG CTA GTT ATT GAG ATA TGT TGG TCA ACT TTG GC
SPC78000.63	63	CTT GAT TCA TCG ATT CCC TCT CTA CTT TAT ACA TTG TTT TAA TTA TGT CTT CCA TCA CTT CAA
SPC78600.57	57	TCA TTA ACA TCA CAG CAA TGT GGT ATC GTC GTG AAG AAC AAC CCA TGC GTA CTC TAT
SPC79100.63	63	TTC ATT TCT GGT GTC TTG TGC AAG GTC TTT AAA AAT GGT CGT CTT TTA ATC GGT GTT TTT ATG
SPC79650.59	59	CAT TTT AAT ATC TCG TTA ATG AAT GAT CCT TAC TCT GCC TAT TGC GGT CAC ACG TTC CC
SPC80200.59	59	TTA GCA CCA TAT TCT CAA TTA AAA ATC CCT ACA AAA CCT GTG ATC TAA ATA CAA CAG CT
SPC80800.61	61	CAA TTG CAC TTG TAT TAC TCA ACA ATC TCA TGT CTG AAA TGT CTT TGA CTA TTG CGT TGC C
SPC81250.59	59	CCT TTC TGG GTG GTA TAT TAG CGA AGG CTT CTA TGC ACT TGA CAG ACC CAA TTT GGA AT
SPC81750.60	60	CCC GTA CAG CCC TAT TTT GTT TTA TAT TCC ATA ATG CAA TTT GGA TTT AGC GTT GGT TCG
SPC82300.58	58	ATA TAT AAT TTC TGT ATT TCA GCT TGG TAT TAC CAT CTA AAA CAA AAC TTG CCG CTA G
SPC82900.57	57	ATA GAG TTG ATC AGC CTC AAA GGC AGC CTT CTC GCT TGA GTA CTG TAG CTA GTA TTT

SPC83300.61	61	TGC CAA TGG ATA TAC GCT CTA CTT TCG ACC ACT GAT TGT TTC CTG CAT TTC TCA AAT AAT A
SPC90900.63	63	CAT TAT GTA TAT AGC AGT TTT CTT TCC ACG ATT GAT ACG GTT TGT CTC ATA ATT AGC CGA ATC
SPC91450.62	62	GTT TTC TAG ATG CAA TTT TCT TCA GCA CAT GCT TTA TAA AAA TAC AAG ACG CGA TTC ATC GT
SPC92000.63	63	CAT TAA TAT CAA GTT TGC TCC CCC TTC CAA CAA AAA GAG AAG ATT AAA TAG AAC ACA TTA TGA
SPC92450.57	57	TTC AAG CTC GGA GCA ATC AGA CGA TTC TTC CTC TTC TTC CTC TTG TAT ATC GGA TTC
SPC93050.62	62	ATA TTA AAA ATC TGT TCG ATA CCT TGT TTA GCA TAC GTT GTA AAT TTC TTG GTG TGC TTC GT
SPC93600.63	63	GCT AAC GGT ACT TAT TAG GAT GAG CCA GTG ATT AAG CAA TTA ACC TCA TAC ATC ATT TCT CGA
SPC94100.58	58	GCT GTT GCT TAT GAC AAA TGA GGA TTT CAA AAA TTT GGT TCC CAA GCT GAT GGA AAT T
SPC94650.63	63	TAT TTT CAA CTT ACA AAG TGC ATA TCT GTC TTC TTC TTA AAG TCG CTT CTA ACG ACA CTT CTT
SPC95200.58	58	GGG TAC TGC TGA CTC TCA AAA TCG CAT GTG GTT ACT TCC AGA GGA GAC TTT GTA TCT T
SPC95750.63	63	TTT AGG ATA TGT GTG GTT AGC TCT CAG GAT ACA TTA TTA CCT ACT ATT TTT GAG ATT GAT GCG
SPC96300.60	60	TTC ATT ACT AGA CAT TAC TAC AAG TAA AGT TGA AGG CTT AGC TTT TAA ACA AAT CCA CAT
SPC96900.58	58	TGT TTC GCA CCT TTT GGA GAA GTT TGC ATT TAA AAC TAC GGA AGA GCG TTC TGC TTT A
SPC103550.59	59	ATG TTT GCA TAA TTA TTC AAA ACG ACA CCT TCA GCA AGA TAT GTA TTA AAT TTT CTT TT
SPC104050.57	57	TCT GGC AGT GTT ACA AGT CCT TGA TTT GAT ACC ATG GGC CTC ATC AAG TAT AAT TCG
SPC104600.58	58	TGT TGT CGA ATC AGA CGA TAA TGG GTT CTT TCA

		TGT GAT GGA ATT GAT TTT GCT TGT G
SPC105150.57	57	CAC GCT CGA CTT GTT GAT TTG AAG CAT TCC CCT TTG AAT CAC CAT TAT CAT CTA GAA
SPC105650.57	57	TAT CCT CAG CAT ACT AAG GAT TTC TTT AAA ATG TCA AGT GGA AGT CGG GTG AGA GGG
SPC106300.60	60	TCA GGA TTT GTG TAT ACG AGT TAT TGC TGA ATA CAT AAA TGA TAT CGA AGC TTT TGG GGA
SPC106800.59	59	ATC TTG ATG ATG AAT GCG TCC GTC TAT TAG CTG GTT GTA GAA ATC TTG TCA GTT TGA AG
SPC107400.63	63	TAG GGC GTG AAG TTC AAT AAG GAT GGA AAT ACT TTC TTT TTC AAT TTT ATC AAA GGT TTT GGT
SPC107900.57	57	TAA TGG TAT TGC TTG CCA GAC TAA AGA GGT TTA CTG CTG CTC GGA AAC GCA AAG CTC
SPC108500.63	63	CCT TAT ATA CTA TCT ATT TTG CAG AGC TAA AGG GAC GAT TGT TAA AGC GGA GAT AAT GTC AGC
SPC109050.62	62	GTT AAA AAC AAC TTT AGT TAT GTT GAA CTT CCT TGG CTG TGT TGA ATA AAG AAT TCG CAG TT
SPC109650.62	62	ATG AAC TCT TTG AAA CTA AAA CAT AGC AAA GCT CTG CGC CCG AAT AAT AAA AAT ATC AGT CA
SPC116650.59	59	ACG ACA TGT GCA GAG ATG CCG ACG AAG CAT AGT TAA ACT GGG ATG GTA AAA TCA ATT AA
SPC117150.62	62	GCT TCT GTC AAA GTT TAA CAA TAT TTC TTT TGG TTT AAA TCA AAT CTT CCA TGC GAT TAA GA
SPC117700.57	57	TTT TTG ACC ATT GAC TAG GAG GAC TTT GAG AAA TGG AGG ATG AAG CTG TCT CCC TGG
SPC118300.63	63	GTG CTT TAT CTC CAT CAT TTC AAA CGC CCT TAA ATA AGG AGC TTA TCT TTT TGA TTC TAA GTT
SPC118900.63	63	TAG TAG TAA TAA ATT TGT GTA AAA GTT AAA GCA AGT GGT ATA AAA TCT TTA GGA GTA ACG GCG
SPC119450.63	63	AAA GTT TGA TTC ATT TTC AAG GTT GTT TAA TTG CTT TCA GAA AAG TCA CTG TAG TTT CCC CTC

SPC119950.57	57	TGA CTG GTA ATT GAT AAG TAC GTA TAC AAA CAC ATT CCA CGT GTC TAT TCA GCA ATT
SPC120300.57	57	CTC CCT TTC TCA TGA TGC AAG TAA CGG AGT TAG ATG TTT AGA ATC TCA TTT TCT TTT
SPC120900.63	63	CTA GGA TCT TTG CTT AAT GTG TTT TGC ATT TCA AGG ATC CCC ACA TAA AGA TTA ATT GCA TGT
SPC121400.59	59	GTG TAA ATG GAA ATA CTA GCT TGG TAG CCG TTC ACG GTA AAG TAT GCT ATA AAT GCC AC
SPC121900.58	58	ATG GCC AAA AAT GTG AGC AAA CGA AGA TAC TAC CAT TGG TCG GAA GTA CAG AGT ATA A
SPC122450.57	57	TGG TAG ATA AGC GAT TCA AGG AGT TTT ATG CCG ATG AAG CCT ATA GTA TAG CTC GAT
SPC129900.63	63	TTT ATT TAT AAT AAT ATT TAT AAT AAG AGA TAT GAG CAG CTC GAA CTC CCG AGT GCA TGA GGA
SPC130450.60	60	ATT AAC GAA TTA CTT CCC CCA GAC CAA AGA AAA TTT GAC GTT GTT TCT CTA CAG TTT TGC
SPC130950.63	63	TGG ACA GGA AAA AGA AGC AGC AGG ATT TTA TTT AGC ATT TGC CTT TGA GAA GAG GGG AAT TTA
SPC131500.60	60	TAG TCT CGA TTG ATT AAA TGA GTC ATC TGC TTT TCG CGT ATT ACC TTT GCA ACG ACA TTC
SPC131950.63	63	CGA TAA TAG GAA ATG AAC TTT ATG GCT TTG GTA CTA ACG TAA ACA ATA CGC TCG GAA TAG CAA
SPC132550.57	57	CTA CTC CAA GGA GAG TTG CTG GCT TAT TAT CGC CTG TCA TAC ATG CCG TTT GTA CTA
SPC133050.61	61	TAC AAA TAT AAT AGA ATT GAA AAT TTA CAA ATG GTT GTT GGA GTC CGT GCG AGC GCT TCT G
SPC133600.60	60	TTT CTC CAT GGC ATC TTG ATT CAC GCT TTT CAC CTT TAA AAG AAA ACC TTT CAA AAC TTG
SPC134150.63	63	TCT AAT GGA TCT TTC ACT TAG TTA TCA CGC TGA GGA GCT TTA TTC ACG ATG TAT TGA TTA TGC
SPC134650.57	57	ACA ATA TAG AGT CTG CTT CGT CAG GAG GGG ATA

		AAA CAC AAT TAA ATG GTC CTG GGG
SPC135100.59	59	CAT CTC AAG TAA TTT ATC AAG AAT CGA AGA AGC GCA TTT CCA ATG GTT CTC CAA CTA GT
SPC135650.63	63	ATT TGC TAC ATG GTG AAT TAT TTA TTA AGC TTT TGT TTC ACG ATA CCA ATC ACA CAT TTT ACA
SPC142550.63	63	CGT ATC GCC AAA GTT GAC TTA AAA ATT CCA AGC TTT GTT CCT CCT GAT GCA CGG GAT CTT ATT
SPC143150.63	63	ACG GGA TAT TAT AAA ATG GGA GGG GAA GGG AGT TGA CAT AAT AAC TAT CTG ACT TTT GTA AAC
SPC143650.60	60	TTA CCG GGA ATG TAA CCT CCA GTA CTT TGT GCA TCA CAA GGC TGG CAA GAA ATG TCT TTA
SPC144150.58	58	ATT GCA AGC TCT TTC ACA TAA AAC AGT ACA ACG CAA AAG TTT TGC AGT TTA TAA CGT T
SPC144850.57	57	GAA AAT GGC TAC ATC TGT AGC TCG CCA AAA TCT GAT GTC GTT GGG TAT TTG CTT TGG
SPC145450.60	60	AAA GAG GTT TTG GGC TTA ACT GCA ACC GTA TTG GTT TGG CAT TTT GGG TTG TGT GGT TTT
SPC146000.63	63	AAA GGA TTA ATT TGA GAA ATT GAA GCG AGA TAA CGA AGT AAT GTG ACT AGG TCA TGA GTA AGG
SPC146500.59	59	ATT CAA TTT CCT TAT CCA GAT AGC AAA GAG CAG CAA CTT TAC AAG AAA CTA ATT CGC TT
SPC147050.62	62	TTT TCT CAC TGA ACG CTG CTT TCC CAC AGC CAT GTC CAG CAA GGC CAA ATT TCG AAT AAA TT
SPC147550.57	57	AAC TGC AAT CCA GGA AGC TGT GAA GCA GCA GCA AGC ATA TTC ACT CGA ATA TCA GCG
SPC148050.60	60	TAT CAC TTT TTG TAT CCC CAA CTT CAA TTT TTC CAA TAC GCT GAG CAA GTT CCT CAA TCT
SPC148550.63	63	TAA GTC ATT TTG ACC TAC ACG CTA GTT GTT TTG CAA CGT GTT TTT AGA CAA AAT AGA CTC ATT
SPC148750.62	62	AAA AGA AAA TAA ATG AGT AGC ACT TAC AGA GAC ACG AAT GAC GAA CAG TTT TTC AAT TAG TG

SPC149250.59	59	TTT AAA TGC TCT GCA ACC AGG TCA GCC ATT CCT TGA ATA GCT GTC ATA CTT CCA TTA AT
SPC149750.63	63	CCA TTC GAA ATC TTT TTC CAT TCC CTT TTC AAT AAG TTT TCT TCG TAA TTC ATT GAG AGA AGC
SPC150300.57	57	ACC CAT GTT CAT CAT GAC AAC AGC CTG CAC ACC TCT GTT AGT AAT ATA ATA ATT AAA
SPC150850.62	62	AGG TCT TAA TGC AAC TGA TTC GCG TTA CTC GTA TTA TAG CTG TAT AAG TAC TAC AAT AGG TT
SPC151250.63	63	ATA AAA TGG GAG TCT TGA AAA AGA GGA AGT CAA GAA ATG GCA AAG AAC TAT CAG GAT TTA CGA
SPC151800.60	60	TGA ATA ATG CTG ATA ATA GAA CCG TAG AGA CTT AAT TGA CCG ATG ACT ACG TAA AGC GGC
SPC152400.63	63	CAA AAT CTT TTA TCT AGT AGA AAA TAA AGG GGT TTT GCC TTT GGA GTT GAC TCC ACC TCT TCC
SPC152900.57	57	CAA TAA AAA AGT GAG AAG GAG AGT TCA ACA TCG CAT GGA AAA CCT CTT TCA GTG CTG
SPC153300.63	63	AGA GAA ATT CAA TGA AGA TGG GAT GAA TGT TAT GGA CTA GTA TTA AAT ATA AGA ATC TCG ACG
SPC153850.61	61	AAC TTT TTA GAA ATT AGC TCC GAT TCA GTT GGT TAA ATC TAG GCT CTT TAG TTT GTC TTT T
SPC154350.63	63	CTT GAA TAT GTA AAC ATT ACA ATA AAC GGA ACC AAA AGT CAA AGC ATG CAA AGC GAA GTG AAT
SPC160950.60	60	AAT AAC CGT ATC TAT GAA AGC AAG GAA TGA ACA CAT TAC TCC AGT TTG TAA ACT TTG GGT
SPC161450.57	57	GGA ACA GCC CGG CAG CTA AAT TGA TAA TTC CTG CGA TAA AGG CAA TCA ATC CAG TTA
SPC161950.57	57	AAG CTA GTT TTA GGG CCC AAA AGA TGG TTT TCT GAA GGA GAC GAC ATT GCT ACA AAC
SPC162450.57	57	CAA GAG AAA ACA CTG AAA AAC AGT ATG AAG CAG TCG AAA CAA GTG CTT TCA TGA ATT
SPC163050.62	62	TGA GTT GAA GCC ATT CAT TTT ACA ATG AAA GAA

		GTT AGA GTA GTA ATT TGC GGT GAC CAA GG
SPC163550.57	57	GTT CTT TGC GAA AAT AAA TCA GAG GAT CTG GAC AAC TAT CAG GGA CTT CAT ACA ATT
SPC164100.63	63	TTT ATG AAG GAC AGT TGA CTT TGC CTG GGT TTT TAG CTT ATA ATC GTG TAC AAG TTG AAA ATG
SPC164550.57	57	GAT TAT AAA ACA ACA TTG GCT TAT TTG GCG TAC CTT GGG TTT GAC ACT GAT GGA CGT
SPC165100.57	57	TTA GAA TTT ATA CCC GGA TCT TCA AAA GAT TCC CTG CGT TTT CGC TGC TAC CAA GGC
SPC165650.62	62	ATA TAT TGC CAA GTA TGC TAG AAT CGC ATT ATG CAT TAT TGG TTA TTT CGT AAA ATT AGG GC
SPC166200.59	59	CAT TTT GAA GTT CGG GCA AAT CAA AGA CAC CAA TTA AAA CGC ATG TCG CTA AAC TAA AG
SPC166700.60	60	GCT TCG TCA CTG TCC ATG TCT TCA TTA ATT GAC TGC AAA GAT AAA TGC GAT TGA ATA TGA
SPC173900.58	58	ACA CAC ACT CTC TTG GTA AAC TGT TAA TCT AGA AAA CCG GGA AGT TTC CGT TCT TTT C
SPC174400.57	57	CAA CAA CGA TAA CTC CTA TGG TGG AAA CAA CAA CAA TTC TTC CTA TGG CAG CAA TGA
SPC174900.57	57	AGC ATG GTA AGC ACC ATA AGG ACG ACA ACT CCT ATG GAA GCA ATG ATA ACT CCT ACG
SPC175300.63	63	AAC AAT TAA ACG TCA ACG ACA ATT CGT CGA ACA ATA ATT CAT CTG GCA ATA CAG ATA GTT CCA
SPC175750.57	57	CTA ACA ACA ATT CCA ATA CAT CCA ACA ACA ATT CCA ATA CAT CCA ACA ACG AGT CCA
SPC176250.61	61	TCG TTC TCA TTT TCT TTG GTA AAC AGG TTG GTG AGC GCA TTC TAG TTC TTT TGA TAG TTA G
SPC176800.58	58	CCG CTG AAA GTT GGA ATA AGC AAT CGA TAC TTT TAT CTC GGC TGA TCG CAA ATT AAA A
SPC177250.57	57	GGA AAA GGC TAA AAT TGC AAA GGC TAA AGG AAG CAC TAG ATT TTG TAT GGG AAG TGC

SPC177700.58	58	CCC TGT AGG AGA CGC TGT AAA AGA GAG ACT TCC GAT TCA TCC TTT CCT TCG CTC CAT T
SPC178350.61	61	TAG AAC CCT TTG AAT TTA CTT GTG CTA ATG TGC TGT TCA TTC GAT AAA TCA ACC GTT AGG T
SPC178800.59	59	ACT GAC TTT TAG AAA ATT GGT CTT CTG AAA AAC GTT GGT AGA GCG ACT TTA TTG CAT TT
SPC179300.62	62	AAC AAA ATA GAT GAG GAG GAT TTT ACC TGG GAA CTA AAG GAT TCG AGT TTA GAA CTA GAT CC
SPC186250.57	57	AGC GTG CAG ATG AGC GCG AGT TTG AAC GAA CGT AGC TGG GTA AAA CAT TAT TGA CGG
SPC186750.57	57	CTT TGA GGA TTT CCG TAT GAA GAA GGC AAA CCT GTT TGC GAA GAA CCA TAA ACT TCC
SPC187500.63	63	ACT TCA ACA GTT TGA TTT CAA GCT AAA ATG GTA AAT TTT AAT GCA TAA AAT ATA CCC ATT GAG
SPC187900.63	63	CAG GTC GCA CGT ATA AGT CTA CCG TAT GTT ATA AAC AAA GAT AAT AAA TTT TAA TAA GAT GAC
SPC188350.62	62	TAG TGT CGC ATT ATT TAC ATT ACC ACT TAT AAC TTA TTT TTG GAC GTT GAA GAC TTT ATT CA
SPC188950.63	63	CAC GAA TTA GCA AAA TGC AAT TGA TAT GGC AAT TTC AAC CAA TTT TAT AAT TAA ATT TTC CTT
SPC189450.61	61	AGG ATT CAT GGA AAA GAG AGA TGA TAC CTT CAA TGC TGA TCA AGT GGA ATA TCT TCC TTC C
SPC189900.58	58	TTT GGA CCG TAA TAA TGG ACC TTC TTA TTC TCT TCT AGC ATC TCT GGC CTC TCA GCT T
SPC190550.58	58	CTG TAT ATA TGT AAA ACA ACT TCG TCC GAA GCG CTA CAA CAG CAT GCA AAC GTT GAA A
SPC191050.60	60	AAC AAG AAA ATC TTT GGC CGA GGA ATG GTT CAT GAA TAT ACT TTT AAA AAT CGA AGC TCC
SPC191550.62	62	TTA AAG ATA TCC CAT GAT AAT TTT TCA GAC TCC AAC GAT TTC TCG TTC CAA AAG AAA CTG AA
SPC192100.63	63	TAG TAT CGA TTT TGA AAG TAA TTT CTT TTG CTT

		TAT TTA ATC ATT AAA AGA TGC CGT AAA GGA
SPC198900.57	57	CTC ACG TTA TGC AAT TCC TGA AGA AGC TTC ACT CTT TTC GTT TGA GCA TCA TGC AGT
SPC199350.63	63	AAA CGT TTC TAT TTA CTG ACA GAT TGA TTA ATA TAT TTA CGG ACT GGT TGT AAA GGC AGG CAG
SPC200000.61	61	CAA CTG ATT GAG GTA TAA ACA TTA AGA GAC TGT AGC TAA TCT CCG ATT TTT CTG CTC CAC T
SPC200550.59	59	TTT CCG CTG ATC GTG GTT ATC TCT GTT GTC AAT GCA GTT ATC GAG ATG ACT TTC TGT AT
SPC201050.63	63	ATT GAT TCG AAG GAA GGG TAT AAA TAA GCA AAC CTG AGA AAT ATG TGT TAT CGG TTT ATT ACT
SPC201750.58	58	TTC AGA ATA GCT TTC ACA TAT TTT ACA CCA GAA CGC CAA ATG CTT GTA ATT CTA TTT T
SPC202300.58	58	CTT AGA GCA AGT TTG GTC ATC AAG GTG AGA TTC GAA ATT TAG TTG CTT GTT TAC ATT T
SPC202850.63	63	ATA TGT ATA GAA CTG TAT CAC CAA AAG AAT CAT ATA ATA GTC ATG CTT TAA GCC TCT TCA GTC
SPC203350.63	63	TTA AAA CAA TCA TTA TTT GCT AAC AAC TTC TGT CTT CAG TTT GCT ACC TGA TAT CGT TTA ACA
SPC203850.58	58	TCC CTC CCT ATT TGA GTA TTG CAG CCG TTT GAT AAT CAT TTA AAA TAC GAT TCC TTC G
SPC204350.63	63	GCG GTC CTC ACC TCG GTA CTT TTC TTC TTT CTC TAT ATT ATA TCT TTT GAG ATT TTC TAT ACA
SPC204850.57	57	TAC GCA GAC AGG TAT GAT CCT GTC ACA GAT ACG TAT TCG TAT TCT TCT GCT CGC CAA
SPC211900.63	63	AAA TTT CCA TTC ATT CTC TCT CAT TCC TTT TGT TTG AGT CTT GTA TTT TGC GTA TTC GCT CTC
SPC212400.57	57	ATC ACC AAC ATG AGA GAT TTC CAA TCT CGT TTC GCA GAT CGG TAT AAT CCG GTA ACA
SPC212850.58	58	GTA TCC ATG GGT ATT AAC TTG CTT TGG GGT ATC ATC ACC TTC ATC GGT ATT TCT TTC C

SPC213300.57	57	CCA TTG ATT ATT GGT GCT CTT TGG CAA TCC ATT ACT TTC TTC ATT TAC GCA GCT GTT
SPC213850.57	57	AAG GTA TTG AGG ACT CTA GCA ATG ACA TTA GTT CCA CAA CAT CTT CGG ATG GCC GTG
SPC214300.63	63	TAA AAG TCT CTT TTC CCT TGC TGT CTT TTT ACA ATT TAT TCC AAT GGT CGT TAC TGT ATT GAA
SPC214900.63	63	AAA TGT TAA AAG TAA GAA TGT ACA CTG CCT TCT AAG TTT AAA CCA CCA CTC TCG TCT ATC ATT
SPC215350.63	63	CCG CCA ACC ACA ATT TCA AGG AAA TCT TAG AAA ACT GTT CGT TGA AAA AGT TAA CAC GAT TAT
SPC215850.63	63	TTT TTA ATT GTT TAT TTG ATT TGG TTA ATA TAT GAC GAT GGG TTT GTA ACA GGA AAA GAT CGG
SPC216300.57	57	GTA TCC AAG AAA TCC TGT GTT AGC ATT CTT TCC CGT AAG CCA GGT GAT AGT GTT TTG
SPC216800.63	63	GAA ATC TAT GAG GTT CAT GAA GAG TAC ATC CGA AGA CTT GAA GGC TTA TGG AAC AAA TAC AAA
SPC217300.63	63	TAA TTC AGT AAG TCT TAG CGA TCG TTT ACA TTT GAT GAG TTG ACA GGG TTT GTA AAT TCA AAC
SPC225800.63	63	TAC TTT ATA TAT TTC TTG TAG ATG GCA AAT ACT CTG CGT ATT GCA ATG CAG CGG TTT ACA ACT
SPC226250.60	60	TAG ATA CTA CGC TTA TTT ACT GTT CAG CAT ACC TTT CTA CTC AAT GTC TAC AAT TCG CAA
SPC226750.57	57	AAC AAA CTT ATG CGT AGC GAT CTT GTT CAT CAA CAT AGT TAG ACT CAG GCT TAG CAC
SPC227300.57	57	AAC GGC ACC GGC ACG ATG GTT TGA GGT ACC ATT CTT ACG AGT AAG AGC ACG ATT ACC
SPC227800.57	57	CAG GAA GGA AAG AAA TAC CGA TGA ACA TGA TAA TAC CCC AAA GCA AGT TAA TAC CCA
SPC228350.63	63	TCC GGC CAT GGA CAC GAA GAC CAA CAT GAC AAT GGT CAA TGT CTT ACC CAT TTT AAG TAA ATT
SPC228750.62	62	AAG CCG AAG GAA CTG AAA CCC CGT CGA ACA

		AAA GTT AAA TTA ACC GTA AAC CAC AGA TTT TA
SPC229250.62	62	CGA ACC TCT GCT ACT TTC TAC CAA ATG TCA CAT TGA TAG CAA AAC TGT CCA TAT TGT ACT TG
SPC229750.58	58	ATC TAT TAG ACG CAT CCT TAC AAA ATC TTA TCG TAG TTC GTA GGG AGG CGA TAT AAG A
SPC230250.60	60	TTA CCG TGT ATT GCC AGA CGT GTT AAA TCC CTT CAT GAT AAG AGT AAA GAC CAA AAG TTG
SPC230800.58	58	GGA AAA GTA GTG CGG TAG GAA TCT CTC TAT TCG TTG ATT AAG GAG GGG AAG AGG GGT A
SPC231250.57	57	CAG GTA TAG AGA AGG AAG TAG GAA GGA AAG GAG AAG AAT AGC AAA GGG AAA TGA ATG
SPC40109.60	60	GCT ATT GTC AAG TGA TGC GTC CCT TCC TGA GAC ATT TCC TGT AGA ATT CCC CTC AGG GCC
SPC40229.60	60	GCT TCT ATA CCA AAC AAC TGC CTC CCG TGA CTG ACT TCA AAT TGC ATT CTT GAC GCA AAT
SPC40349.60	60	AGA CCT GAG ACA AGG TCT TCA GAC CAA AAA TAG TCG CAA AAT GAT AGC ATG CTC GCT TGC
SPC40469.60	60	TTC CAT CTA TTA TTG TCA TTA CGT CTT GGT ATC ACT ATA CAC GAG TAG AGA AAC AGT ATG
SPC40589.60	60	ATT ATA CCA ATA AAA TTA ACA AAT AGT TAG CAT CAG GGA ATT TGT ATT TTA TGA TGA TGT
SPM451.63	63	AGT GAT AGT GAA GAG TAC CGT AAG GGA AAG CAA TGA AAT AGT TAA TCT ATA AGC GAA GAT AGT
SPM2401.62	62	TCT ATT TAA AAG GTT ACG TGA GTT GGG TTA AAT CCG TCG TAA GAC AGG ATG GTT CCT ATC TA
SPM4451.60	60	CAT TTG TTT TCC CAA GGT GTT GTG CAA TTA GTG TTA AGT CGA AAT AAG GTA ACC GTA GTG
SPM6401.59	59	AAA AAT AAA TAT TCC CGA GGA TCA CAA CAT CCA AAG GAT AAA TAA AAT AAG TTC AAG GT
SPM8051.62	62	TGG TAT GGT CTC ATC ACT TAT TCA CAG TTG GTT TAG ATG TTG ATA CTA GAG CTT ATT TCA GT

SPM10201.63	63	TAG CTT TGG CTA ATA ATT ATA TGA TTG ATG CTC CTG AAC CTT CAA ATA TTT CAT ACT TCT GGA
SPM12201.61	61	GAA TTT GAT TAC AAA GGT CCA ATA GCT CGT AAA AGA ACT TCT GAA TCT AGA CAT CTT CAT T
SPC154800.63	63	TAC ATA TAA AGT GCA TCG CAT ACG CTT GAA ATT AAC TTT TGA TTC AGA AAT CTT CTC GTG TCT
SPC155250.61	61	CGA TTG CCG ATT AAA GCA TTC CGA CAT TTT TGT TTT CAT CAG ATT ATT CAA GTA ACG AGT T
SPC155700.60	60	AAC GAA AGA ATT TAC TGT TTA ATA TTC TGT TGA TTC CTT AAC TTC TTG GAG GAT CTA TTT
SPC156150.58	58	CTT TTC TGC TTA ACC CAG TAG GAT CAA GTC CTG GTA AAT TGC CCG TGC CTT CAA GGA A
SPC156750.63	63	CAA ACA ATT TCA CCA TTA TTC AAA GGG ATG TTG TAA GGC AAC AAG AAA ATG AAT ATC ACA AGG
SPC157250.57	57	CAG CCG AAT AAC TTT GCA CAA ACG TCG TTC GAT TAC CAA CCT AAT CAT CCA AAT GCA
SPC157800.63	63	GTT TCA AAA TTA AGG GAA ACA GTC CTA TGT GTT TTG TAG AAT TTG AGG AAG TTT GCC ATG CTG
SPC158300.61	61	TTT TAA CAT TCA GAA TCA GTT GCA TGA CAG AAT TAT TTT ATC CTT TGT GAA AAA TTC CAC A
SPC158750.62	62	CGA AGG CTG TTG CTT TCG GAT TGC TAG GAC AAT CAA TTA ACA TAT AAA TTC GAT AGC TTC TC
SPC159300.61	61	AAG GTT CCT TCT TTT CCA TGA ATT GTC ATT GTC AAT ACT TTT AAC AAA ATC TAA AAC ACG C
SPC159800.63	63	ACA TAG ATA TAT AGT TCT ATG GGA AAT AAA TAG AAA GCA AAG TAT AAG CAT TTT GCC ATT GAT
SPC160400.58	58	AAG AAG GAT TAG TTC GAT GGG ATA AAG TGG AAG CTG CAT CTT CAA GGA TAA ATT CTG C
SPC167200.63	63	AGC ATC TGA AGC TTG AGT ACT TTT GGA ACG CAT CGA AAT TGA CTC AGC ATC TGA TGT ATA ATT
SPC167850.57	57	GAC TAG GTA TTG AAG GAG CTA ACT GGG GAG

		AAT TTT GCG CGA ATT GGG CAT TAT ATG
SPC168450.58	58	AAC GTT CAT TGA CTG TAG GTG CAG ATG AAA CAT TGA CAG AAT TGA TGG AAT GAA AAG G
SPC169000.63	63	TTC TAG AAA AGA ATC CTA AAA TTC GAT CAA TAA ATT AAC GAA CCT GAT TTA AGG GGA AGG GAA
SPC169500.57	57	AAT CGT TGT GTT GTG GAG AGA CAG TCA GCA ATT TTA CCC AAG GAA ATC GCT ACG CTA
SPC170050.63	63	ATG CTA TTC GAG ACG CAG TAA AGG TAT GGA AAA CGT ATG ATT AAT CCA AAT ATC TAC AAT CTT
SPC170550.57	57	CAA AGC ACT AAA CTA CAC AGG CGA AAA CGG TAA CAA CAA CGT AAA TAA TGA ATT TAC
SPC171100.57	57	AAA GCG AAT GTC AAA ACC CCA GAA AAT AGT CAG GAG ACT GCA TCA AGA CAT GAT TCG
SPC171650.63	63	TTA AAA TTT GCA TGG TGC TGG TTT ATT AAT TGA TAT AGG ATG AAG TGT ACA GCG TAA AGG AAA
SPC172350.57	57	CGT ATT TGC CAT GGG AAA CGA ATC TTC TTC TTG TCT GCT TTC CCT CTT TCA AGA AAT
SPC172900.62	62	AAC TGG CGT AAA TCG TAA CAT TTT AAG AAG TAT CAA ATA AAT TAG TAT CTC CGA TGT TGC GA
SPC173400.63	63	ACT CAT TGT TGT ACA CAA ATC TTC GTA CTT TTG TTA CTA AGT GTC CCT TTC CTT ACT ATG TGC
SPC179850.57	57	AAG AAG CAA CTC AAC TGG CGA AAG ATA CTG GTC TTT CAC TTA AAA TTT GCA CAG ACG
SPC180350.63	63	ATA GTC CGT GAA CGT GTT GCT GCT TGG TGG GCT AAA GAG GCA GAT ACT TCT CCT AAT AGT ATT
SPC180850.59	59	GTA ATT GAT GTA GGT TTC ATC ATG GAA CAG AAA ACA ATA GTC GTA AAG TTA CAG AAC GT
SPC181400.62	62	TTG GTT GTA TTC TCA TAA ATT TAG GCC AAC CTT TGT CTG TCA CTT CAA AAG CTA TAC GAT GC
SPC181900.60	60	TCT CTG AAG GTG CTG AAT GAG TGA AAA TGT TTT TTA TAC TTT GAT ATA CAC GTT CCA ACT

SPC182450.57	57	GTT TGC AAT TAA GTC TGA GAA TCA TGG TTT AGA ATA CTA GTT AGT TCC GCG TTC TTG
SPC182900.61	61	ATA GCA ACG GTA TAT AGG ACA ATA GAC CAT GAG CAT CGC CAC CAA CAA TAT TCT AGG ATA T
SPC183500.57	57	CTC CGC ATC CGT TTG TTT GGA ATT ATT ACT GCT GCT GAC CTT GAA CTC TCC ATC ACT
SPC184050.63	63	ATC CTG CTT CCT TCT TCT TTC ATT TTA GGA AGT AAT TCT TTC AAC ACT CGA ATT TTA CCT GAA
SPC184700.58	58	GAA AGC GAG CAA GGA AAC TAA TTC TTT GAG ATT GTT TTG TAA AGG AGT ACC AGT CAG C
SPC185150.59	59	CTC ATC AGC CAA AAT CCC TGA CAG TTT CTG TTG ATA AAG CAA ATG TAA CCA ATT AAC GC
SPC185650.58	58	TAC TGC TGT TTA TAA AAT TGA GTA CAG TTC TCT CTA ACG CCG ATG TAT CCC TGT GAA C
SPC192600.63	63	GAA AGG AAC AAA TAA ATA AGG TAT TTT GAC AAT GGA CAG AGA AAT TTG TTT GAA GAT TCT TCC
SPC193050.57	57	CCA AAT GAT ATT GCA GAA AAC AAT ACC TCT AGG ATA AGA AGC CTT TTC GTA ACC GAA
SPC193600.60	60	GGT TTA TCA AGT AAC TAT GAT CGG TTC GGT TGA GGA GAT AAA CGC ATT CAA AGT TAG AGT
SPC194100.63	63	AAA CTT CTA AAA GTG AAT TTT GAT TGA AAT TAA AAG TCG AAG AGA ACT ATT AGC ATT GCC GTT
SPC194600.57	57	TGC GGT ATG TTA AAT AGT TGG TTT TTG TGA ATC ATA GTC GCT TCA TCT CTC CTT TTT
SPC195150.57	57	GAC CTA TAG CCG TGT ACA ACA GTT CAC CTG TCA CGG AAA TAA CAG AAA TAA AAC AGG
SPC195650.59	59	CGG CGG ATG GAA CTA TTA CAC TCT TTG ATC AAC TGA AGC AAA CAC AAT ATT CAA TTG AC
SPC196150.57	57	GCA TAC AGC TCC TGT TAG TGC TTT GAG CTT CAG TCA TAA CGG TCG CTA TCT TGC TAC
SPC196700.63	63	GTG AAC GCA AAT TAG TGG CTT CGT CTA ACA TAA

		ATC AAA AGG AAC ATT ACT ATA GGT AAA GAT
SPC197250.63	63	GAC ATC GAA TAA ACC GGG AAC CAT ATG TTC GAT ACA ATC AAC GAT ATA ATC CCT AAA GGT GTC
SPC197800.57	57	CTA TCT AGT TCA GGT TCA CTT TCA CGT TTC ATA AAT GTG TCA ACG CTG CTC TGG AAC
SPC198350.57	57	TGG ATT TAG CAT CGC GTA TAC TTT TCA ACA CTT CAT CAG AAG ATT CTT CAG CCT CGA
SPC205350.63	63	TTA CCT TGC AGT TAA AGG AAG AAA TGA AGA ATG TAT GAA AAT TCT TAC TCG TAA TGC TGG CCT
SPC205900.57	57	TAC GTT ATC GTT GGT GAA TCG TAT CCC ATT CGT TAT CGT TCC AAG TGT GCT GCC GTC
SPC206400.63	63	TTT CAA GTC CCA ATA TGA AGA CAC TTA GAT TTT AAG TCC AGA TAA AAT TGG ATT TCC GCC GAA
SPC207350.58	58	ATG AAT AAC ATG CGT AAA TAA TAC GGA TTG GAG AGT GAA TGA ACA AGT ATC TGC ATT T
SPC207850.57	57	GAA AGG AAA AAG AAA AAG GCA CGC GAT GCT AAT GAG TAG CAC AGA TTT TGA AAG AAG
SPC208300.61	61	AAC GAA ACG AAA CAA AAC GAG AAA ATA ATC TTT CAT TTG ATG TTT GCG AAT CAT TAG CGT T
SPC208800.59	59	CCG ATA ATA ATT TTC ACC AAT GAC ATA AAT TTG CGA CCA ACA GAA TAG AAG CCT GTG CA
SPC209250.57	57	TAT TAT GTG TTT TAT TTT TGT TTA TGT ATG TGT GTA TGG AAA AGG TTT GAT GGG AAT
SPC209800.60	60	GGA TGG TTT CAT AGC TGC ATC ACA CAT TGC AAA AGT AGA ACT TGT TTT CAA ACT CTA GTG
SPC210300.63	63	GTT TAG TTA AAA CAT TGT TAG TAC ATA GTT AGT AGA TAG CTA GGA TCA CAG GGA GAA GCA ATT
SPC210800.57	57	TGA TAA GCC ATG ACA CCA AGA AAT CCT AGT AAT TAT GAT TCC CTC TTC GAC TCA TCG
SPC211350.57	57	CAT CCA TGC GTC GTT ACT TGT CCG TTC GGT AAA ACT GTC AGC AAG GCA ATT TTG TGG

SPC217850.60	60	GTA GGG AGT CGA TCT CAG ACT GCA AAT CGC TAA ATT TCA AAA TTT GAT AAG GGT AGA GTC
SPC218350.58	58	TCA TTA ACA ATA CTA CCC TAT GTA CGT TAC GTT GTC ACC GAT AGA TGT AAT GTT TGC A
SPC219300.57	57	TTT TAT CAT CTT TAC CGC AAG TTG TTG CAA GTC TGG TCA GTA AAT GCG TAG AGC AAG
SPC221050.61	61	AGT TAA GCT TAT ATC TCA CAA ACT TAT AAG GCA AGA GAA TAA TTA AAG TAT GTT GGT GCT T
SPC221600.58	58	CTT TCG TTA CAA TAA CGA CAC AAC TTT GTA CAC CCA GCA CTT GAA ACA AAC CAA AAA G
SPC222200.63	63	ATA TTA GTA GTC TCA GAA GGG AAT ATA TAT TTG GAT GTG AAT CTT TGA ATT CTT GTC TTC TTT
SPC222700.63	63	ATT TCT TAG AAA CAA TCT ATT ATT GAA TGA TAT ACA CAC AAA TAG TTC AAT TGT TGT GAT TGC
SPC223300.62	62	TCA CTC TTT AAC AAC TAC AGT GGA ATA TAC AGT TAG AAC GTA CAT TTT GCG TTG CTT TCA CG
SPC223750.62	62	TAA CCT CGA AGA AGG TGA CGA AGT TGT ATT CGT CTC AAC GTT ACA CGG TAA AAC AAT ATT TA
SPC224250.60	60	CCT TCA TCC TCA TCT GAT GAG AAT GGA TGT TGC CTG GAA TGA ATA TTG TTA TGA GAA GGA
SPC224750.58	58	CGA GAT ATT CCC ACC ACG GGT GTT ATC GCC ATT TAG GCT GGA ATC ACT CTC AGC ATT T
SPC225300.57	57	TTA AGC ATC AAA TTT TCA CTT ACC TCG TAA CAA TAT GTA TGA CCG CAA TGA GTG GTA
SPC231700.57	57	TAA AAG TCC ATA AAA CGC TCA TTA TGA GTC GTG GAT TCA GAG AGG CTG CTT CTT TCC
SPC232400.58	58	AAT TGA GAG AAA TAA TAG AAC GAG ATT TTC GCG TAT TTC GAT GGA TTG ATC GGC CAG C
SPC232950.57	57	TGG AAG AGA GGT CTT CAG AAT CGG AAT TGG AAT CAG CGT GCT CCA AAT ATC TGG CTG
SPC233450.57	57	AGA GCA CGA TTA CCA ACA GCT GCA TAG ATG AAG

		AAA CAA ATG GAT TGC CAA ACA CCA
SPC233950.57	57	CAA GTT AAT ACC GAT GGA TAC ACG CCA TTG AGC AGT TTT ATA AAG TTT GTG GGT ACC
SPC234550.63	63	AAG ACC AAC AAC GAT GTT GTG AAG TTA AAG GAT ATT TTG TTT TTA AGA CTT TGT AGG CTC TGC
SPC235150.57	57	GCG AAG GGT AAG TTT AGG AAA GAT CAG AGA ATG TAG CTC GAA AGA TGC CTT TAC CTT
SPC235600.60	60	AGT AGA AAG TTA GTT AAT CGG TTG AGT GGG ATG AAT GGG ATT TGA GAG AAA GGT TAT GGG
SPC236100.57	57	CAT TAC TTC CAC TCG GTT GGC AGG ACG AAT GAT CGT TGG CTA AGC GAT CGT TTA GGT
SPC236600.59	59	CGC CGG GTT TCC TTC GGG TTC CTT TCC TTT GGC TTT TAC AAC ATT CTC GTC TTG ACA TC
SPC237100.60	60	TTT CTT AAT CTT TTT GCT AGT TTT TCT TTC TAT GTG AAG GGA GCG GAC GAG TGT CCT TTC
SPC237600.57	57	TTG ACG TGA AAG AAG ACC CTT TCT TAG TTA CTT GGC AGT CTC CAA CCG ATC CTA AAA
SPM14301.63	63	CAG ATT CGT TTA AGC AGA TAT AAG TCA ATT GGC AGA CCT TCT GAT TTC CAC TTA GAT TAT GTT
SPM16601.63	63	AAT TAA GGT TAG ATT GTC TTA TTC TTC AAA TCT TTA ATT GTT TCC CCC TTC AAT CAT TTA AAG
SPM18651.63	63	ACA CAC TTA AAT GAT TAT CTT ATG TTC TAT CTA ACT TTC ATC TTT ATC GGT GTA ATC TAT GCT
SPA4665000.57	57	CCA AAA TTA CGT GAA CGA TTA GAC GAC CCT GAC ACT TCC GTT GTA AAT GCT GCT GTC
SPA4665550.63	63	AAA ATA TCC GTA CTA TCG TTA AAA CTT TAA TGC TCC AAT TAA TCG TTT CGT CGG ATG AAT CAG
SPA4666200.61	61	TTT AAT CAA GTC AGT GAC ATT GTT AAC ACC TCT GAT ACT ATG GAA GTT CTT GAG TTA CAA A
SPA4666600.62	62	AAC CTG AAA ATT TAA ATG TTG AAA CGT CGA TGT CAG ATG AGG CTT TCA ATG CTG ATA AAG TG

SPA4667100.58	58	CTT TGA AAA GAG TCT TAA ATC CAA AGA AAT GTT ATT TGG TGC AAT TTA CAA AGT GTT T
SPA4667600.63	63	TTC GAA GAA CAC GTA ACA TAC TGG CTT TAT TTT CTA CTG CAA CAA CTT CAA AAG AAA CAA GAT
SPA4668200.63	63	GGT ATT TGA TAA AGT CAT TTG CTG GGT CAA GAT TGA CGA TAA TAC TAC TCC TTC CTA TGG CAG
SPA4668700.63	63	TCC TTA TGT AAT ATT GAA GTT TGA GAT GAC ATT TAA AGA GTT GTG GCA GTT ATT GAA ATT TGG
SPA4669150.58	58	AAA ATG AAA ACA TTA CAT CCT GAT GTT GCT CAG GGA AAA GAC GCT GCC TTA GCT CAA A
SPC149000.63	63	ATA TGG AAA TCT TCG TAG GTA GGA ATT CTT ATT TTT GAT GCT TAG CCA TAC GCC AAA TGA TTT
SPC149500.60	60	CCT TTG ATT GCC AAG CGT TCA CAA ATT TAT TCT TGT AAT TTA AAC GTT TCA TAA CGG CTT
SPC150050.58	58	ATT TTG AAC TTT AGG TGT CCT ACG TTT CGC AAT GAA TTT TCC TAA TGA ATT TTG AAA A
SPC150550.58	58	TTC TTT TTA GTG CAA AAC ATG ATG GGC AAA ATC GAA TCA TTT TGG CAT GAG TAG TCG A
SPC151050.63	63	GAA TGC CTT CTA CAA GAA TGC GTT ATA TGA ACT CAC TTA CAT TCA ATG ATT AGA ACT CTC CGA
SPC151550.62	62	TAT AAT TAA GAC GAA GGC AAT GGG GTA CAG CCA GTA CAC ATG ATA TCC AAA CAC ATG AAT AC
SPC152100.58	58	GAA TCG AGT AAA CTT GCA CTA AGC ATG TTT GTG TAT TGG TAG GCT AGT ACA ACG AAA T
SPC152650.63	63	TTG AGG AAT TAT GAG TGT AAT GAC TTT TAT CGA AGT CAG AAG TAC TGC TTT CCT GGT ATG ACT
SPC153100.62	62	ATG CAG TAG TAA ATT AAT AGT AAA AGC TCG ACT TTT GAA ATT AAT ACC TGA GTT GGC TTG CT
SPC153600.63	63	TAT CAC GAG TAT TAA CGA TTT GGT AAA TGT TAA GAC AGT TTA TTA TGT CCG GTC CTC CTG TTT
SPC154100.63	63	TAA GCT TGC TAG GGG TTT TAA CGA AAT TGC GTT

		ATG TCA TTC TGT GTA ATA AAA TTG TCA CCA
SPC154600.63	63	TTT TTG AAA GCA AAT TCG TAA GGA GAA TAA TAA TAG TAT TAT AAA CAC CAC TTG ATC GCA GTG
SPC161200.59	59	GAT TTA AGA TTT CGA CAA CCT ATG AGC AAC GAT ATT GTA ATA CAT GAA AGT ATA GTG GT
SPC161700.57	57	TGC AAA TGA TAA TGC AAT AGG CAC AGA GAG ACA AGC AGT CGA GCA GCC AGC TAA CAC
SPC162200.63	63	GAA ATA AGT CAT AAG ATT TAA CTT GCA CTT AGT GAA TTG ACT GTA TTA TCA AAT TGG TAG GTC
SPC162800.59	59	CTT TGT ACT TAT AGT AAT AGT TGT TGT TCC CTT GGC AGC CCA AAA TCA TTT TAA TGG TT
SPC163350.63	63	CGT GTA GCG GAT AGC AAT GAA AGA GAA TAT TTA GCA GCA GAA ATA AAA AAG GCA AAT GTG ATT
SPC163850.63	63	AAT AAT CTT ACT AGA CAA TGA GAA AAG CTA CAA TTC ATG CTC TGA GTA GGA TCT TCT GTA CGT
SPC164350.63	63	TGG TAG ATC TCT TTT ATC AAT TCG ATA GAG ATA ATG ATG GTG CCT TGA ACA ATG AGG AAC TTT
SPC164850.62	62	GTG TTG AAT TCC AAA GCA CAC AAA GAT ATT TGG TGG TAT GGG ATA TAT TTA TTT AAT ACT TT
SPC165400.63	63	CAG TAT CCA GCT TCC TCT ATT ATC CGC ATT CCT GAA GAA GAT TCT AAT AAG ACA AAT TAC CAG
SPC165950.61	61	TTA CAA AGT CGG GTG TAT AAA TAT GCA TCT TTA AAA GTA GCA GAG TCC ATT TCT TGA TTG C
SPC166500.57	57	GAT CTG GTG ACT GGA AAA CAT CAG AAA TGT GTC CCA CTA TCA ATG ATA AGC GTA CAA
SPC166950.63	63	CTT TCA TAG GAG TAT TAT GAA ACA CAT GAC CTA AAG TAG CAA ATG GTG GTC GGT TAA CAA GTG
SPC174150.61	61	AGG CCT ACA ACG AGT ACC AAA AGT ACA ACA ACA ATG ACA GCA ACA ATA ATA GCA GCA ATA A
SPC174650.57	57	AAT GAC AAC TCC TAT GGA GGA AAC AAC AAC AAC AAC TCC TAT GGT GGA AAC AAC GAC

SPC175100.58	58	GGT AAC AAC AAC TCC TAT GGT AAT TCT GGT AAT GGC AGT GGT TAT GGA AAC GAC TAT T
SPC175500.63	63	CTC ATG AAT AAA ACA AAT TAC TCT GCT AAT CAA TCC CAG AAT GGG AAT TCT CAA AAC AGC GGT
SPC176000.63	63	GAA CAG TAA TAA GCA GAA TTA CAA CAA CAA CAA TAA CCA GAA TTA TGG CAA TAA CAA CAA CCA
SPC176500.63	63	AGT TTG TTT TAG TAG CTT TTC ATC TAA ATA GTT AAT ATC TGA TAT ATG CTG AAA TCT ACA GTG
SPC177050.60	60	TAC GAC ACA CCA TTG ATC GAT TTG ATT TTT AGA GCC GCT TCC ATA CAT AGA AAA TTC CAC
SPC177500.60	60	ATT ATC AGC ACT AGG ACA TAT GAT GAA CGG TTA AAT ACT ATT GAC AAT CTT CGG AAA GCT
SPC178100.62	62	TAA AAG CTA ATT TTG GTT TTC CCA GCC TGG TTC TTT GAT GTG ATT CCA AAA TCA AAT CAG AC
SPC178550.62	62	AAC AAC TTC TGT AAA TTA TCT CAC TCA TCA CGA CAA CGC TTG ATT GCT AAT TAT AAA TTT GT
SPC179100.63	63	TTT AGG AAT ATT ATT CGA GAT ATC TAG ATT CTG CGC TAG TTC TTG CAA TTG TAA ATG ACA GTG
SPC179600.61	61	CCT TTG AGT AAA TTT CAA TCC AAC ACT GAG GAA GTC AAT GAA GAC CCA ATT CTA AAG CCT T
SPC186550.59	59	AAT AAC CAA AAT TGT TGT TGA TTG CAG CGG GAG CAT CTC TGG AAT AAG CAT TGG ACA TT
SPC187050.62	62	AAC GGT AAT TAT CGC AAT GCC GTA TAT ACT TTA ACC TGT CAG GAA TAG TAC ACT ACG CTA TG
SPC187700.57	57	CAC CAC TGT TTG ATA CAC CCT TGC CTG CTA AAC AAT TCA ATG TGA GTT TGT AAA TTG
SPC188100.58	58	ATG GAA AGG AAA TCC CAA GTT TCA GAT ACA AAT AAT AAC TCT ATA CCT ACG TAA GTT T
SPC188650.63	63	CTA AAT CAA ATT TTG AAG CGG ATT TTC CTT CTG CCA ACG TAT TTA GGT GAA TCA TTC GAT GTC
SPC189200.63	63	AAA GTT TGG TGT TAA TTA TTT ATT TTA TTC TGT

		TTT CTG CTG ATT CAA AGA CTT CGA ACA TGT
SPC189650.58	58	GAT CTG ATA CAA ATT TTC AAG GTG AAA ACG AAC AAA TTC CTC AAG CTA CAG CTC CTT T
SPC190200.63	63	GAT TAA GCG AGA AAG ACT TTT CAA ACC CTT AGG AAT TCA ATT TTA TAT TAT CTT TTC AAC ATC
SPC190800.63	63	ATA AGC AGC GCC AAG AAA TAA TAA TTA TTT TTA GGA CGT GTT CAA GGT AAG TAA TTG AGA ATC
SPC191300.59	59	GAA GGT TGA AAC TGC TCG TTT GGA GAA TGA ACA GAA ATC TTA CGA AGA AAT GAA ACA AA
SPC191800.63	63	AAC TTA AGA TAA AGA ATA TCT AGA CAT AGA AAA GCT AGA GCT TAA GAC TTG ATG ACT TCT ACA
SPC192350.63	63	ATA ATT ATA AAA GCA AAG CAA TAC ACG GGT CTT AGA TAG CTA GTC AAA ACA ATA GCA ATG CGT
SPC199150.57	57	TTA GCT TTC GAT CTT AAT TCC AAG CAT CCT AAC AAG GAC CAC TTA CCT CAC ATT GTC
SPC199700.63	63	TGT TCG AAT TTA CCT ATA GTG TTC AAA TCG TCT ACG AGG CTC ATT TAA GTA ATA ATA TAT ACT
SPC200250.63	63	TTT CAG AAA ATA TTT ATG TAG TAC TCA GTA AGT AGA TGG AAA GGT TTA AAG ATT GCT ATC TCT
SPC200800.57	57	CAC TGT TCA AAC GGC ACA AGG AAT TCG TTT CGT CAA CTT TTG GTA GTA TGT GTA TTG
SPC201400.60	60	TAT GGT AGC ACC ATG GCG TAC TAG TTT TCT GAT TTG TCG GTT ACT GAA TAT TTA TTG AAT
SPC202050.63	63	TGA AAT TTT TCT CAA AGT CTG CAC AAA CAA TTT TGT CAT TAT CCA TTG GAT CTA TAG AAG ACC
SPC202550.58	58	TGT GTC AGT TTG CAC TGA ACT TCA CTC AGT CTG TTA GTG TTT TCG TTA CTG AAG AAT T
SPC203100.61	61	GTG ATG AAA TTA CAC AAA ACT CCA ATC ATG CTG ATT CAT GGT TGA GAC AGG ATG ACT GAA T
SPC203600.59	59	AAC TTG AAC GTT TTT GCA CTA AGT GCG GGG TTA TAA TTG ATG AAT TAG CTA GTT GTC TG

SPC204150.58	58	TTT TCG TTT GTT TAC ATC CCA TTT TTC TCC ACA TTA TTG ACA CTG TTT GTT TGT TTT T
SPC204550.59	59	TCT CTA TAC TTA CTC TTC TAT TCT TTC CTT CTT TCT CTA ATA TTA CTA CAT CGT TGC GA
SPC205050.57	57	AAT TCC ATC TTG GGT TCA AAT GAT GGT AGC TAA AAT TTG GAC TGG TCT TGG CAT TGG
SPC212150.59	59	TAA CGA TAG TTA AAC TTC GGT TTC GGT TCT ATT TTG GGC TTT TGT CTT GTC ATT GAA AG
SPC212550.60	60	ATG GAT CGC ATT GGT AAG CGT ACC TCT ATT ATG TTT TGG ACT ATT GTT TAT TTG ATT GGT
SPC213100.57	57	GTT ATC GTA CCT TCC TTG GTC TTG GTG TCA TGT CTC TTC AAC AAC TTA CCG GTG ATA
SPC213550.58	58	TGT GGG GTT TCC TTA TTT CCT TTT TCA CTC CCT TCA TTA CCA ACT CCA TTG GAT TCA A
SPC214050.63	63	TTT ATT ATA TCC CTG AAA ATT TCT GAT GTT TGT TAG GAG TAG CGG TGG CAT TTT TTA AAG ATT
SPC214550.59	59	ATT TTT CAA GTT TGT TAA AAA TTT GAA TGA GTA AGG TGA TTA GAG GAA CGG AAT TAT GT
SPC215100.57	57	TTG ACT TGC TTG CAA AGG GAA AAG TGT CTT TCC AAT ATC TGA CTG TCA ATG AAA GAG
SPC215600.63	63	TGT AAG TGA GCA TCT TTT CTG ATG TTA CAA TAT CAA TGA TTT CTG AAA GTG AAA CTT CAT TTG
SPC216100.60	60	GGA AAT TAG AGG CTA CTC TAC ATC AGT GTT TTC CCT TTT AAT TCC ACA CAT AGT TTT CTT
SPC216550.60	60	TTT GAT GTT GTT TAC TTT AGT ATC TTG CTA ACC AAC ACA AGA CAT TTT TGA ACA GGT TGA
SPC217050.60	60	ATT TAC AAT ATA ATG TTT GAG GCA CTT TAG ACT TGT TTA TTT CTG TTC CCT TCT TCA ATT
SPC217550.57	57	ATG GTT CGA AGG TTC AGA CCT GCA ATA AGT TTA ATT AAT TCA AGC AAC CAT GTA CGC
SPC226050.63	63	TCA CGT ATA GTT TTA AGT ATA ATA TTT TCA ACC

		ATA GCT TTC AAT GTT TCA CTC TAA GTA GAA
SPC226500.63	63	ATT ACC AGT TAA AAA GCA TGC ACA CCA ACT TAA TAT TAA TAT CTC TAA CAG AAA GGA CGG AAC
SPC227000.58	58	ATA CAG TTC GTT AAT CTC TTC CAA GGT AAG ACC TTT GGT TTC ATG GGC AAA CAA GAA A
SPC227500.61	61	AAG GAG AGT TCA TAC CAG TAC CCT CAA ACA CCT CGA AAC CAT AGT AGA AGT AGT AGT TAT C
SPC228100.62	62	TAA TAC CAA TCA AAT AAA CAA TAG TCC AAA ACA TAA TGG AAA CAC GCT TGC CAA TGC GAT CC
SPC228550.63	63	GTA TTA ATG TAA AGG GAA TGT GGA TAA TAT GAA CCA GCG AAA TTA AAT AGA AAA GAA TAG AGG
SPC229000.57	57	TGA AGA GAG GAA GCA ATT ATG CTC AGT TTT GCA CTC CCA AAT GCG GAA CCA TTT ATG
SPC229550.57	57	TAG GAT GGT CAA ATT AGT ACG GCA CTA TTC CGA ATG CAC CAT AAC TGT CCA CTC CGC
SPC230000.57	57	GAC AAG TAT AAA CAA ACC AGA AAA TGA TAT GAA GAG ATG AGA CGA GAT GAT CCA AAA
SPC230550.60	60	GGG CAC CCA CTC AAC GAG GCA TAA ATG ACA AGA GAC TAT GTA GCG GCA TTT GAC AAA TTA
SPC231000.57	57	TAC AAA GAA ATA TCC ACC CTT TTC GAT ATC AAA GCA AAC GTC CAT CCC TCA AGA AAA
SPC231450.57	57	CTT CAT TTT AGG GTT GTG TGT CGG CTT GCG GAA GCA CGT GTA TTA ACG GTG CGA AAA
SPC40169.60	60	CTC GAA ACT GTC ACC CCT TCA TCT TCA TCT TCC TGC ACT GCC TTC AGA TTG TCC ATA CCG
SPC40289.60	60	AAT TGA ATG AAC AAG TCG CAT TGC TCA CAA CCA TGG TAT AAT CTA TCA AAA AGC ACA TCT
SPC40409.60	60	AGG TGT ACA ACA TCT ATT ACA ATT CAA CGG AAT ATT ATT CCT GTT TCA TAA TAT TTC AAT
SPC40529.60	60	AAC TCT TAA GAA TGT TTT TGA TCT TTT AAT AGA AGC CTT AAT TGG TAT TGA ATA ATT ATA

SPC40649.60	60	TAA TTA CTA TGA AAA GTA CAG TAC TGT TGA GTG TTT TTA CCG TCT TAG CGA CAT GGG CTG
SPM1451.63	63	GAA TAT CTT AAG GTG TTG AGA GAA CTA TGC TTA AGG AAC TCG GCA AAT TAG CTC TGT TAC TTC
SPM3401.61	61	GTA AAA GCT TAA CAA GCT CAA ATC GTA ATC ATG ACA CTA AGT GGT GCT CTG ATC ACA TTG G
SPM5401.57	57	CTT GGG GAA AAT TCG CCT GGT TAA ATC CAT CAG AGA CTC GTT TCT ATC ACA ATT AGA
SPM7051.60	60	AAG TTT CTA TTC ATC TCC CTA AAC ACT TAA AAC CAG CTA ATG ATT CTC AAT TTG GTC ACT
SPM9051.63	63	CTG TTT AGC AGC TAC ATT ATA TTT ACA TGG ATA CAA ACA TTC ATC AGT ATT CTT TGG AAT CAG
SPM11201.57	57	AAA ACC TAC GTC TAT CAT CTC AAA CAT ATA TAC AAC TAC GGG TCC CGC TAA GGT CAG
SPM13351.63	63	AAC CCT GCA AAG GGA ATG CGT CTT AGT TCA ATG ATT TGA TAA CAA TAT TTA TCT TCT TAA TCG
SPC155000.60	60	AAC CAT TCT ATT CTT TCT CTT TCC TTG CTA CCA CCA TCC ATT TCT AGT CAT TCA GTT CAG
SPC155450.57	57	CCT CAC ATC GAC GAC ATA AGC AAA TTC CGT TCT CTC ACT AGG CCT ATT GAC CGA TTT
SPC155900.60	60	TGA ATA ATG AGT CAA CTT TAT GCT TTA AAC GGA CAA CCT CTA GCA GGG AAA TCT GAT GAC
SPC156450.58	58	CTA GCT TGG TAC CGG AAC ATT ACT TTG ATG ATA CCG ACA AAT CAG TAC ATT CTA AAA G
SPC157000.59	59	CCC AAT AAA GGC TCC TAA AGC TGA ATA TCT TAA TCA ATC ACG AGC ATC AAT GGG AAA CC
SPC157600.60	60	ATA TGT CCG TGT TTG TGA ATT AAA ATG CCT TCA GGA AGA AGA ATT TGA TAA TAA AGT GCT
SPC158050.59	59	TGC TGT TGA ACG AAA ATA CCA TTT TCA AAA TAT GAC TCC AAA ACC AAA TGG AAC CAA CA
SPC158550.61	61	TAA CGA CGT CTT CCA TTT AAA GTC GAA TTA ACG

		ATT TAT GTT TGT TAG TTC GTC ATG GTT T
SPC159000.62	62	AAC AAT GGA TTT ATG GTG ATT GTC GTT AAA CCT CTG GTT CGA TGA TGA TGT TAG CAA TGA AG
SPC159550.60	60	GAA AGG AAC TTG CGG AAT CGT TAC ACT TCA ATA AAC TAA TTT CTC GAG CCC TAA CGA ATC
SPC160050.61	61	GTG ATC GTA ACC ACC GAT GAG ATC CAA AAG ACC ATG TTT TTC AAA AAT ATG AAG AAC TTG C
SPC160700.61	61	GCA TGC AAG TAG GTA CAG TGG AGA ACA CAG GCA TGA TAA AGA ATG TTG CTA ACA AAA TAA G
SPC167550.58	58	GAG GAG GTA AGT AAA CTG GAT TGT TTG GCT TGG GAG GGG AAG GAA AGT TAC TAG ATA A
SPC168100.60	60	AGA AGA TGA TTT AGC AGA ACC AGA GGC TCG CTT TTG TTT TTC ATC AGC ATT TCG ATT AAA
SPC168700.63	63	TAA CAA AGA TTC AAA CGT TAA AAA CAA TTA AAG AGA AAA TGC AAA ACG CAG TTA AAA TGA TGA
SPC169200.58	58	AAT CAA AGT CAA ATC CCC AAC AAG CAA GGC TAT GAC AAA GAA GAT GAG CGA AGC AAA A
SPC169800.63	63	ATT CTG TAT TTA ACT CCC AGA AAA TGA ATT GTA GGT CAA CGA TAT GAT AGT TGA ACG ATG GAA
SPC170300.63	63	CAA TGA ACC TCG AAT TAT GTA AAC AGC GTA TTC AAA TTA TCG GAA AGA GTA ACT GCT AAC CAG
SPC170850.57	57	TCC GAC CCT CTA CAC GAT TTG ATT GGT TTG CTA CAG CTT AAT GTG CAT CCT AAC CCC
SPC171400.62	62	ATA TGT GAT CAT GTT AAT GGA GGC ATT GCA CTG GTT TTA AGT CAG ACT CTA GTT TAA CTC GG
SPC172050.58	58	ATA GTA TCG TGG CTG GTC CAA ACC CAT CAT CAA ACT TAA CCG ACT CAT TTT GTT TCA T
SPC172700.60	60	GCA TTA CTC AAG CTG GAT ATA GGG AGA GAA TGA GTT TGT ACG AAA TGG AAC AAG ATT AGT
SPC173150.57	57	TTA CCT CGT CCT CGG TTA CTA TTC TAG TGA CCA CTA CTG CTG TTG TTG CTA TTC CTA

SPC173700.57	57	AAC TAA ACC TCT TGA TGC CTC GCC TAC TGT TCC CTT CCC TAG CGT TAC CCG AGA AAG
SPC180100.61	61	AGA TTG TAA CCG TCT TGC CGA AGA GTA TTT GGA ATT ACG AAA TAT GCA ATA TTC TAA CTC G
SPC180600.63	63	TAA CTT GCG CAT TGC TAG TCT GCC ATA TTC TTT TGG TTT GTT CTA TTT ATT TAT TTT CGC TCT
SPC181100.57	57	CGC CGC TGT CAT ACA GAA ACC GTG AAT TGG TGC GAG AAA ATG GAA CAG TAT TAT TTC
SPC181650.59	59	GAA AGG AAT ATT AGA AAG TTT TGC TTG AGC TTG CTC TTT GTA TTT TTC TTT AAG ATA GT
SPC182150.57	57	TGA TCT TAA GTA GGC ATC AGT AAG TAG TGA ATG AAT GAA TGT AAC TCG ACG CGT GAT
SPC182700.63	63	AAA AGA TGA AAA GCA GAA GGG TAA GAC TCA ATA ATA CCT TGT ATA ATC GAT GAA TTA TTA ACT
SPC183200.59	59	CTT GGT AGT TTT TTC TAT TGT GCT TAC GAC AGG TGA CGC CTC GGT GTC TAA ATC TAA AT
SPC183800.58	58	ATC ATC GAA TGG ATT ATA AGA GCA GTC ATA CAA TAT CAC CAC ATT TGC ACA CGC TAA A
SPC184450.63	63	CAT ATC GAT TGT AAA TTT CCA ACT GGT TTT CTG ATA ACT TAC AAT GTT CAA TAA TCT GAG TCT
SPC184900.60	60	CTA ACT GAT AAG TAG TAA CCA AAA TGT CAT ACT TTA TTT CGT TTT CTT CTA TTG CTT CGC
SPC185400.63	63	AGA AAC TCG CAT TTG GCA ATT AGA TTA TCA ATT GCA TAA TAT CCT TCC ATG GTT TCC ATA CAA
SPC185950.62	62	TTG GAA TAC CAA GGT CAC AAC TCA AAT ATT TTT GAG ATA AAA TAG AAA GTG CTG CAG TCT CA
SPC192800.59	59	GTT TGG AAT CAC TGT TAA AAA CTG ACA GAA AAG CAA GTA CTC CTT ATG ACC CCA TGA AC
SPC193350.63	63	CCT ATA TGG ATT AAT ACA GTT AGT AAA TTA GAA AAT GAG AAG TTG AGA TAA GAC ACT GCA CCT
SPC193850.62	62	TTT AGA GCG GTC TTA TAT CAG AAA GAA GCA TTC

		GCT TGA CCT TGA AAC ATT TGT AAT AGA AA
SPC194350.62	62	TAA CTA ATA CTT CCA CAA CTA CTA AAA ACC GTC TCA ATG GGA ATT CGT CGA CAA TCC TTC AA
SPC194900.60	60	AAT GAA AAT CCT TGT AAA CTC CCT TTC TAT CTC ACC GTA CCT CAT ATA TGC GAA ACG ACA
SPC195400.57	57	AAC AGA TTA TTG TCT GCA ATC TTC TTT CGA TAA CCA AGA TAT CAC TGG CGT TGC TTG
SPC195900.62	62	TAG TGA TAC CAC TGT TAA GGT AAA TTT CTT TTG TCA TTC TAA TAA TTA CTG TTT TCT AAC GT
SPC196450.58	58	TAT CCA TTT ATT ACA GTA GGT TCG TTG GGT CAT CAT CTA TAG TCA AAT ATA TAC ATA T
SPC196950.57	57	GTT TTG AAG GAC TGC TTC AAC CAA GGA CAA AGG TGC ATC CCA GTA TGC TTC TTT CAG
SPC197500.63	63	TCA AAG GTG TTG AGC AAA GTG TTA TCC ATA AAG CCA GCA CAC TTT AAA TTA TCA TTA GCT ACT
SPC198050.57	57	AGA CTG CTT TAG TTA CAA GGC CAA CAA GAC ACT CTT GAT AAA TGG AAG CAA ATG TTT
SPC198600.59	59	TTC TCG AAG CTT AGA AAG TTT ATA ATG GGC GCG TAA CAA ATT TGG CAT ATC AGA GTC AA
SPC205650.63	63	TTT GAT TCT TGA TGC TGT CAA CTT TGG ATG TAC GTT TGG TGC TAT CTT TGT TCT TGA ATA CTT
SPC206200.57	57	CAA AGC GAC TCA GAA AAG GAG CGT GGA CCT ACT TCT AAG TTA CAT GAA TAC GTT GAG
SPC206700.62	62	ATG TTG CAT TTG TTT GAA GTG GTG TTT GTG AAT GAA GGA GTA AGA GAG GAG AGA AAG ATA TG
SPC207600.57	57	ATG AGA GGA GAC GAG GAG AAT GGA ATG GAA TGA TCA TGG CCT AGC TGA TAT AAC GAT
SPC208050.58	58	TCT CAT GGT TCG GTG TTA TAG GGT GCT GGA ACG AGT AAG ACG ATA AGA CGC AAG CCT A
SPC208550.59	59	TAT TAT CAA ACA CCG AAA TAA GGG TAC GCA CTA GGG AAA GAC TCA TGT GAT GGA AAG AA

SPC209050.57	57	GTT GAA TGC AAA CAT TTG AAA GAC AAT GGA TGA GGG TGT CAA GCA TGA AGA AAG AGA
SPC209550.60	60	TGT AAC TAG GTG AAT CAC CAT GTA TCC AAG CTG TAC GTA TCC ACG CAA AGT TAA GAT TAC
SPC210050.57	57	AGT GTG TGT TGT GGA TCT TGT GGA TCT TAT GGG ACA CGA ATG TTT TGA GGG GCA GGG
SPC210550.57	57	CGT ACT GCC TGA TGC TGG AGA TAG CTT GTC ATG AGG AGG AGG AGG ATG ATG CAT ACC
SPC211050.57	57	GGG TCA GTA CCG TTA CAT CCC TTC CCA TCT CTC ACT GTC TCA TTG TCT CTC TCT CTC
SPC211600.63	63	TAT ATA TTA TAA CTG CAA ACC CGG TCT TTT CCT TCG TCC AAC AAT CGG TTT TAT CCA TCT CTA
SPC218150.63	63	TCT AAA CAT ACT GGA ATA GAT TAT ATG TGA ACT AAG GAA CAT TCG AAT AGA TCA TCT TCA TAG
SPC219050.63	63	TGA AAT ACT CTA AAC ACA TCA TGA TTT GTT TGG TAA ATC ACA ATT ATA TAG ATG CTT CTT TCC
SPC220450.63	63	AAT AAT AGC AAG TGA AAA GTA AAA TTC CAA AAC TCA TTT TCA AAA TGT AAT TGC TTC TGA CAC
SPC221400.58	58	AAA GAA GCT ATA CCA TTC CCG ATA AAT TTG GGC ATC CTT TTG AGA GCA CCT GGA CGT T
SPC221900.58	58	CAA TTA CCT TTA CAC ATT TTG CCA AGA AAA CTG CAG TCA CCT TTA CAC ATT GTG CCA A
SPC222500.61	61	AGA CAC AGC AAA TAA TCC TTG ACA CAC TCA ATT CTT TAT CTT ATC CAC CAA CAA CAC AAC A
SPC223050.57	57	TAG TAA TTT CAA TAT AGC GGT ATG CGC CCT GTT AAC TAA CAA TAT TAT ATT TAT TGT
SPC223500.63	63	CTT TCA AAC CTT GGC TCA ACT AAG TTG GGA TCA TTG ATT GAT TAT ATG GCT CAA ATA ATA GAA
SPC224000.58	58	AAA TCA TTC ATC ATG AAC AGA AAC CAT AGA CCA GGT TAA GGT TAC TTG ATT TCT CGG T
SPC224500.57	57	CCA AAG CAG TAA TAC CAT TTC CAT TAC TTC CAT

		TGA CAG GAG CTA CGA AGT CAT CGT
SPC225050.59	59	ACT GTT TGA GGC AGC AAC CAC ATT CAT TAT TTC ATA TAC AAG ATA TGC TGG TGA AGG TT
SPC225550.62	62	GAA CCT CCA TTT TCT CAT CGT CCA AGT TAT TAT TAT AAT TTT CTT CAT CGT CTT CAT CCA AA
SPC232150.62	62	TTA GGT TCA CTC TAC TCT CCC AAT GTA CCG TGG ATT TTT ATT GCC TTT AGC TAT TAA TTT CT
SPC232650.58	58	TAA ACA TCC TTA AAG CAG AAA ACG GAA TCA AAC AAA AAA CAA TTA AAT ACA GGT GTG T
SPC233200.60	60	GAT GTA ACC ATA CTT GAA TCC AAT GGA GTT GGT GAT AAA TGG AGT GAA GAA GGA AAT CAA
SPC233700.57	57	TTC CCA AAA ATG TTC TGT ATC TGA TTT CCT TAC CAA AAA TCT CTG GCC ATG TGC ATG
SPC234200.58	58	TGA ATG ATA ACG CCG ATG AGG TAA ATA ACA CAG AAC CCA ATA ATG GAA TTA CGC TTA C
SPC234900.57	57	TTT TAC GCT TAC GAT TGT TGG ATT GGA ATA GGC AGT TGA ATG AGA GAG AAT GGG AAT
SPC235400.62	62	GTT GCA TAC CAA GCC CAA ATC CAT TCA CTT CCT CTA TGC CTG TTA AAC TCA AAC GCA AAT TT
SPC235900.59	59	TCC AAC TAT TCC TCT CTA CCT TAC TCC CCT ATA CCT TAT CCT TTG CAT CGG AGG GTT TT
SPC236350.57	57	GTC CAA TTC ACC ATT AGC TCA TCC AGG TAA GAT GTC TGA ACA CAC ATC GGA TCA CCT
SPC236850.63	63	CTA AGG TGA TGG TTG GTA AAA CGT AAA CAA TAG ATA GTG AAA CTG TGG GAT ACT ATA AAT CAA
SPC237350.63	63	TTC TTT TGT ATT CAT TTA GTT TTA TTA AAC TCG TTA TTT GCT TTG CTT GGC ATC TCT TGA CTG
SPC237850.57	57	CAT GAT TAG CGC TCC TTT ATC GGA AGT TTT TGG ACG TCG TAT GCT ATT GCA AGT TGG
SPM15401.63	63	ATT GGT TTC TTA CCT ATT ACA GTT TTA GTC GCT ATC TCT TTA TTA GAA TTT GGT ATT GCT TTT

SPM17651.57	57	GGT TTA AGG AGC GAT ATT TGA GCT ATT GTG ATT TAT TCA TCG CAG GTT CGA ATC CTG
SPM19201.57	57	ACG GAC AAT GTA GTG AAC TTT GTG GTG TAC TTC ATT CAT CTA TGC CTA TTG TTG TTC
SPA4665250.63	63	ATA CAC ATG CGA TGT CAC TTT TAT ATG AGT GTA TCA ACA CAA TAG TTT CCG GTA ATA TGC TTG
SPA4665900.62	62	TTA CCT GCC ATC ATA TGG TGC CTC GGT GAA TAT GCT GAA TTT ATT GAG GAA TAT CTT GAT AT
SPA4666400.63	63	TGA AGT ACC AAA AGA GCT ACT TGA AAA CAT CAT CCA AAG TGA TGA TTC GTT AAT TAA TTT TGA
SPA4666850.62	62	ATA CTC CTT AAT TTC TTT GGA CCC TCC ACT TTC TAC CAA TCA AGG ATC AAT GGG TGA TAT TG
SPA4667300.58	58	AGT AGC AAA AGA TGA AGA TGA CAC AAG CAA GGT TGA GTA AAA TCT TTC AAC TTT CAA A
SPA4667900.57	57	CAG AAG GAT TCG TGC AGC AAT ATG CAT CTA CAA GTT GTA CGG ACA CAG GTC TAT AAT
SPA4668500.63	63	AAT ACA AAG ATT GGT AGA ATT TCC TTA AAT TTT TCA AAT GTA GAC GTT AAA TTC ACC GAA CTG
SPA4668900.60	60	CTT TAG GCC CTT TAT TTC TTT TGC ACA GAA GAG TTT ATT CAA TCG ACA AAT TAC TGG AAA
SPA4669450.57	57	ATA CGG AAA GCA ATG GAG TCT TCA AAT TGG GAT AGT GCT CTG CTA TAT GTC AAT CGG