

**INVERTED REPEATS AS A SOURCE OF EUKARYOTIC GENOME
INSTABILITY**

A Dissertation
Presented to
The Academic Faculty

By

Vidhya Narayanan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Biology

Georgia Institute of Technology

August, 2008

**INVERTED REPEATS AS A SOURCE OF EUKARYOTIC GENOME
INSTABILITY**

Approved by:

Dr. Kirill Lobachev, Advisor

School of Biology

Georgia Institute of Technology

Dr. Yury Chernoff

School of Biology

Georgia Institute of Technology

Dr. Michael Goodisman

School of Biology

Georgia Institute of Technology

Dr. Todd Strelman

School of Biology

Georgia Institute of Technology

Dr. Gray Crouse

Department of Biology

Emory University

Date approved: June 27, 2008

ACKNOWLEDGEMENTS

I would like extend my sincere gratitude to my advisor Dr. Kirill Lobachev for his immense help, guidance and scientific inspiration throughout my graduate career.

I would also like to thank my colleague and good friend Hyun-Min Kim for assistance with experiments.

I thank all the undergraduate students of Dr.Lobachev's lab.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES	x
LIST OF FIGURES	xi
ABBREVIATIONS	xi
SUMMARY	xiii
CHAPTER 1: INTRODUCTION	
1.1. The pros and cons of chromosomal rearrangements: contribution to evolution of eukaryotic genomes and human diseases.	1
1.2. Hairpin- and cruciform adopting sequence motifs as sources of genome instability	2
1.3. Mechanisms of palindrome-mediated chromosomal fragility	5
1.3.1. Replication arrest at the location of hairpin structures and subsequent breakage	5
1.3.2. Center-Break palindrome revision mechanism	8
1.3.3. Cruciform resolution model	8
1.4. Implications and concluding remarks	10
1.5. REFERENCES	
CHAPTER 2: The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks	
2.1. Summary	16
2.2. Introduction	17

2.3. Results	19
2.3.1. Experimental system	19
2.3.2. Inverted <i>Alus</i> strongly induce arm loss events characterized by a specific pattern: terminal deletion coupled with adjacent inverted duplication	21
2.3.3. Chromosome rearrangements in strains with direct <i>Alu</i> repeats	27
2.3.4. Hairpin-capped breaks trigger a palindrome-dependent recurring chromosome instability	29
2.3.5. Extrachromosomal amplicons resulting from hairpin-capped breaks are linear inverted dimers	31
2.3.6. Intrachromosomal amplification is an alternative outcome of the repair of hairpin-capped broken molecules	34
2.3.7. Elevated levels of deletions and amplifications in <i>mre11</i> strains	39
2.4. Discussion	40
2.4.1. Mechanisms of GCRs triggered by hairpin-capped DSBs	40
2.4.2. Palindrome regeneration cycle leading to continuing genetic instability	41
2.4.3. Double minutes versus homogeneously-staining regions	42
2.4.4. Implications for human genome stability	43
2.5. Experimental Procedures	44
2.6. Acknowledgements	45
2.7. Supplemental Data	45
2.8. References	56

2.9. Contribution to the publication	60
CHAPTER 3 Intrachromosomal gene amplification triggered by hairpin-capped breaks requires homologous recombination and is independent of non-homologous end joining	
3.1. Summary	61
3.2. Introduction	62
3.3. Results and discussion	65
3.3.1. Experimental system to study intrachromosomal gene amplification mediated by hairpin-capped DSBs	65
3.3.2. <i>Dnl4</i> , <i>Hdf1</i> and <i>Hdf2</i> are not required for inverted repeat-induced intrachromosomal gene amplification	67
3.3.3. Intrachromosomal gene amplification depends on homologous recombination	69
3.5. References	74
3.6. Contribution to the publication	78
CHAPTER 4: Post-replication repair mediates the inverted repeat fragility under conditions of compromised replication.	
4.1. Summary	79
4.2. Introduction	80
4.3. Results	84
4.3.1. Experimental system to study gross chromosomal rearrangements and chromosomal fragility resulting from inverted repeats	84
4.3.2. Defects in DNA replication leads to elevated levels of chromosomal fragility and GCR formation	85

4.3.3. Replication fork stalling at <i>Alu</i> -IRs in replication	86
mutants triggers elevated hairpin-capped DSB formation	
4.3.4. Chromosome fragility due to compromised replication	88
requires some components of PRR	
4.3.5. Mrc1 and Tof1-Csm3 are required for the hyper	90
fragility observed in <i>pol3-P664L</i> strains	
4.4. Discussion	91
4.4.1. Replication-dependent mechanism of <i>Alu</i> -IR	91
fragility requires PRR pathway	
4.4.2. <i>Alu</i> -IR-induced replication arrest activates Tof1-Csm3	94
and Mrc1 proteins	
4.5. Materials and Methods	95
4.8. Supplementary Information	99
4.9. References	101
4.10. Contribution to the publication	105
CHAPTER 5: Estimation of mutation rates	106
CHAPTER 6: CONCLUSION	106
6.1. References	109
CHAPTER 7. List of manuscripts (published and in preparation)	113

LIST OF TABLES

Table 2.S1	Induction of GCR events by homologous and homeologous inverted <i>Alus</i> in wild type and <i>Δmre11</i> strains	50
Table 2.S2	Can ^R Ade ⁻ isolates of strains with direct and inverted <i>Alu</i> repeats	51
Table 2.S3	Cu ^R Fh ^R isolates containing intrachromosomal amplicons	52
Table 3.1	Intrachromosomal gene amplification in wild type, NHEJ and HR mutants.	73
Table 4.1	Effect of compromised replication on Alu-IR -induced fragility	97
Table 4.2	PRR and checkpoint requirements for fragility in wild type and pol3-P664L strains	98
Table 4.S1	Length- and orientation-dependent size variations in expanded GAA and TTC repeat tracts.	99

LIST OF FIGURES

Figure 1.1	Secondary structures formed by inverted repeats and AT- and CG-rich micro and minisatellites.	3
Figure 1.2	Models of hairpin or cruciform-stimulated DSBs	7
Figure 2.1	Experimental systems to study GCRs resulting from hairpin-capped DSBs	20
Figure 2.2	Analysis of arm loss events triggered by hairpin-capped breaks.	22
Figure 2.3	Model for chromosomal rearrangements triggered by hairpin-capped DNA breaks.	24
Figure 2.4	Recurring instability of chromosomes with large palindromes derived from the $\text{Can}^{\text{R}}\text{Ade}^-$ strain	30
Figure 2.5	Analysis of extrachromosomal amplification events induced by hairpin-capped breaks.	33
Figure 2.6	Analysis of intrachromosomal amplification stimulated by hairpin-capped breaks	35
Figure 2.7	Model for generating intrachromosomal amplicons.	38
Figure 2.S1	Alu quasi-palindromes are located at the center of symmetry of the inverted duplications in strains with terminal deletions, and in strains with extrachromosomal and intrachromosomal amplicons.	53
Figure 2.S2	Analysis of non-reciprocal translocations in $\text{Can}^{\text{R}}\text{Ade}^-$ isolates	54
Figure 2.S3	Structural analysis of arm loss events in TP strains containing direct <i>Alus</i>	55
Figure 3.1	Experimental system to study intrachromosomal amplification	66

	resulting from hairpin-capped DSBs and the formation of dicentric intermediate	
Figure 3.2	Analysis of intrachromosomal amplification resulting from the breakage of the dicentric in $\square hdf2$, $\square rad51$ and $\square rad52$ mutants.	68
Figure 3.3	Model for BIR-mediated intrachromosomal amplification	71
Figure 4.1	Experimental systems to study chromosomal fragility induced by <i>Alu</i> -quasipalindromes.	84
Figure 4.2	2D gel analysis of replication fork progression in wild type and replication deficient strains with <i>Alu</i> -IRs.	86
Figure 4.3	Analysis of hairpin-capped DSB formation at the site of <i>Alu</i> -IRs using in-plug digestions with rarely-cutting restriction enzymes.	87
Figure 4.4	Replication-dependent and –independent mechanisms of cruciform-mediated DSBs.	93
Figure 4.S1	2D gel analysis of replication fork progression in <i>pol3-P664LΔmrc1</i> and <i>pol3-P664LΔcsm3</i> mutant strains.	100

LIST OF ABBREVIATIONS

<i>Alu</i> -IRs	inverted <i>Alu</i> repeats
BFB	breakage/fusion/bridge cycle
BIR	break-induced replication
CGH	comparative genome hybridization
CHEF	contour-clamped homogeneous electric field
DM	double minutes
DSB	double strand break
FISH	fluorescent <i>in situ</i> hybridization
HR	homologous recombination
HSR	homogeneously staining regions
NHEJ	nonhomologous end-joining
PRR	post-replication repair

SUMMARY

Chromosomal rearrangements play a major role in the evolution of eukaryotic genomes. Genomic aberrations are also a hallmark of many tumors and are associated with a number of hereditary diseases in humans. The presence of repetitive sequences that can adopt non-canonical DNA structures is one of the factors which can predispose chromosomal regions where they reside to instability. Palindromic sequences (inverted repeats with or without a unique sequence between them) that can adopt hairpin or cruciform structures are frequently found in regions that are prone for gross chromosomal rearrangements (GCRs) in somatic and germ cells in different organisms. Direct physical evidence was obtained that double-strand breaks (DSBs) occur at the location of long inverted repeats, a triggering event for the genomic instability. However, the mechanisms by which palindromic sequences lead to chromosomal fragility are largely unknown. The overall goal of this research is to elucidate the mechanisms of DSB and GCR generation by palindromic sequences in yeast, *Saccharomyces cerevisiae*.

CHAPTER 1

INTRODUCTION

1.1. The pros and cons of chromosomal rearrangements: contribution to evolution of eukaryotic genomes and human diseases.

DNA, the genetic material of living organisms, constantly incurs stress from exogenous agents including radiation and chemicals and from endogenous sources such as free radicals generated during essential metabolic processes. Chromosomal double strand break (DSB) is one of the detrimental outcomes of DNA damage. DSBs often lead to gross chromosomal rearrangements (GCRs) that can have either deleterious or advantageous consequences. Besides predisposing individuals to disease and cancer (see below), GCRs also have an evolutionary role.

The last two decades has seen a burst of technological developments that have enabled the scientists to perform high resolution analysis of genomes of many organisms including humans. Locus-specific fluorescent *in situ* hybridization, CGH microarray analyses using BAC and PAC clones have replaced cytogenetic techniques such as G-banding. These methods have been successfully employed to characterize deletions, duplications and translocations (Bruder et al., 2001; Shaw et al., 2004; Veltman et al., 2003). Global analyses of rearrangement breakpoints in human genome have unveiled the chromosomal loci enriched with repetitive sequences, such as low-copy repeats and AT-rich palindromes to be hotspots for rearrangements (reviewed in Shaw et al., 2004).

Comparisons of the genome sequences of related species suggests varying patterns of chromosomal rearrangements in different evolutionary lineages. One example is in the light of primate evolution. Karyotypic analyses of humans and our

closest ancestor, the chimpanzees reveal 98.8% identity (Navarro and Barton, 2003). However, several key chromosomal rearrangements (nine pericentric inversions and one acrocentric fusion) have aided in shaping the human karyotype (Navarro and Barton, 2003; Warburton et al., 1973). Rather than single nucleotide alterations, GCRs seemed to have to been the driving force for the evolution of the human genome.

GCRs and other karyotypic abnormalities are a hallmark of many cancers (Albertson, 2006; Hoeijmakers, 2001; Markowitz, 2000). Gene amplification plays a crucial role in the origin and development of tumors. In addition, while amplification of genes involved in metabolism of drugs can lead to appearance of cancer cells that are resistant to chemotherapeutical agents (Albertson et al., 2003). Chromosomal anomalies also cause hereditary diseases and are often associated with many human syndromes that arise due to defects in DNA repair genes (Hoeijmakers, 2001).

1.2. Hairpin- and cruciform adopting sequence motifs as sources of genome instability

Research over the recent years has demonstrated that chromosomal regions containing repetitive sequences that can adopt non-B DNA conformations are susceptible to breakage and subsequent aberrations. Human polymorphism research and studies from model organisms clearly indicate that DNA composition is a contributing factor to the maintenance of the genome integrity. These observations suggest a rather intriguing concept that GCRs are not random events, but arise as a consequence of instability mediated by the complex genomic architecture.

Hairpins and cruciforms are non-B DNA structures that are adopted by a broad range of repetitive sequences characterized by internal symmetry including inverted

Several physical and genetic parameters govern the extrusion of these non canonical structures. These include the repeat length, nucleotide composition of the sequence, length of the spacer DNA separating the repeats, location of the repeat relative to the replication origin, genomic location and the genetic background (reviewed in Ehrlich, 1989; Leach, 1994; Lewis and Cote, 2006; Sinden, 1994). Unlike hairpins, the formation of cruciform that requires a prenucleation step involving unwinding of 10 bp at the center of symmetry (Zheng and Sinden, 1988) is thermodynamically and energetically unfavorable. However, in negatively supercoiled DNA cruciforms can be adopted stably without the thermodynamic constraints. *In vivo*, processes such as transcription and replication that involve the separation of two strands can provide hypernegative supercoiling that can drive the extrusion and also stabilize the four way junctions (Potaman and Sinden, 2004 ; Sinden, 1994). Besides inverted repeats, expanded tracts of trinucleotide repeats such as CTG/CAG and CCG/CGG can also adopt stable hairpin and cruciform structures (Sinden, 1994).

Early studies have demonstrated that IRs are more wide-spread in the eukaryotic genomes than in prokaryotes (Cox and Mirkin, 1997; Schroth and Ho, 1995). In higher eukaryotes, palindromic sequences of about 200–800 bp exist at sites of some inherited chromosomal rearrangements in humans. For example, the sites of recurrent translocation between chromosome 22 and 11 harbor palindromic AT-rich repeats (PATRRs) (Edelmann et al., 2001; Kurahashi and Emanuel, 2001). Inverted repeats can also arise as a consequence of transposition event. Mobile elements can integrate adjacent to one another in inverted orientation giving rise to palindromic symmetry. A direct connection between hairpin-and cruciform-adopting sequence motifs and carcinogenic rearrangements was established by several studies.

First, Tanaka et al., demonstrated that palindromes are wide-spread in a variety of human solid tumors and can predispose the regions for gene amplifications or translocations (Tanaka et al., 2005). Two genes identified in this study that contain palindromes, are associated with translocations in T-cell leukemia/ lymphoma 1A (*TCL1A*) and myeloid leukemia factor 1 (*MLF1*) indicating that a similar phenomenon can happen in hematologic neoplasms. Recently, this has been shown to be the case by Neiman et al (Neiman et al., 2006). Palindromes were readily detected in both preneoplastic-transformed follicles and end-stage B-cell lymphoma cells carrying amplified loci. These data comply with early works where these sequence motifs were found to be a part of the amplified c-myc locus in HL-60 human leukemia cells (Feo et al., 1996; Mangano et al., 1998).

Overall, these data indicate that hairpin- and cruciform-forming sequence motifs are potent sources of genome instability. The physical link between the secondary structure forming repeats and their ability to induce GCRs came from studies in model organisms that demonstrated the DSB formation at the location of non-B structures. Several models have been proposed to explain the mechanism of fragility and consequent chromosomal aberrations (see below).

1.3. Mechanisms of palindrome-mediated chromosomal fragility

1.3.1. Replication arrest at the location of hairpin structures and subsequent breakage.

Hairpin secondary structures when adopted on the lagging strand template or in the displaced flap of Okazaki fragment can obstruct the progression of replication fork (**Figure 1.2.A**). The arrested fork can then be processed to DSBs. Several mechanisms are conceivable to explain the breakage at the site of hairpin-arrested

fork. First, the single stranded DNA that exists at the vicinity of the halted replisome can be subjected to sheer mechanical fracture. Second, the unusual structures that are generated at the fork stall site can be actively targeted by structure specific nucleases. One likely candidate to be processed is the hairpin. Prokaryotes and eukaryotes have evolved nucleases that specifically attack stem loop structures. In bacteria, SbcCD can cleave the hairpins formed by inverted repeats, a process that depends on DNA replication (Eykelboom et al., 2008). Eukaryotic homologs of bacterial SbcD and SbcC are Mre11 and Rad50 respectively (Sharples and Leach, 1995). *In vitro*, yeast Mre11/Rad50/Xrs2 (MRX) and human Mre11/Rad50/Nbs1 complexes that possess both endonuclease and exonuclease activities can actively cleave hairpin substrates (Paull and Gellert, 1999; Trujillo et al., 2003; Trujillo and Sung, 2001). Studies in *S.cerevisiae* have verified the hairpin-processing capabilities of the MRX complex. Lobachev et al demonstrated that the endonuclease activity of Mre11/Rad50/Xrs2 complex and function of the Sae2 protein are required for the inverted *Alu*-mediated homologous recombination (Lobachev et al., 2002). Analysis of DSB intermediates in wild type and in *mrx* and *sae2* mutants revealed that the Mre11 complex was needed to process the hairpin-capped broken fragments. In a recent study, Lengsfeld et al showed that Sae2 exhibits endonuclease activity and acts in conjunction with the MRX complex to process hairpin-DNA (Lengsfeld et al., 2007). Third, the arrested replication fork can be converted to Holliday junctions via fork reversal. The four-way junctions can then be cleaved by resolvases. In *Escherichia coli*, RuvABC complex resolves the Holliday junctions generated at the site of stalled fork (Michel 2000). There are no known eukaryotic homologs of RuvC resolvase that function in the nucleus (reviewed in (Heyer et al., 2003)). However, mitochondrial cruciform cutting endonucleases that cleaves Holliday junctions formed during recombination of

mitochondrial DNA have been discovered in both *S.cerevisiae* (Cce1) and *S.pombe* (Ydc1) (Kleff et al., 1992). These proteins act specifically on the mitochondrial DNA and lack a nuclear role. Yet another candidate that was suggested to function as a eukaryotic Holliday junction resolvase is the Mus81/Mms4 complex (Kleff et al., 1992). But analysis of *Alu*-stimulated recombination and DSB formation in *mus81* or *mms4* mutants argue against this possibility (Lobachev et al., 2002 and unpublished data).

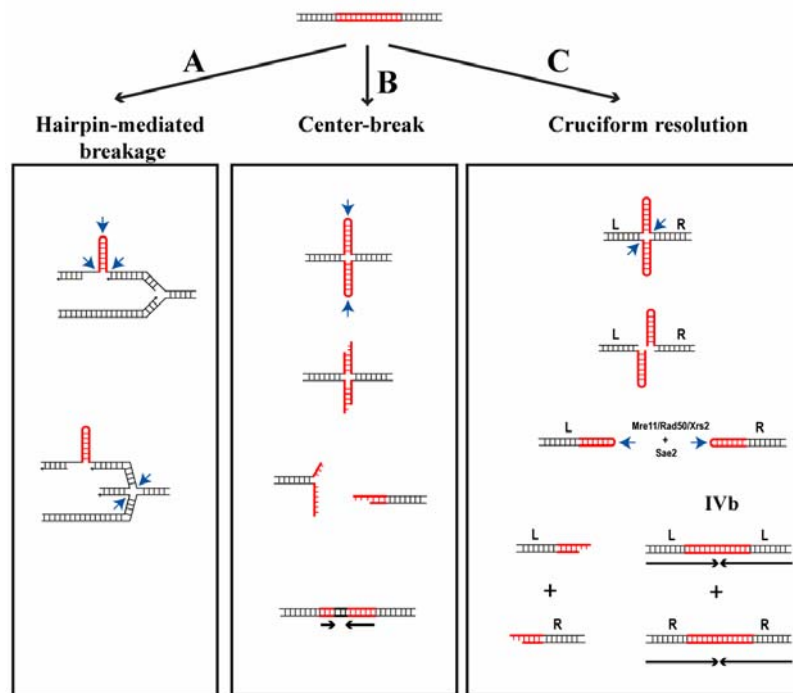


Figure 1.2. Models of hairpin or cruciform-stimulated DSBs. Blue arrows indicate putative or experimentally proven nuclease cleavage sites. Red color depicts hairpin/cruciform forming sequence motif. Horizontal arrows correspond to the repeats in rearranged molecules. L and R in cruciform resolution panel denote left and right regions flanking the secondary-structure forming sequence. A detailed description is presented in the text.

The existence of Holliday junction resolving protein (besides Mus81/Mms4) in mitotically dividing cells has been shown by genetic and biochemical experiments.

The identity of this putative protein is still largely unknown and its contribution to secondary structure-mediated fragility is yet to be determined.

1.3.2. Center-Break palindrome revision mechanism

The schematic of the center-break model is illustrated in **Figure 1.2.B**. The initial step in the model involves the isomerization of inverted repeats from linear duplex DNA to cruciform structure. As discussed previously, cruciform extrusion is an energetically unfavorable process (Potaman and Sinden, 2004; Sinden, 1994).

However, the increased negative superhelicity that promote cruciform extrusion can be provided by cellular processes that alter chromatin organization. The cell responds to the cruciform formation by recruiting enzymes that nick the hairpin ends of the structure. The degradation of the opened hairpins will then generate a linear broken DNA that can be repaired by either non-homologous end joining or homologous recombination. The center-break revision model was proposed to explain the asymmetry in the center of the recovered stabilized molecules after long palindromes were introduced into mouse or hamster tissue culture cells (Cunningham et al., 2003). Several key candidate nucleases that can participate in this mechanism can be envisioned. These include the MRX complex along with Sae2 endonuclease (see above), *RAG1/RAG2* and Artemis that nick hairpins formed during V(D)J recombination in mammals (Besmer et al., 1998; Havas et al., 2000) and *ERCC1-XPF*, structure specific nuclease that participates in nucleotide excision pathway (de Laat et al., 1998).

An alternative model that involves the resolution of the extruded cruciform structure was proposed based on studies done in yeast model systems (see below).

1.3.3. Cruciform resolution model

The initiating lesion in this model is also an extruded cruciform structure. The cruciform junction is attacked by putative nuclease that introduces symmetrical cuts on opposite sides of the four-way junction (**Figure 1.2.C**). Detailed genetic and structural analyses in yeast *S.cerevisiae*, demonstrated that 320 bp *Alu*-quasi palindrome induce DSBs containing hairpin-capped termini (Lobachev et al., 2002). Upon DSB generation, Mre11 complex along with Sae2 target and process the hairpin ends. The resected fragments can then be repaired via homologous recombination. In *mre11* mutants, unprocessed hairpin-capped DSB molecules accumulate giving rise to highly unstable inverted dimers and subsequent chromosomal rearrangements. It is conceivable that a similar mechanism albeit involving NHEJ repair pathway, might also exist in higher eukaryotes including humans, given that many cellular processes are conserved from yeast to humans.

The cruciform secondary structure closely resembles the four-way Holliday junction intermediate that occurs during homologous recombination process. It is quite possible that Holliday junction-resolving enzymes might misrecognize cruciform structures and process them. As mentioned previously, the identity of the proteins that possess such an activity is still unknown. A closer look at the model reveals that the generation of hairpin-capped DSB molecule necessitates self-rejoining of the broken end following resolution. Many proteins in both prokaryotes and eukaryotes are known to have such a breakage and rejoining activity. One group of proteins function in prokaryotic organisms that replicate their linear genomes via hairpin-ended intermediates rather than telomeres (Kobryn and Chaconas, 2002). The breakage and hairpin formation, a process termed telomere resolution in these organisms is essential to ensure conversion of the circular inverted dimers to linear DNA following replication. Examples include ResT in *Borrelia burgdorferi*, TelN in

Escherichia coli phage N15, Tel PY54 of the *Yersinia enterocolitica* phage PY54, and TelK of the *Klebsiella* phage Φ KO2, enzymes related to tyrosine recombinases and type IB topoisomerases (Deneke et al., 2002; Hertwig et al., 2003; Huang et al., 2004; Kobryn and Chaconas, 2002). The second group of enzymes, is comprised of prokaryotic and eukaryotic transposases, the V(D)J recombinase *RAG1*, and retroviral integrases. These processes involve the formation of a hairpin-ended molecule during the breakage and rejoining process (Zhou et al., 2004).

The eukaryotic protein that carries our cruciform resolving function could have evolved from any of the above described classes of proteins. Identification and characterization of the enzymatic activity of the protein will help in the better understanding of the palindrome-associated instabilities.

1.4. Implications and concluding remarks

With the growing body of data indicating that fragility at the location of secondary structure forming repeats is a contributing factor to eukaryotic genome instability, this area of research is gaining increasing attention. Elucidation of the molecular mechanisms underlying chromosomal fragility mediated by unstable repeats is clearly important for studying the predisposition of different individuals to diseases, the origin of inherited disorders, the cancer diagnostics and treatments as well as for our understanding of the fundamental processes that determine the architecture and dynamics of eukaryotic genomes.

1.5. References

- Albertson, D. G., C. Collins, F. McCormick & J. W. Gray: Chromosome aberrations in solid tumors. *Nat Genet*, 34, 369-76(2003).
- Besmer, E., Mansilla-Soto, J., Cassard, S., Sawchuk, D.J., Brown, G., Sadofsky, M., Lewis, S.M., Nussenzweig, M.C., and Cortes, P. (1998). Hairpin coding end opening is mediated by RAG1 and RAG2 proteins. *Molecular cell* 2, 817-828.
- Bruder, C.E., Hirvela, C., Tapia-Paez, I., Fransson, I., Segraves, R., Hamilton, G., Zhang, X.X., Evans, D.G., Wallace, A.J., Baser, M.E., *et al.* (2001). High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH. *Human molecular genetics* 10, 271-282.
- Consortium, I.H.G.S. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Cox, R., and Mirkin, S.M. (1997). Characteristic enrichment of DNA repeats in different genomes. *Proceedings of the National Academy of Sciences of the United States of America* 94, 5237-5242.
- Cunningham, L.A., Cote, A.G., Cam-Ozdemir, C., and Lewis, S.M. (2003). Rapid, stabilizing palindrome rearrangements in somatic cells by the center-break mechanism. *Mol Cell Biol* 23, 8740-8750.
- de Laat, W.L., Sijbers, A.M., Odijk, H., Jaspers, N.G., and Hoeijmakers, J.H. (1998). Mapping of interaction domains between human repair proteins ERCC1 and XPF. *Nucleic acids research* 26, 4146-4152.
- Deneke, J., Ziegelin, G., Lurz, R., and Lanka, E. (2002). Phage N15 telomere resolution. Target requirements for recognition and processing by the protelomerase. *J Biol Chem* 277, 10410-10419.
- Edelmann, L., Spiteri, E., Koren, K., Pulijaal, V., Bialer, M.G., Shanske, A., Goldberg, R., and Morrow, B.E. (2001). AT-rich palindromes mediate the constitutional t(11;22) translocation. *American journal of human genetics* 68, 1-13.

Ehrlich, D.S. (1989). Illegitimate recombination in bacteria. In *Mobile DNA*, D.E. Berg, and M.M. Howe, eds. (Washington D. C., American Society for Microbiology), pp. 799-832.

Eykelenboom, J.K., Blackwood, J.K., Okely, E., and Leach, D.R. (2008). SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. *Molecular cell* 29, 644-651.

Feo, S., Di Liegro, C., Mangano, R., Read, M., and Fried, M. (1996). The amplicons in HL60 cells contain novel cellular sequences linked to MYC locus DNA. *Oncogene* 13, 1521-1529.

Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P.A., Lilley, D.M., and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell* 103, 1133-1142.

Hertwig, S., Klein, I., Lurz, R., Lanka, E., and Appel, B. (2003). PY54, a linear plasmid prophage of *Yersinia enterocolitica* with covalently closed ends. *Mol Microbiol* 48, 989-1003.

Heyer, W.D., Ehmsen, K.T., and Solinger, J.A. (2003). Holliday junctions in the eukaryotic nucleus: resolution in sight? *Trends Biochem Sci* 28, 548-557.
Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366-374.

Huang, W.M., Joss, L., Hsieh, T., and Casjens, S. (2004). Protelomerase uses a topoisomerase IB/Y-recombinase type mechanism to generate DNA hairpin ends. *J Mol Biol* 337, 77-92.

Kleff, S., Kemper, B., and Sternglanz, R. (1992). Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease. *Embo J* 11, 699-704.

Kobryn, K., and Chaconas, G. (2002). ResT, a telomere resolvase encoded by the Lyme disease spirochete. *Molecular cell* 9, 195-201.

Kurahashi, H., and Emanuel, B.S. (2001). Long AT-rich palindromes and the constitutional t(11;22) breakpoint. *Human molecular genetics* 10, 2605-2617.

Leach, D.R. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* 16, 893-900.

Lengsfeld, B.M., Rattray, A.J., Bhaskara, V., Ghirlando, R., and Paull, T.T. (2007). Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Molecular cell* 28, 638-651.

Lewis, S.M., and Cote, A.G. (2006). Palindromes and genomic stress fractures: bracing and repairing the damage. *DNA Repair (Amst)* 5, 1146-1160.

Lobachev, K.S., Gordenin, D.A., and Resnick, M.A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* 108, 183-193.

Mangano, R., Piddini, E., Carramusa, L., Duhig, T., Feo, S., and Fried, M. (1998). Chimeric amplicons containing the c-myc gene in HL60 cells. *Oncogene* 17, 2771-2777.

Markowitz, S. (2000). DNA repair defects inactivate tumor suppressor genes and induce hereditary and sporadic colon cancers. *J Clin Oncol* 18, 75S-80S.
Navarro, A., and Barton, N.H. (2003). Chromosomal speciation and molecular divergence--accelerated evolution in rearranged chromosomes. *Science (New York, NY)* 300, 321-324.

Neiman, P.E., Kimmel, R., Icreverzi, A., Elsaesser, K., Bowers, S.J., Burnside, J., and Delrow, J. (2006). Genomic instability during Myc-induced lymphomagenesis in the bursa of Fabricius. *Oncogene* 25, 6325-6335.

Paull, T.T., and Gellert, M. (1999). Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev* 13, 1276-1288.
Potaman, V.N., and Sinden, R.R. (2004). DNA: alternative conformations and biology. In *DNA Conformation and Transcription*, T. Ohyama, ed. (Landes Bioscience), pp. 1-16.

Schroth, G.P., and Ho, P.S. (1995). Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. *Nucleic acids research* 23, 1977-1983.

Sharples, G.J., and Leach, D.R. (1995). Structural and functional similarities between the SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Mol Microbiol* 17, 1215-1217.

Shaw, C.J., Shaw, C.A., Yu, W., Stankiewicz, P., White, L.D., Beaudet, A.L., and Lupski, J.R. (2004). Comparative genomic hybridisation using a proximal 17p BAC/PAC array detects rearrangements responsible for four genomic disorders. *Journal of medical genetics* 41, 113-119.

Sinden, R.R. (1994). *DNA Structure and Function* (San Diego and London, Academic Press, Inc).

Tanaka, H., Bergstrom, D.A., Yao, M.C., and Tapscott, S.J. (2005). Widespread and nonrandom distribution of DNA palindromes in cancer cells provides a structural platform for subsequent gene amplification. *Nat Genet* 37, 320-327.

Trujillo, K.M., Roh, D.H., Chen, L., Van Komen, S., Tomkinson, A., and Sung, P. (2003). Yeast *xrs2* binds DNA and helps target *rad50* and *mre11* to DNA ends. *J Biol Chem* 278, 48957-48964.

Trujillo, K.M., and Sung, P. (2001). DNA structure-specific nuclease activities in the *Saccharomyces cerevisiae* Rad50*Mre11 complex. *J Biol Chem* 276, 35458-35464.

Veltman, J.A., Jonkers, Y., Nuijten, I., Janssen, I., van der Vliet, W., Huys, E., Vermeesch, J., Van Buggenhout, G., Fryns, J.P., Admiraal, R., *et al.* (2003). Definition of a critical region on chromosome 18 for congenital aural atresia by arrayCGH. *American journal of human genetics* 72, 1578-1584.

Wang, Y., and Leung, F.C. (2006). Long inverted repeats in eukaryotic genomes: recombinogenic motifs determine genomic plasticity. *FEBS letters* 580, 1277-1284.

Warburton, D., Firschein, I.L., Miller, D.A., and Warburton, F.E. (1973). Karyotype of the chimpanzee, *Pan troglodytes*, based on measurements and banding pattern: comparison to the human karyotype. *Cytogenetics and cell genetics* 12, 453-461.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., *et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520-562.

Zheng, G.X., and Sinden, R.R. (1988). Effect of base composition at the center of inverted repeated DNA sequences on cruciform transitions in DNA. *J Biol Chem* 263, 5356-5361.

Zhou, L., Mitra, R., Atkinson, P.W., Hickman, A.B., Dyda, F., and Craig, N.L. (2004). Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature* 432, 995-1001.

CHAPTER 2

The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks

2.1. Summary

DNA palindromes often co-localize in cancer cells with chromosomal regions that are predisposed to gene amplification. The molecular mechanisms by which palindromes can cause gene amplification are largely unknown. Using yeast as a model system, we found that hairpin-capped double strand breaks (DSBs) occurring at the location of human *Alu* quasi-palindromes lead to the formation of intrachromosomal amplicons with large inverted repeats (equivalent to homogeneously-staining regions in mammalian chromosomes) or extrachromosomal palindromic molecules (equivalent to double minutes in mammalian cells). We demonstrate that the specific outcomes of gene amplification depend on the applied selection, the nature of the break, and the chromosomal location of the amplified gene relative to the site of the hairpin-capped DSB. The rules for the palindrome-dependent pathway of gene amplification defined in yeast may operate during the formation of amplicons in human tumors.

2.2. Introduction

Amplification of chromosomal regions plays an important role in tumor pathogenesis. Increase in copy number of oncogenes is critical for the occurrence and progression of a variety of solid tumors, while amplification of genes involved in metabolism of drugs can lead to appearance of cancer cells that are resistant to chemotherapeutical agents (Albertson et al., 2003; Fletcher, 2005; Naeem, 2005). Two types of amplification events have been detected cytogenetically: extra- and intrachromosomal amplicons (Debatisse, 2005; Stark et al., 1989; Windle and Wahl, 1992). Extrachromosomally amplified molecules or double minutes (DM) have up to several hundred copies of a genomic segment, forming mini-chromosomes with inverted symmetry. Intrachromosomal amplicons described as abnormally banded or homogeneously staining regions (HSR) can be organized either as head to tail or tail to tail tandem repeats, typically represented at early stages of amplification by less than 10 copies. In addition, some gene amplifications are accompanied by other chromosome aberrations such as aneuploidy, deletions or translocations (Albertson et al., 2003; Fletcher, 2005; Naeem, 2005).

The common feature of models to explain gene amplification is that the initiating step that involves double-strand break (DSB) formation. In yeast and mammalian cells, the induction of DSBs with *I-SceI* or HO endonucleases increases the frequency of extrachromosomal and intrachromosomal amplification (Coquelle et al., 2002; Pipiras et al., 1998; Watanabe and Horiuchi, 2005). Amplification is also induced as a result of treatment of cells with DNA damaging agents which can directly or indirectly cause DSBs (Kuo et al., 1994; Paulson et al., 1998; Poupon et al., 1996; Yunis et al., 1987). In addition, in human and rodent cells, chromosomal regions containing fragile sites, that are natural hot-spots for breakage and

recombination, are highly susceptible to amplification and are found to frame the early amplicons (reviewed in Debatisse, 2005). DSBs can trigger gene amplification through a variety of mechanisms which include unequal sister chromatid exchange, rolling-circle replication, break-induced replication, foldback priming processes, and the breakage-fusion-bridge (BFB) cycle (Kobayashi et al., 2004; Kraus et al., 2001; McClintock, 1941; Rattray et al., 2005; Watanabe and Horiuchi, 2005).

The BFB cycle (McClintock, 1941) is the most popular model to explain intrachromosomal amplicons. The HSRs in cancer cells are often organized as an inverted ladder and are associated with a deletion that spans from the location of the amplicon towards a telomere (Debatisse, 2005). In accordance with the BFB model, such a complex rearrangement is a result of an initial DSB followed by replication of the broken molecule, fusion of sister chromatids, formation of a bridge during anaphase, asymmetrical breakage due to mechanical tension generating one chromatid with an inverted repeat at the broken end, and subsequent repetition of the cycle. Although the fusion of broken chromatids has not yet been directly demonstrated, this step seems likely because of the robust non-homologous end joining machinery in mammalian cells. Anaphase bridges are often observed in cells undergoing gene amplification, implicating a role of the BFB cycle during this process (Coquelle et al., 1997; Ma et al., 1993; Shimizu et al., 2005; Toledo et al., 1993). In addition, end-to-end chromosome fusions are well documented in cells where telomere dysfunction leads to recognition of unprotected chromosome ends as DSBs (Chan and Blackburn, 2004).

Recently, it has been found that DNA palindromes are abundant in human cancer cells and often co-localize with the chromosomal regions that are predisposed to gene amplification (Tanaka et al., 2005). Palindromic sequences are implicated in

early steps of gene amplification and found to be hot-spots for other types of gross chromosomal rearrangements (GCRs) in many organisms (Fried et al., 1991; Zhou et al., 2001 and references therein). However, the molecular mechanisms by which palindromes can cause GCRs, including gene amplification, in eukaryotic genomes are poorly understood. Previously, we showed that, in the yeast *Saccharomyces cerevisiae*, a quasi-palindrome comprised of two human *Alu* repeats induces DSBs that are terminated by covalently-closed hairpins. Unprocessed DSBs lead to the formation of acentric and dicentric rearranged molecules characterized by inverted symmetry (Lobachev et al., 2002). In the present study, we developed yeast strains that allowed us to investigate the fate of these rearranged intermediates and to determine how they can be directed into chromosomal aberrations. We have found that chromosomal arm loss, extrachromosomal and intrachromosomal gene amplification events are different consequences of the hairpin-capped breaks. Using the yeast model system, we define a novel, palindrome-dependent pathway of gene amplification that can mimic the formation of oncogene amplicons in humans.

2.3. Results

2.3.1. Experimental system

To characterize the GCRs that result from secondary structure-mediated DSBs, we developed an experimental system based on the loss of *CANI* and *ADE2* genes, and amplification of *CUPI* and *SFAI* genes located on chromosome V. (Figure 2.1). Two sets of haploid yeast strains were constructed where the left arm of chromosome V in the region of *CANI* gene was modified. The system is based on the assay for GCR developed by R. Kolodner and colleagues (Chen and Kolodner, 1999). A *LYS2* cassette containing homologous and homeologous inverted or direct *Alu* repeats was

placed centromere-proximal to *CAN1*, such that the region between *LYS2* and the telomere does not contain essential genes and can be deleted. The *Alu* repeats are 320 bp long and separated by a 12 bp spacer. 100%, 94% and 86% identical inverted *Alu* repeats (*Alu*-IR) were used as a source of hairpin-capped breaks (Lobachev et al., 2002; Lobachev et al., 2000). As a control, 100% identical direct *Alu* repeats that cannot form a secondary structure were inserted at the same chromosomal location. The *ADE2* gene was moved telomere-distal to *CAN1*, while *CUP1* and *SFA1* genes were positioned either telomere-proximal (Figure 2.1, TP strains) or telomere-distal (Figure 2.1, TD strains) to *CAN1*.

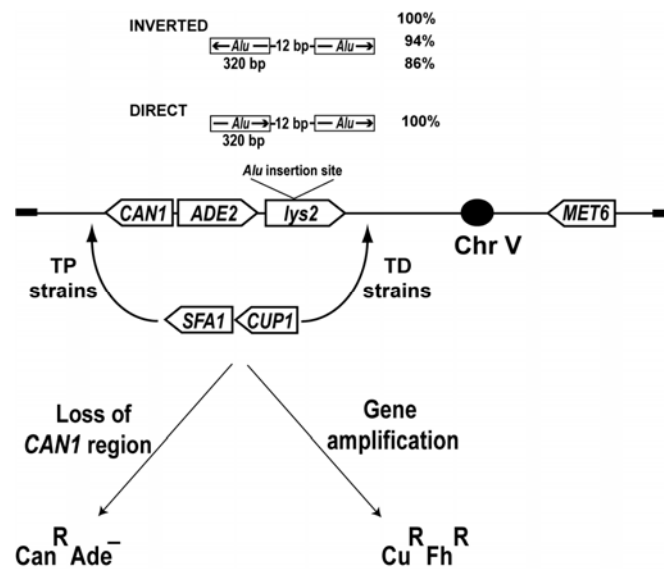


Figure 2.1. Experimental systems to study GCRs resulting from hairpin-capped DSBs. Two sets of strains with different modifications of chromosome V were constructed. Both sets contain the *ADE2* gene placed telomere-proximal and the *LYS2* gene moved telomere-distal to *CAN1*. 100, 94 and 86 % identical inverted and 100% identical direct *Alu* repeats (open boxes with arrows inside) were inserted into a *Bam*HI site of the *LYS2* gene. The TP strains contain the *CUP1* and *SFA1* genes next to 3' end of *CAN1* gene while the TD strains contain the *CUP1* and *SFA1* genes next to 3' end of *LYS2*. The *MET6* gene was used as a target to highlight the right arm of chromosome V in Southern analysis. The telomeres, centromere, and various relevant genes are shown as filled rectangles, a filled circle, and open arrows, respectively. Depending on the mode of selection, these strains can be used to detect deletions or amplifications. Selection for canavanine-resistant strains that are also Ade^- results in isolates with a 42 kb telomeric deletion. Selection for copper- and formaldehyde-resistant derivatives ($Cu^R Fh^R$) results in either extrachromosomal amplifications (in TP strains) or intrachromosomal amplifications (in TD strains) of the regions adjacent to the *Alu* repeats.

This system allows for selection of two types of GCRs induced by unstable motifs. First, a hairpin-capped break can cause deletion of the chromosome V region including *CAN1* and *ADE2* genes resulting in canavanine-resistant red colonies ($\text{Can}^{\text{R}}\text{Ade}^-$). Second, breakage at the location of the inverted repeat can cause amplification of the telomere-proximal or telomere-distal regions including *CUPI* (encoding copper chelatin) and *SFA1* (encoding formaldehyde dehydrogenase) genes that serve as convenient gene dosage markers (Resnick et al., 1990; van den Berg and Steensma, 1997). Clones carrying the amplified regions of chromosome V can be initially selected on medium containing a high concentration of copper (Cu^{R}) followed by replica plating the Cu^{R} colonies onto media with a high concentration of formaldehyde.

2.3.2. Inverted *Alu*s strongly induce arm loss events characterized by a specific pattern: terminal deletion coupled with adjacent inverted duplication

The *Alu*-IRs dramatically increased the rate of terminal deletion (see Table 2.S1 in the Supplemental Data available with this article online). The rate of *CAN1* region loss was nearly 25000 fold higher in strains containing 100% homologous inverted repeats than in control strains with direct repeats. The majority of these $\text{Can}^{\text{R}}\text{Ade}^-$ colonies were small in size (Figure 2.2A). 40 to 60% of the cells in these colonies were “large-budded” suggesting ongoing DNA damage in these isolates (Figure 2.2B). These colonies, when sub-cultured, gave rise to a heterogeneous population of single colonies of small and normal size (Figure 2.2C). Using CHEF gel electrophoresis, we examined the structure of chromosome V in $\text{Can}^{\text{R}}\text{Ade}^-$ clones selected initially on canavanine-containing media as well as from cultures of small and normal size colonies occurring in their sub-cultured progeny.

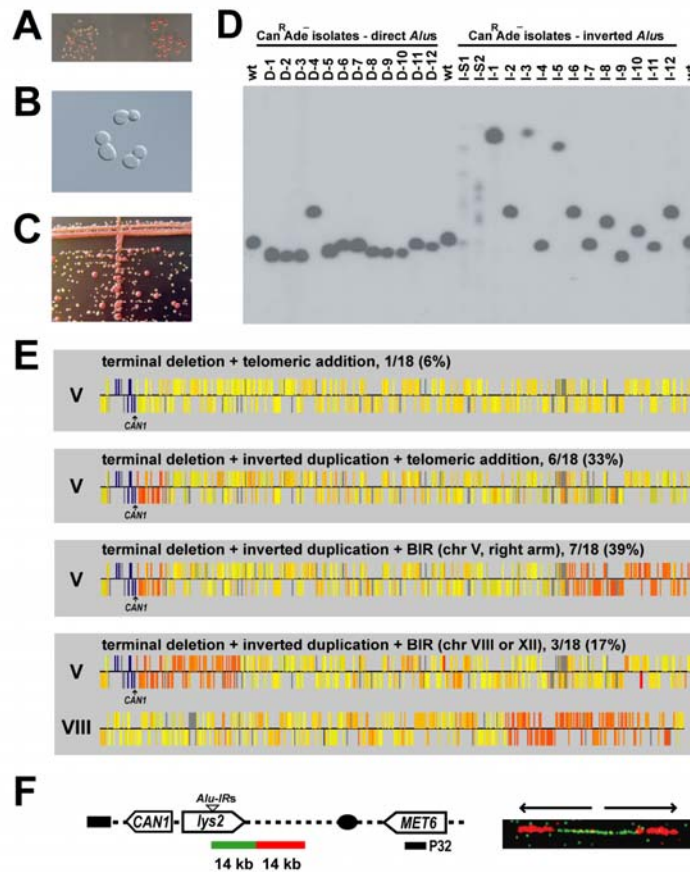


Figure 2.2. Analysis of arm loss events triggered by hairpin-capped breaks.

(A) Small size colonies of primary $\text{Can}^{\text{R}}\text{Ade}^-$ isolates from TP strains. $\text{Can}^{\text{R}}\text{Ade}^-$ colonies arising from TP strain (left side of the plate) were grown adjacent to colonies from isogenic strain carrying point mutations in *CAN1* gene (right side of the plate).

(B) Cells of the primary $\text{Can}^{\text{R}}\text{Ade}^-$ isolates often are arrested with large buds, characteristic of a DNA damage checkpoint response.

(C) The primary $\text{Can}^{\text{R}}\text{Ade}^-$ isolates give rise to a heterogeneous population of small and normal size colonies.

(D) Analysis of rearranged chromosome Vs in $\text{Can}^{\text{R}}\text{Ade}^-$ isolates by CHEF gel electrophoresis and Southern blot hybridization. Lanes D-1 to D-12 are $\text{Can}^{\text{R}}\text{Ade}^-$ isolates from TP strains with direct *Alus*. Lanes I-S1 and I-S2 are small size primary $\text{Can}^{\text{R}}\text{Ade}^-$ isolates, and lanes I-1 to I-12 are normal size $\text{Can}^{\text{R}}\text{Ade}^-$ isolates from TP strains with inverted *Alus*. Lanes labeled with “wt” are TP strains with the wild type chromosome V. Chromosome V was detected using a *MET6*-specific probe.

(E) Microarray analysis of GCRs stimulated by hairpin-capped DNA breaks. The DNA microarrays used in this experiment contained almost all yeast genomic ORFs. Color coding is as follows: gray, not present on the array; yellow, single copy sequences; red, repeated sequences; blue, deletions. Only those chromosomes that had a deletion or duplication are shown in this figure. Complete data for these experiments is online at <https://genome.unc.edu/>. From the top panel to the bottom, the classes of rearrangements (number of isolate in parentheses) are: Class I (I-9), Class II (I-7), Class III (I-12), and Class IV (I-16).

(F) The duplicated regions adjacent to terminal deletions are organized as inverted repeats. The left panel shows the regions on chromosome V that are fluorescently-labeled for FISH, and ^{32}P -labeled for Southern analysis. The right panel is an example of an inverted duplication visualized by molecular combing (isolate I-8). Arrowheads above the panel depict the repeat units in the amplicon.

In small size colonies, chromosome V was undetectable or had diminished intensity in ethidium bromide-stained CHEF gels (data not shown). Hybridization with the chromosome V-specific probe revealed the presence of multiple bands of different sizes (Figure 2.2D, I-S1 and I-S2 isolates). In each of the derived normal size colonies, a discrete chromosomal band was detected (Figure 2.2D, I-1 to I-12 isolates). The cells in these colonies had the morphology characteristic of logarithmically-growing culture (data not shown). These results suggest that arm loss events lead to the generation of a mixed population of cells carrying unhealed broken chromosome Vs. These broken chromosomes are detected by DNA damage checkpoints, resulting in the characteristic G2/M arrest phenotype. In the small colonies, the lack of a single species of healed chromosome also results in the absence of a discrete chromosome V band. Subsequent repair of the broken molecules leads to formation of rearranged chromosomes (as described below) and recovery from the arrest resulting in normal size colonies.

Most of the rearranged chromosome Vs in $\text{Can}^{\text{R}}\text{Ade}^-$ normal sized isolates were equal to or larger than the unrearranged chromosome V. This result suggests that the loss of the 42 kb telomeric region adjacent to the DSB site was accompanied by a gain of genetic material from elsewhere (right panel of Figure 2.2D). To determine the structure of chromosomal rearrangements in these strains, we examined genomic DNA from 18 independent $\text{Can}^{\text{R}}\text{Ade}^-$ isolates using comparative genomic hybridization (CGH) on microarrays (Figure 2.2E). Based on the CGH and CHEF gel analyses, we group the isolates into four classes. Class I (I-9) had a chromosome V that was about 40 kb smaller than the wild-type (Figure 2.2D) and, by CGH analysis had a terminal deletion of V with a breakpoint near the *CANI* locus (Figure 2.2E). This pattern is consistent with formation of a DSB near the inverted repeat, followed

by resection of the broken end and *de novo* addition of telomeric sequences. The initiating DSB for I-9 and all of the other rearrangements is likely to reflect processing of an extruded cruciform containing the *Alu* repeats (Figure 2.3).

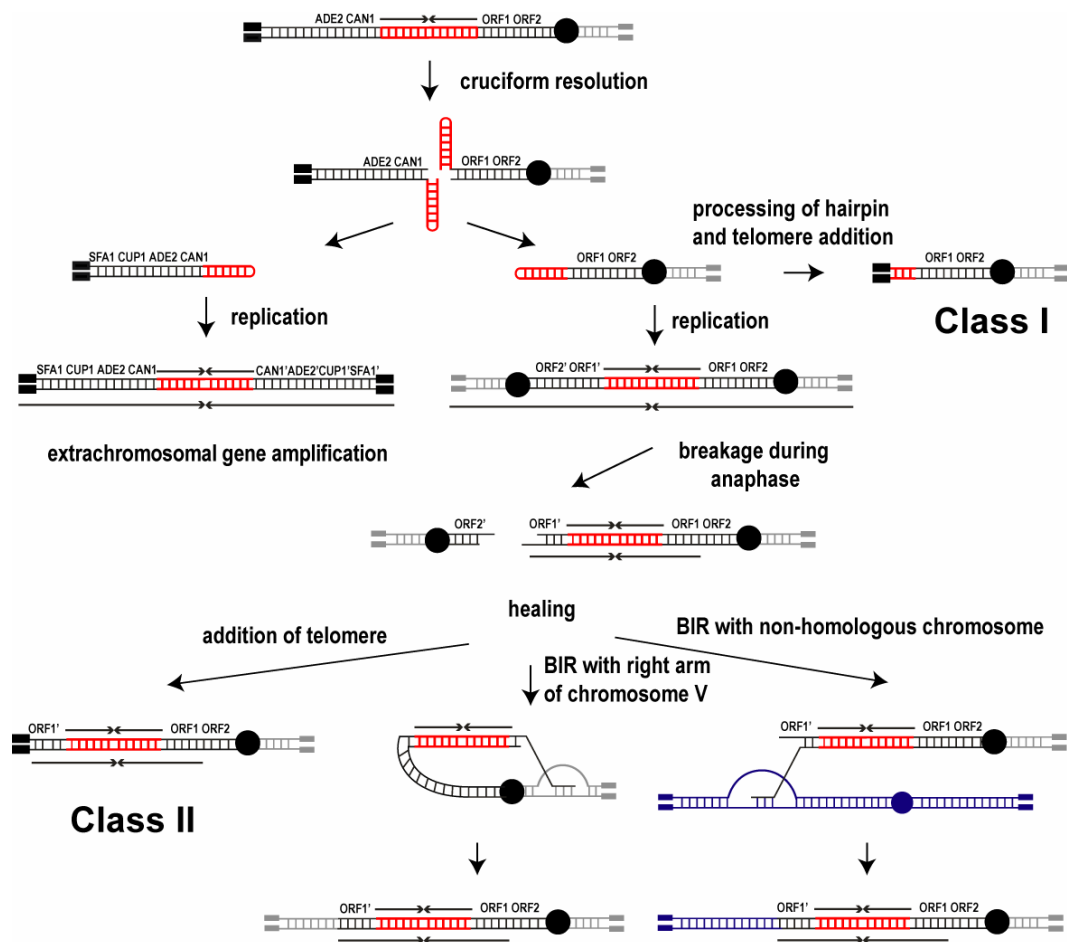


Figure 2.3. Model for chromosomal rearrangements triggered by hairpin-capped DNA breaks. The *Alu* quasi-palindrome is depicted (not to scale) as red arrows. Telomeres (filled rectangles) and centromeres (filled circles) are also shown. Right arm of chromosome V is represented in gray color. A non-homologous chromosome is indicated by blue color. A detailed description is presented in the text. Classes I-IV refer to the various classes of chromosomes with a deletion of *CAN1*, as described in the text. Intrachromosomal amplification is not shown in this figure, but is presented in Figure 2.7.

Class II isolates (I-4, I-7, I-10, I-11, I-13, I-14, I-18) had a deletion of the region of chromosome V centromere-distal to *CAN1*, and a duplication of about 30 kb located centromere-proximal to *CAN1* (Figure 2.2E). The sizes of chromosome V in six of the seven Class II isolates were slightly smaller than the wild-type V,

suggesting that the duplication of 30 kb is within chromosome V, since the duplication of 30 kb compensates for the deletion of 40 kb. For six of the seven Class II isolates, the duplication breakpoint is within a region of about one kb that contains *YELWdelta1* and *YELWdelta2* and *PAU2* (a repetitive sequence found near the telomeres of most chromosomes; Rachidi et al., 2000). As in Class I, the likely initiating event for the Class II isolates is a DSB at the extruded *Alu* cruciform, followed by replication of the resulting hairpin-capped molecule, generating a dicentric chromosome (Figure 2.3). For the Class II isolates, we suggest that the dicentric breaks near the region of the chromosome with the delta elements and *PAU2*. The resulting broken end is either capped by *de novo* telomere addition or repaired by break-induced replication (BIR) (Malkova et al., 1996) utilizing *PAU2*-related or delta-related repeats located near the telomeres of other chromosomes (Figure 2.3).

In most of the Class III isolates (I-1, I-2, I-6, I-8, I-12, I-15, I-17), there was a 30 kb duplication adjacent to the deletion, as observed in Class II. In addition, a region on the right arm of chromosome V was duplicated. The size of the region duplicated on the right arm was somewhat variable. In the isolate shown in Figure 2.2E (I-12), the duplicated region was about 130 kb, with a breakpoint near *YERCTy1-1*; three other Class III isolates (I-6, I-8, and I-17) shared this breakpoint. In two other isolates (I-2 and I-15), the duplicated region on the right arm of V was at *YERCTy1-2*. Thus, one simple way of explaining most of the Class III isolates is that the delta elements located near the breakpoint of the 30 kb duplication (*YELWdelta1* and *YELWdelta2*) initiate a BIR event with a delta or Ty element located on the right arm of chromosome V, resulting in the duplication (Figure 2.3).

In Class IV isolates, a deletion of the sequences centromere-distal to *CANI* is adjacent to a duplication of 30 (I-3) or 100 (I-5 and I-16) kb; both of these breakpoints are near delta elements (*YELdelta1,2* and *YELdelta4-6*, respectively). In addition, in these isolates, duplications of sequences derived from another homologue were observed: chromosome XII in I-3 and I-5, and chromosome VIII in I-16 (shown in Figure 2E). The breakpoints of the duplications on the other chromosomes are near delta elements (delta elements 9-12 on chromosome XII for I-3, delta elements 18 and 19 for I-5, and delta elements 10 and 11 for I-16). The simplest explanation of these strains is that the delta elements at the breakpoints of chromosome V were repaired by a BIR event using delta elements from non-homologous chromosomes.

As a further confirmation of chromosome rearrangements mediated by delta-delta recombination, we used PCR to demonstrate the postulated structure for one Class III isolate (I-6) and two Class IV isolates (I-3 and I-16) (Figures 2.S2A and 2.S2B). The translocations postulated for the Class IV isolates were also confirmed by CHEF gel analysis and Southern blot hybridization both with *MET6* probe and with a probe that hybridizes with the arm of the presumed donor chromosome (Figure 2.S2C). BIR events capable of duplicating very large segments of chromosomes were observed previously by Malkova et al. (1996), and BIR events involving transposable elements that generated non-reciprocal translocations were detected by Lemoine et al. (2005).

The model shown in Figure 2.3 predicts that the intrachromosomal duplication at the left end of chromosome V in Class II, III, and IV isolates will form a large (30 kb to 100 kb) quasi-palindrome with the inverted *Alu* repeats in the middle. This prediction was confirmed in two ways. First, we used DNA fiber fluorescent *in situ* hybridization (FISH) and molecular combing (Conti, 2001) to examine the DNA for

one Class II isolate (I-7), three Class III isolates (I-6, I-8, and I-12), and two Class IV isolates (I-5 and I-10). For this procedure, we PCR-amplified DNA fragments within two 14 kb chromosomal DNA regions located centromere-proximal to the *Alu* repeats (positions shown in Figure 2.2F). One fragment was labeled with biotin-tagged dUTP and the other with digoxigenin-dUTP. Chromosomal DNA from each strain was stretched out on siliconized cover slips and the DNA was hybridized to the labeled fragments. The incorporated biotin and digoxigenin nucleotides were detected with fluorescent-conjugated antibodies containing red and green fluorochromes, respectively. All of the isolates examined had a palindromic pattern (shown in Figure 2.2F for I-8), a green region flanked by two red regions that are approximately half of the size of the green region. We also confirmed the structure by Southern analysis (Figure 2.S1).

In summary, these results show that hairpin-capped breaks primarily induce GCRs that have a very specific pattern: terminal deletion coupled with an adjacent inverted duplication (large quasi-palindromes). This pattern of chromosomal rearrangements is different from GCR events that result from spontaneous or damage-induced breaks analyzed in other yeast studies where *CANI* was used as a reporter (Myung and Kolodner, 2003; Putnam et al., 2005 and references therein).

2.3.3. Chromosome rearrangements in strains with direct *Alu* repeats

In order to compare chromosomal rearrangements induced by unstable motifs with GCRs that resulted from spontaneous DSBs, we carried out structural analysis of the chromosome V in 12 independent $\text{Can}^{\text{R}}\text{Ade}^-$ clones isolated from strains containing direct *Alu* repeats. CHEF gel analysis coupled with Southern blot hybridization of chromosome V (Figure 2.2D), CGH microarray analysis (Figure

2.S3) and DNA combing (Figure 2.S3) identified four classes of rearrangements involving the left arm of chromosome V. In contrast to the GCR isolates from strains with inverted *Alus*, most (8 of 12) of the isolates obtained from strains with directly repeated *Alus* had terminal deletions of the left arm of chromosome V (including the *CANI* region) with no associated duplication (top panel, Figure 2.S3A). In such isolates (D-1, D-2, D-3, D-5, D-8, D-9, D-10, D-12), it is likely that the deleted chromosome is capped by a telomere addition, although other possibilities are not excluded. One of the isolates (D-7) has a 20 kb interstitial deletion (second panel from the top, Figure 2.S3A). The remaining three isolates had chromosome rearrangements that resemble Class II and Class III isolates derived from strains with the inverted *Alu* repeats. Two of the isolates (D-6 and D-11) had a duplication of about 30 kb adjacent to a terminal deletion of about 40 kb (third panel from top, Figure 2.S3A). We also found a single isolate (D-4) that had a terminal deletion, an adjacent duplication of 30 kb, and a duplication of about 130 kb from the right arm of chromosome V near *YERCTyI-1*. By molecular combing, we showed that the 30 kb duplications in D-4, D-6, and D-11 were inverted repeats (D-6 shown in Figure 2.S3B).

How are inverted repeats generated from yeast strains in which the *Alu* sequences are directly repeated? It is possible that a spontaneous DSB centromere-proximal to *CANI* triggers the resection of the broken end. The processed single-strand end can fold back on itself (utilizing very short inverted repeats), priming DNA synthesis to form a dicentric in an intramolecular reaction (Rattray et al., 2005). It should be stressed that strains with large inverted duplications are about 10^5 -fold more frequent in chromosome rearrangements derived from strains with the inverted *Alu* repeats than in strains with the direct *Alu* repeats. Supplementary Table 2 summarizes the

data on the Can^RAde⁻ isolates derived from strains with the direct and inverted *Alu* repeats.

2.3.4. Hairpin-capped breaks trigger a palindrome-dependent recurring chromosome instability

Since a hairpin-capped break located at the inverted *Alu* repeats often results in the formation of chromosome V rearrangements with larger (30 kb or 100 kb) quasi-palindromes, we hypothesized that these large quasi-palindromes could also cause hairpin-capped DSBs that would trigger a new round of chromosome V rearrangements. We observed previously that one phenotypic manifestation of hairpin-capped DSBs is the occurrence of small colonies containing a high percentage of large-budded cells. Thus, we used these criteria to follow the fate of the individual palindromes by isolating on non-selective medium small, non-*petite*, colonies in the progeny of Can^RAde⁻ clones. Depending on the GCR isolate, the frequency of the occurrence of the small colonies varied from 1 to 10%. These small colonies were streaked on non-selective medium and normal size colonies derived from individual small colonies were purified. The structure of chromosome V in these isolates was assessed by CHEF gel electrophoresis, followed by Southern blot hybridization with the *MET6* probe, and by microarray analysis (Figure 2.4).

Using this approach, we found that chromosome V containing a 30 kb quasi-palindrome with added telomeric sequences at the end can give rise to differently sized chromosomes that have either changed the length of the quasi-palindromic sequence (30 kb or 100kb) and/or acquired telomeres via BIR. The derivative chromosome V containing a 100kb quasi-palindrome was also unstable producing a next generation of chromosomes with changes in size and structure (Figure 2.4).

These results show that hairpin-capped breaks can trigger ongoing chromosome instability.

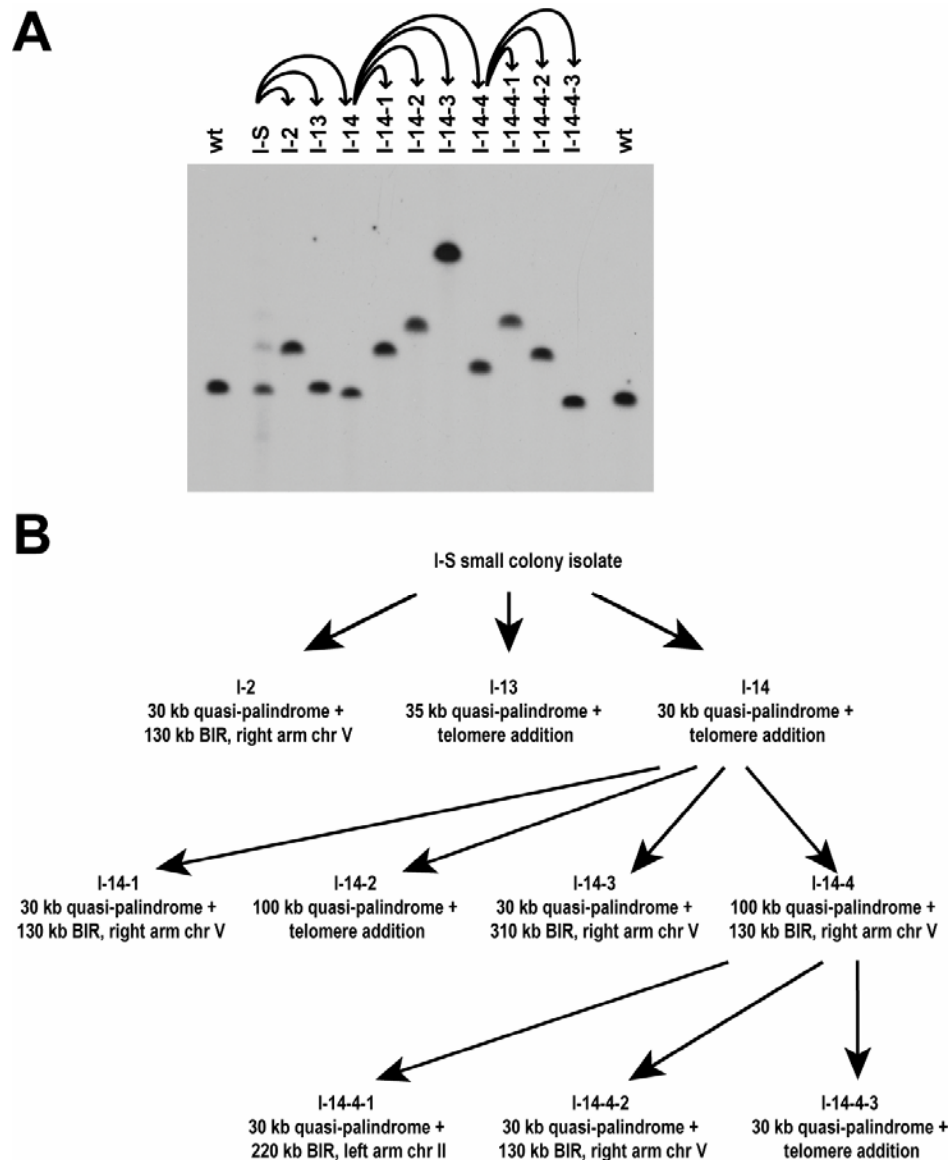


Figure 2.4. Recurring instability of chromosomes with large palindromes derived from the $\text{Can}^{\text{R}}\text{Ade}^-$ strain. (A) Analysis of chromosome V rearrangements by CHEF and Southern blot hybridization. The DNA samples are derived from the following strains: lane I-S (small colony of a primary $\text{Can}^{\text{R}}\text{Ade}^-$ isolate), lanes I-2, I-13 and I-14 (normal size colonies derived from I-S), lanes I-14-1 to I-14-3 (colonies derived from I-14), lanes I-14-4-1 to I-14-4-3 (colonies derived from I-14-4). Lanes labeled “wt” contain samples from the progenitor TP strains. Arrows above the lanes indicate the origin of the specific isolate. (B) Chromosome rearrangements in strains derived from I-S. The sizes of the palindromes and the presumptive nature of stabilizing the broken end are indicated. The sizes of the palindromes and the extent of the BIR-related duplications are based on microarray analysis.

2.3.5. Extrachromosomal amplicons resulting from hairpin-capped breaks are linear inverted dimers

A hairpin-capped break at the location of inverted *Alus* is expected to split the chromosome V into acentric and centromere-containing fragments (Figure 2.3). In the TP strains, the *CUPI* and *SFAI* gene dosage markers were inserted on the left arm of chromosome V, telomere-proximal from the break site (Figure 2.1). Using these strains, we could isolate strains that amplified *CUPI* and *SFAI* by selecting for derivatives that had increased levels of resistance to copper and formaldehyde. We found that both homologous and homeologous inverted *Alus* greatly induced amplification of *CUPI* and *SFAI* (Table 2.S1). $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ colonies occurred rarely in the progeny of strains containing direct repeats (2×10^{-9}). There were 11000-, 2000- and 250-fold increases in the rates of amplification in strains with 100, 94 and 86% identical inverted repeats, respectively.

We performed structural analysis of the rearrangements in $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates from strains with inverted and direct repeats using Southern blot hybridization of chromosome V following CHEF gel electrophoresis, and microarray analysis (Figure 2.5). All analyzed $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ clones from strains with inverted repeats (IA-1 to IA-12) had extrachromosomal amplicons that were approximately twice as large as the 42 kb fragment that would be expected as a result of breakage at the location of the inverted *Alus*. The size of these amplicons was also about twice as large as the amplified region detected with CGH analysis (Figure 2.5B). No size changes were detected for all 16 chromosomes (including chromosome V). Depending on the isolate, the copy number of the *CUPI* and *SFAI* was 5-13 times stronger than for the single copy as determined by Southern analysis or by microarray analysis using the Cluster Along Chromosomes (CLAC) program (Wang et al., 2005). Extrachromosomal

amplification was sometimes accompanied by non-disjunction of chromosome V and/or II (data not shown). Using molecular combing (Figure 2.5C) and Southern analysis (Figure 2.S1), we showed that the extrachromosomal amplicons were inverted dimers containing the *Alu* quasi-palindrome at the center of symmetry. This structure suggests that the extrachromosomal amplicons are generated by duplication of the unprocessed acentric hairpin-capped fragment that results from resolution of the cruciform (Figure 2.3).

The structural organization of amplicons in strains with direct *Alu* repeats (presumably reflecting spontaneous DSBs) was different. In these strains (DA-1 to DA-12), the amplicons were variable in size (Figure 2.5A). Although we have not analyzed these events in detail, using molecular combing, we found that only three of the twelve amplicons had an inverted repeat structure (data not shown). In addition, about 10% of these Cu^RFh^R isolates were disomic for chromosome V. Hence, there was an approximately 46,000-fold increase in the generation of extrachromosomal palindromic amplicons in strains with inverted *Alus* in comparison with strains containing direct repeats.

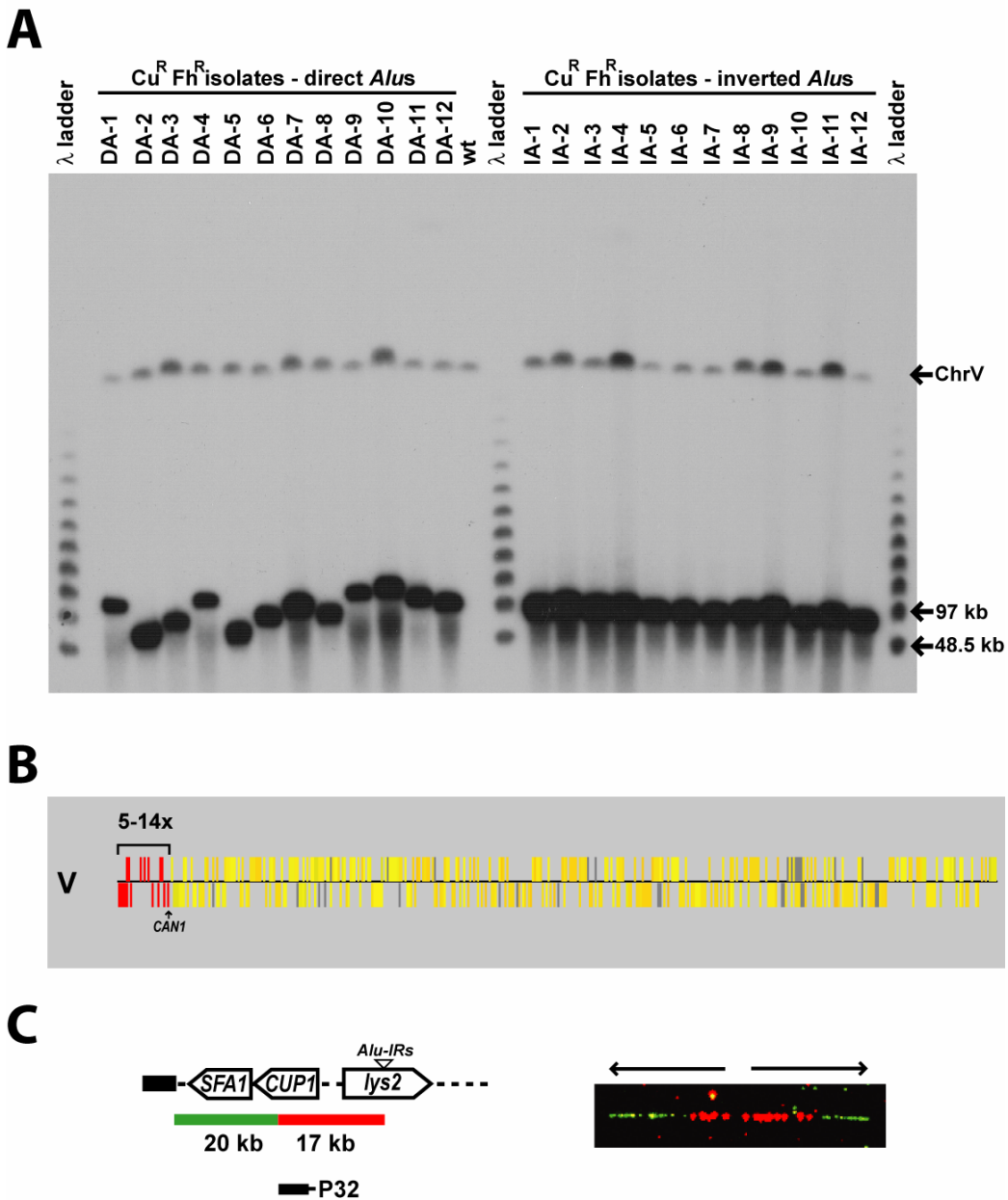


Figure 2.5. Analysis of extrachromosomal amplification events induced by hairpin-capped breaks.

(A) Analysis of karyotypic changes in Cu^RFh^R isolates derived from TP strains by CHEF and Southern blot hybridization. Lanes DA-1 to DA-12 have DNA samples from Cu^RFh^R isolates generated from TP strains with direct *Alus*, and lanes IA-1 to IA-12 contain DNA samples from Cu^RFh^R isolates of TP strains with inverted *Alus*. The lane labeled “wt” contains DNA from the progenitor TP strain. A 48.5 kb lambda ladder was used as a molecular size standard. The CHEF gel was hybridized simultaneously with *CUP1*- and lambda-specific probes.

(B) Microarray analysis of an extrachromosomal amplification in an Cu^RFh^R isolate of a TP strain with inverted *Alus*. The color coding is the same as in Figure 2E. The amplified region is bracketed, with the degree of amplification in different isolates varying between 5 and 14.

(C) The extrachromosomal amplicons are arranged as inverted dimers. The left panel shows the positions of the fluorescent and ³²P labeled probes used in FISH and Southern hybridization, respectively. The right panel is an example of an amplicon visualized by molecular combing.

2.3.6. Intrachromosomal amplification is an alternative outcome of the repair of hairpin-capped broken molecules

To determine if the centromere-containing broken molecules formed as a result of hairpin-capped DSB had the potential for gene amplification, we created TD strains, where the *SFAI* and *CUPI* were placed centromere-proximal to the *Alu* quasi-palindrome (Figure 2.1). In the TD strain with direct *Alu* repeats, the Cu^RFh^R isolates were uncommon and were usually a consequence of duplication of chromosome V or extrachromosomal amplification of broken fragments from the left arm of chromosome V (data not shown). In contrast, in Cu^RFh^R isolates from strains with inverted *Alus*, CHEF gel analysis showed that chromosome V was larger than the wild-type chromosome, and no extrachromosomal bands were detected (Figure 2.6A), indicating intrachromosomal amplification. The chromosomes containing the intrachromosomal amplicons were highly unstable and upon propagation often gave rise to the secondary rearrangements (data not shown).

There were three classes of the intrachromosomal amplicons. 96% of the Cu^RFh^R isolates were also Ade⁻ auxotrophs and Can^R, suggesting that gene amplification in these isolates was accompanied by a telomere-proximal deletion. This conclusion was confirmed by microarray analysis (Figure 2.6B) of 11 isolates. In Class I isolates (9 of 11), the telomere-proximal deletion bordered the amplified region on the left arm of chromosome V, and the amplified region corresponded to a 100 kb block with a breakpoint near the *YELCdelta4*, *YELWdelta5*, and *YELWdelta6* elements. Within the 100 kb block, there were two levels of amplification with the deletion-proximal 30 kb region (bordering *YELCdelta1* and *YELCdelta2*) being amplified more than the adjacent 70 kb region (Figure 2.6B). Among all of the Class I isolates, the copy number of the 30 kb region varied between 3 and 6 copies, while

the 70 kb region was amplified from 2 to 4 times. The presence of delta elements at the borders of the amplicon and the differentially-amplified regions within the amplicon strongly suggests the involvement of homologous recombination in the generation of these events.

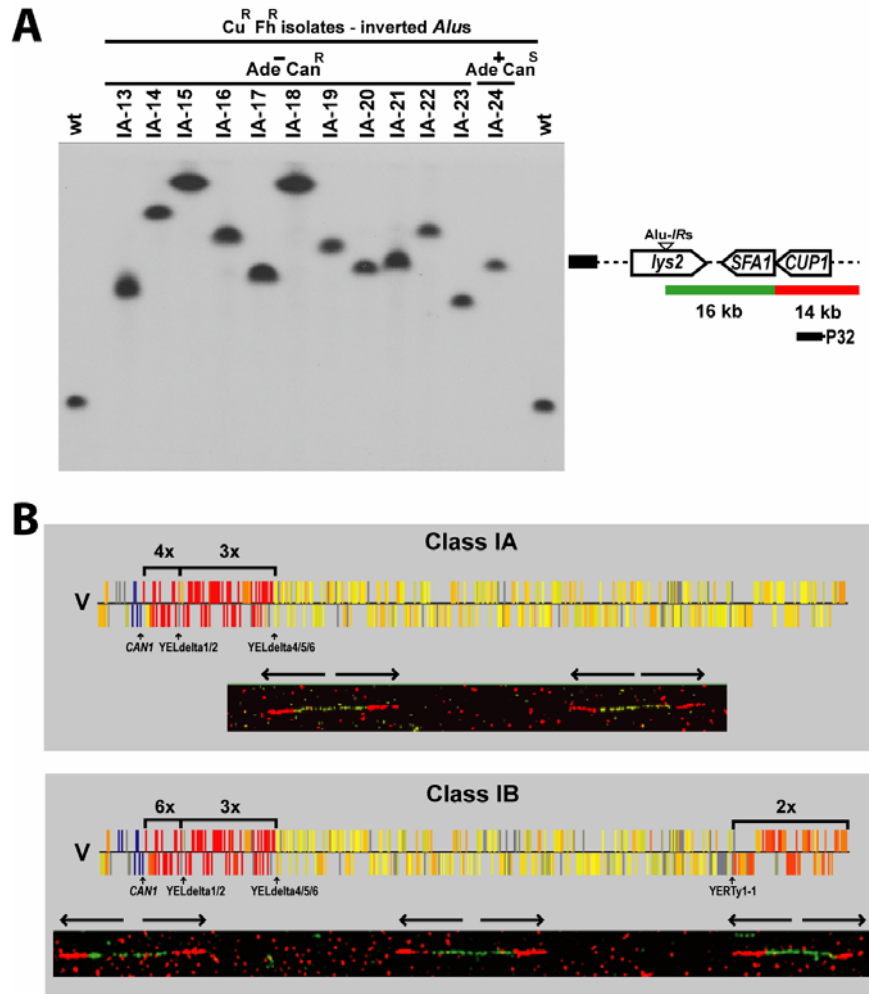


Figure 2.6. Analysis of intrachromosomal amplification stimulated by hairpin-capped breaks. (A) The left panel is a Southern analysis of a CHEF gel containing DNA from Cu^RFh^R isolates of TD strains with inverted *Alus* (lanes IA-13 to IA-24). The gel was hybridized with a *CUP1*-specific probe. The right panel shows the positions of the fluorescent and ³²-P labeled hybridization probes.

(B) In these panels, we show microarray and the FISH/molecular combing analyses for two Class I isolates. The upper isolate is Class IA (IA-17) and the bottom isolate is Class IB (IA-16). In the microarray depiction, bracketed regions indicate differential amplification of the two segments within that isolate. The positions of the *CAN1* gene and repetitive elements near the junctions of the amplification segments are indicated. The mechanism responsible for generation of intrachromosomal amplicons is discussed in the text and depicted in Figure 2.7.

There were two different types of Class I isolates. In Class IA strains (IA-17 and IA-23), there were no additional amplifications detected by microarrays (Figure 2.6B, upper panel). In Class IB strains (IA-13 to IA-16, IA-19 to IA-21), sequences from the right arm of chromosome V were amplified. The breakpoint of this amplification was near *YERTyI-1* (Figure 2.6B, lower panel). It is likely that Classes IA and IB represent two different mechanisms of stabilizing the end of a broken DNA molecule with Class IA reflecting *de novo* telomere addition and Class IB reflecting acquisition of a telomere by a BIR event involving the right arm of chromosome V. Although for most of the Class I isolates, the sizes of the chromosome Vs were as expected if the amplifications were intrachromosomal, for two isolates (IA-15 and IA-17), the chromosome Vs were larger than expected based on the regions amplified; we have not determined the source of the extra DNA in these isolates.

In the Class II isolates (IA-18 and IA-22), a 44 kb region centromere-proximal to the terminal deletion was amplified 3- to 4-fold. The boundary of the amplification events in these clones was determined by microarrays to be near *GDAI*, a region without any repetitive elements. In both Class II isolates, chromosome V was larger than expected by the microarray analysis. A small fraction (4%) of the $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates were $\text{Ade}^+\text{Can}^{\text{S}}$ (Class III). We examined only one of these isolates (IA-24) by microarrays and we found that, similar to the Class I isolates, there was an amplification of a 100 kb region spanning from the *CANI* locus up to the *YELCdelta4*, *YELWdelta5*, and *YELWdelta6* cluster. In this isolate, however, there was no deletion of the telomere-proximal region. Southern blot hybridization with an *ADE2* probe demonstrated that the telomere-proximal region in this isolate is a part of the chromosome V (data not shown).

To determine whether the intrachromosomal amplification events in the Cu^RFh^R isolates derived from the TD strains resulted in direct or inverted repeats, we examined the genomic DNA of six different Cu^RFh^R isolates (IA-13, IA-16, IA-17, IA-20, IA-22, IA-24) by molecular dynamic combing and dual-color FISH (Figure 2.6). In all six of the analyzed Cu^RFh^R clones, the amplified copies were organized as inverted repeats. In most of the isolates examined, the inverted repeats labeled by the fluorescent probes were separated by about 70 kb of unlabeled DNA. Southern blot hybridization and restriction analysis showed that *Alu*-quasi-palindromes are present in the center of the amplified units (Figure 2.S1). The observed structure of the intrachromosomal amplicons derived from the strains with the *Alu* inverted repeats is strikingly similar to that observed for HSRs in chromosomes of human cancers (Debatisse, 2005).

We suggest that the first steps in the intrachromosomal amplification process for Class I isolates are similar to those shown in Figure 2.3. The inverted *Alu* sequences extrude as a cruciform that is processed to yield two hairpin-capped molecules. The centromere-containing fragment is replicated, and the resulting dicentric chromosome breaks. In Class I events, we hypothesize that the DSB occurs near the delta 4 element on the right arm in order to produce a 100 kb duplication (Figure 2.7). The delta 4 element at the end of the chromosome invades one of the pairs of delta 1, 2 elements, setting up a “rolling” circle replication intermediate. This intermediate will produce tandem arrays of 130 kb repeats (two copies of the 30 kb repeat and one copy of the 70 kb region separating the 30 kb repeats). We postulate that this reaction will be terminated by a DSB break within the circular part of the replication structure or by a break at the replication fork. Once the DSB occurs, the broken end must be stabilized. We suggest that the break is stabilized either by telomere addition (Class

IA) or by a BIR event utilizing the delta elements of *YERTy1-1* as a template (Class IB). The Class II and III isolates are not explained by the model shown in Figure 2.7 and will be the subjects of future experiments. Supplementary Table 3 summarizes the data on the isolates with intrachromosomal gene amplifications.

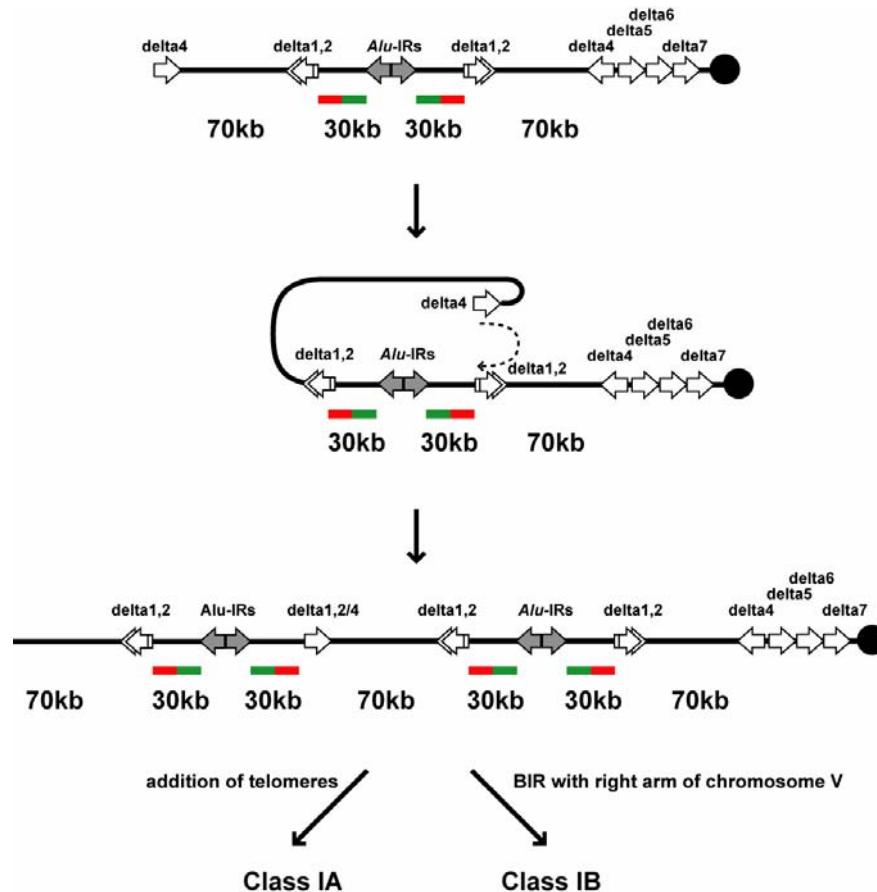


Figure 2.7. Model for generating intrachromosomal amplicons. Arrows indicate the orientation and location of the delta elements on left arm of chromosome V. *Alu* repeats located at the center of symmetry of the duplications are shown as solid gray arrows, and the centromere is depicted as a filled circle. We suggest that a derivative of chromosome V with a palindrome centered on the *Alu* repeats is generated by the pathway shown in Figure 2.3. In this particular derivative, the DSB resulting from resolution of the dicentric chromosome is near delta 4. The red and green regions indicate the positions of the fluorescent probes described in Figure 2.6. As indicated by the dashed arrow, delta 4 invades one of the two delta 1 elements, setting up a BIR. In the orientation shown, a rolling circle replication intermediate would result. Continued synthesis would produce tandem arrays of 60 kb (containing the inverted pairs of labeled segments) separated by 70 kb. DNA synthesis would continue to produce longer arrays until the rolling circle intermediate is broken. The broken end could then be healed by telomeric addition (Class IA events) or by delta-mediated BIR event using the right arm of chromosome V as a template (Class IB events). In this figure, the 30 kb segment is amplified four times, while the 70 kb segment is amplified three times (similar to the isolate shown in the upper half of Figure 2.6B)

2.3.7. Elevated levels of deletions and amplifications in *mre11* strains

In our previous study (Lobachev et al., 2002), we showed that Mre11p was not required to make the DSB at the extruded cruciform, but was required to process the resulting hairpin. In the absence of Mre11p, we observed an increased level of chromosomes with inverted duplications. Consistent with these observations, in the present study, we found that the *mre11* mutation elevated the rate of *CANI* deletions and amplifications in strains with inverted *Alu* sequences (100% identity) about five-fold (Table 2.S1). The observation that all three types of rearrangements were affected to similar extents is consistent with the hypothesis that all are initiated by the same event: a hairpin that is generated by resolution of an extruded cruciform. Failure to process the hairpin-capped molecules by Mre11p results in an increased level of acentric and dicentric intermediates that, in turn, increases the level of downstream chromosome rearrangements.

2.4. Discussion

In previous yeast studies, several different modes of amplification have been detected including chromosome aneuploidy (Whittaker et al., 1988), non-reciprocal translocations (Dunham et al., 2002), tandem duplications (Koszul et al., 2004), extrachromosomal palindromes (Dorsey et al., 1992), and intrachromosomal palindromes (Moore et al., 2000). In our experiments, we selected deletions and duplications of markers located close to an inverted pair of *Alu* elements. In three separate types of experiments, we selected for loss of a marker distal to the palindrome, duplication of markers centromere-distal to the palindrome, or duplication of markers centromere-proximal to the palindrome. Depending on the nature of the selection different chromosome rearrangements were recovered. Our

study demonstrates that palindromic sequences that can adopt hairpin and cruciform secondary structures are potent sources of GCR events, including gene amplification.

2.4.1. Mechanisms of GCRs triggered by hairpin-capped DSBs

Previously we showed that inverted repeats of *Alu* inserted into chromosome II led to DSBs (Lobachev et al., 2002) and the resulting hairpin-capped ends required Mre11p and Sae2p for their processing. Unprocessed hairpin-capped molecules accumulated in mutants defective in the endonuclease function of the Mre11 complex and frequently gave rise to large acentric and dicentric inverted duplications. In this study, we examined the types of GCR events associated with processing of the inverted duplications. All of the chromosome rearrangements can be explained by a single initiating event: processing of an extruded cruciform to generate two hairpin-capped broken ends (Figure 2.3). If neither end is processed by the MRX complex, two palindromic dimers will be produced: i) an acentric fragment that includes *CANI* and *ADE2* and ii) a dicentric fragment with a deletion from *LYS2* to the left telomere region of chromosome V (Figure 2.3). In experiments in which we select for loss of *CANI*, only one of these products will be detected. As shown in Figure 2.3, the dicentric fragment with the terminal deletion would be expected to break, resulting in a duplication of the sequences adjacent to the deletion (Haber et al., 1984; Kramer et al., 1994). To generate a stable chromosome, the broken end would have to acquire a telomere. This process could involve either *de novo* telomere addition by telomerase, as observed in previous studies (Kramer and Haber, 1993; Pennaneach et al., 2006), or repair by homologous recombination utilizing microhomology or a large repetitive sequence (for example, a delta or Ty element) to initiate a BIR event. Our analysis

suggests that both *de novo* telomere additions and BIR events are used to repair the broken ends.

In all of the Cu^RFh^R clones isolated from TP strains (*CUPI* and *SFAI* located centromere-distal to the inverted repeats), we found that the *CUPI* and *SFAI* repeats were part of a large (about 80 kb) palindromic extrachromosomal molecules (Figure 2.5), as expected from the mechanism shown in Figure 2.3. In contrast, the majority of the Cu^RFh^R clones isolated from the TD strains (*CUPI* and *SFAI* located centromere-proximal to the inverted repeats), the amplified copies of the selected markers were located on the centromere-containing portion of chromosome V. The resulting rearranged chromosome had a deletion of the DNA distal to the inverted repeats and tandem arrays of an inverted repeat containing the selectable markers (Figure 2.6). We suggest that dicentric molecule is a common intermediate for this rearrangement and for the chromosome deletions. A broken end resulting from the breakage of the dicentric could invade a repeat on the same chromosome arm, generating a rolling circle intermediate that would produce a tandem array of repeats (Figure 2.7).

2.4.2. Palindrome regeneration cycle leading to continuing genetic instability

We showed that the inverted *Alu* repeats induced arm loss events that were coupled with the formation of large (more than 30 kb) quasi-palindromes (Figure 2.2). It is important to point out that such chromosomes will be unstable because of the existence of the large quasi-palindromes. It should be noted that the breakpoints of the resulting rearrangements do not co-localize with the initial hairpin-capped break site, but are at the position where the repair of the broken molecules was initiated.

The sequence that triggered the primary DSB and resulting rearrangement is positioned at the center of the duplication.

Other studies showed that long palindromic duplications can be generated by mechanisms that involve very short (4-12 bp) inverted repeats (Albrecht et al., 2000; Maringele and Lydall, 2004; Rattray et al., 2005). It is possible that the rare inverted duplications identified in the strains with direct *Alu* repeats occur via one of these mechanisms. Regardless of the mechanism by which they are formed, large palindromic regions will initiate iterative cycles of genome instability leading to a wide variety of chromosomal aberrations, as shown in Figure 2.4.

2.4.3. Double minutes versus homogeneously-staining regions

We have presented evidence that hairpin-capped breaks can lead to either extra- or intrachromosomal amplification. The nature of the amplicons depends on the chromosomal location of the amplified gene relative to the site of DSB. The chromosomal structure with the order, telomere/amplified gene/hairpin-capped DSB/centromere, results in a double-minute-like amplicon. In contrast, the telomere/hairpin-capped DSB/amplified gene/centromere arrangement yields intrachromosomal amplicons with an inverted ladder-like structure. The structural organization of the amplicons identified in this work bear striking similarity to DMs and HSRs amplification products detected in human cancers. Based on this correlation, we propose that the rules of the palindrome-dependent amplification as seen in yeast may also operate in higher eukaryotes.

We found that extrachromosomal amplicons (up to 14 in copy number) resulting from hairpin capped DSBs were linear dimers. The most likely mechanism of the accumulation of the extrachromosomal copies to this level is missegregation of

the acentric fragments during mitotic divisions. It is interesting to note that the formation of DMs in our system was not accompanied by the arm loss event. This observation indicates that the hairpin-apped DSB was formed during G2 or S phase, after synthesis of the *Alu*-IRs region. We also found that extrachromosomal amplification was sometimes associated with non-disjunction of chromosome V and/or II. Similarly, trisomy is frequently found in cancer cells carrying DMs (Naeem, 2005).

The generation of intrachromosomal amplicons resulting from the hairpin capped DSB was usually coupled with the loss of the telomere proximal region. A similar pattern has been described for HSRs in tumors. This structure is most frequently explained by the BFB cycle (Debatisse, 2005). The key step in BFB model is the repetitive formation of dicentrics through sister-chromatid fusions that are presumably carried out by non-homologous end joining repair (NHEJ). This conclusion is somewhat controversial, since a mammalian cell line deficient in NHEJ had an elevated level of gene amplification (Mondello et al., 2001). In our experiments, it is unlikely that dicentrics are caused by fusions of broken sister-chromatids, since NHEJ is inefficient in yeast cells compared to mammalian cells (Krogh and Symington, 2004). In addition, the palindrome-mediated DSB obviates the requirement for sister-chromatid fusions by generating a broken molecule terminated with a hairpin.

2.4.4. Implications for human genome stability

We find that both homologous and homeologous inverted *Alu* repeats are strong inducers of GCR (Table 2.S1); the latter class of *Alu* repeats are found in the human genome (Lobachev et al., 2000; Stenger et al., 2001). These repeats, therefore,

represent a potential threat to the integrity of the human genome, especially in mutant backgrounds that promote rearrangements (for example, in mutants with an defective MRX complex). The mechanisms of GCRs described in this study could apply, not only to inverted repeats, but also to other repeats that can adopt stable hairpin or cruciform structures, such as certain trinucleotides (CAG/CTG or CCG/CGG repeats) or AT- and GC -rich minisatellites. These types of repeats are often found at rare fragile sites in humans (Sutherland, 2003).

Terminal deletions, duplications, translocations, amplifications and more complex rearrangements are frequently found in leukemias, lymphomas and sarcomas (Albertson et al., 2003; Fletcher, 2005; Naeem, 2005). Our results demonstrate that, in palindrome-mediated rearrangements, the sequence that triggers GCR is located in the center of the duplicated or amplified regions. We propose that the specific patterns of GCR described in our study (terminal deletions coupled with adjacent duplications) can serve as biomarkers in cancer genomic studies to reveal the causative sequence of rearrangements.

2.5. Experimental Procedures

2.5.1. Strains and Genetic Techniques

All strains in this study were isogenic to KS520 (*MATa*, *his7-2*, *leu2-3,112*, *trp1-Δ*, *ura3-Δ*, *lys2-Δ*, *ade2-Δ*, *bar1-Δ*, *sfa1-Δ*, *cup1-1-Δ*, *yhr054c-Δ*, *cup1-2-Δ*). Details of the constructions of the TP and TD strains, as well as description of genetic techniques, are given in the Supplemental Experimental Procedures.

2.5.2. Structural Analysis of the Genome Rearrangements

Chromosome aberrations were characterized using CHEF (contour-clamped homogeneous electric field) gels, Southern Blot Hybridization, CGH (comparative genomic hybridization) analysis, DNA combing and FISH (fluorescent *in situ* hybridization). The detailed description of these techniques can be found in the Supplemental Experimental Procedures.

2.6. Acknowledgements

We thank C. Moon and J. Tucker for assistance in experiments; A. Bensimon and members of his lab for help with molecular combing technique; L. Williams for help with silanization of the coverslips; D. Aiken, S. Woodard and J. Crowe for help with recording images using Zeiss microscopes; D. Gordenin, N. Degtyareva, N. McCarty, S. Spiro and S. Lewis for critical reading of the manuscript and helpful discussions. This work was supported by NSF grant MCB-0417088 (to K.S.L.) and NIH grant GM52319 (to T.D.P).

2.7. Supplemental Data

2.7.1. Strains

All strains in this study were isogenic to KS520 (*MATa*, *his7-2*, *leu2-3,112*, *trp1-Δ*, *ura3-Δ*, *lys2-Δ*, *ade2-Δ*, *bar1-Δ*, *sfa1-Δ*, *cup1-1-Δ*, *yhr054c-Δ*, *cup1-2-Δ*). 4477 bp *LYS2*, 2227 bp *ADE2*, 787 bp *CUP1* and 1619 bp *SFA1* fragments were PCR amplified from genomic DNA of the CGL strain (Lobachev et al., 2000) and inserted consecutively into chromosome V of the KS520 strain. *LYS2* and *ADE2* were inserted adjacent to each other between 34211 bp and 34212 bp (all coordinates here and below are given in accordance with *Saccharomyces* Genome Database). In the TP strains, *CUP1* and *SFA1* were placed between 29616 bp and 29617 bp and between

31676 bp and 31677 bp, respectively. In the TD strains, *CUP1* and *SFA1* were inserted next to each other between 36396 bp and 36397 bp. *SFA1* and *CUP1* fragments were introduced into chromosomal DNA using the *delitto perfetto* technique (Storici et al., 2001). Inverted and direct *Alus* were inserted into *LYS2* gene as previously described (Lobachev et al., 2000). *MRE11* gene was disrupted with the *kanMX* cassette (Wach et al., 1994). Nucleotide sequences of the primers used for integrations and disruptions are available upon request.

2.7.2. Genetic techniques

The rates and 95% confidence intervals of the arm loss and gene amplification were estimated in fluctuation tests using at least 14 independent cultures (Lobachev et al., 1998). The canavanine-containing media was made with a low concentration of adenine (5mg/L) to allow color detection; strains with an *ade2* mutations form red colonies in medium with low levels of adenine. Copper plates were prepared from SD complete media with final concentration of 700 μ M CuSO₄ solution.

Formaldehyde plates were made from SD complete media with a final concentration of 2mM formaldehyde solution (Sigma). To select for amplification events, we replica plated the Cu^R colonies to freshly-made formaldehyde plates. After two days of incubation, these plates were replica plated again to formaldehyde-containing medium to verify the growth. For CGH analysis genomic DNA was extracted from Can^RAde⁻ isolates and wild type strains grown in liquid YPD media, while the Cu^RFh^R isolates were propagated on copper plates.

2.7.3. CHEF gel electrophoreis and Southern blot hybridization

Chromosomal DNA was embedded into agarose plugs using the CHEF Genomic DNA plug Kit from Bio-Rad. Gels were run in 0.5X TBE at 14°C using the Bio-Rad CHEF Mapper XA for 31 hours with switch times of 15.09s-1m5.17s for the analysis of extrachromosomal amplification and for 40 hours with switch times of 36.63s-2m6.67s for the analysis of arm loss events and intrachromosomal amplification events. ³²P-labeled 339 bp, 182 bp, 374 bp, 430 bp and 259 bp probes homologous to the *MET6*, *CUP1*, *CDC25*, *SPC97* genes and HS-*Alu* sequence, respectively, were used in Southern blot hybridization. Southern blot hybridization was performed as previously described (Lobachev et al., 2002). Nucleotide sequences of the primers used to generate the fragments for labeling are available upon request

2.7.4. Comparative Genomic Hybridization analysis

DNA preparation and subsequent microarray analysis were performed according to procedures described (Lemoine et al., 2005). Arrays were analyzed using GenePix pro 4.1 (Axon Instruments) and Gene Spring[®] 5.1 (Silicon Genetics). Genomic ratios and copy numbers of the amplified regions were estimated using Cluster Along Chromosomes (CLAC) analysis (Wang et al., 2005). A complete analysis of the microarrays can be found on line at <https://genome.unc.edu/>

2.7.5. DNA Combing and Fluorescent In Situ Hybridization

Genomic DNA preparation and molecular combing were performed as described in (Conti, 2001). An automated combing apparatus for stretching DNA was designed and constructed in the mechanical workshop of the School of Physics, Georgia Institute of Technology. Dual-color hybridization of rearranged molecules was performed using two adjacent fluorescent probes. The probes were comprised of a set

of 6 to 9 kb long fragments that were obtained by PCR amplification from genomic DNA of TP or TD strains as template. Nucleotide sequences of the primers used to generate fragments for labeling are available upon request. Probes were labeled with either biotin-dUTP or dig-dUTP. Hybridization and fluorescent detection of combed DNA molecules were achieved according to protocols described in (Conti, 2001) with a few modifications. Five successive layers of fluorophore-conjugated antibodies diluted 1:100 in 1X PBST (1X PBS + 0.05% Tween) were used. For the biotin-conjugated probes, the following series were used: 1) Alexa-488-Streptavidin (Molecular Probes), 2) Biotinylated anti-streptavidin (Jackson ImmunoResearch Lab, PA), 3) Alexa-488-Streptavidin, 4) Biotinylated anti-streptavidin and 5) Alexa-488-Streptavidin. For DNA molecules labeled with digoxenin the following series were used: 1) Cy3-coupled mouse anti-digoxenin, 2) Cy3-coupled rat anti-mouse, 3) Cy3-coupled rabbit anti-rat, 4) Cy3-coupled mouse anti-rabbit and 5) Cy3-coupled donkey anti-mouse (all antibodies from Jackson ImmunoResearch Lab, PA) were employed for detection. All images were acquired using the Zeiss LSM 510 Confocal Microscope with 100X objective. The lengths of the fluorescent stretches were converted into base pairs using the conversion that one micron equaled 2 to 2.5 kb (based on the use of phage lambda DNA molecules as the size standard). For each of the scanned images, 20 to 40 full-length molecules were detected, and arrangements of the fluorescent stretches were documented.

2.7.6. Supplemental References

Conti, C., Caburet, S., Schurra, C., Bensimon, A. (2001). Molecular Combing, In Current Protocols in Cytometry (New York: John Wiley & Sons, Inc), pp. 8.10.11-18.10.23.

- Lemoine, F. J., Degtyareva, N. P., Lobachev, K., and Petes, T. D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* *120*, 587-598.
- Lobachev, K. S., Gordenin, D. A., and Resnick, M. A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* *108*, 183-193.
- Lobachev, K. S., Shor, B. M., Tran, H. T., Taylor, W., Keen, J. D., Resnick, M. A., and Gordenin, D. A. (1998). Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. *Genetics* *148*, 1507-1524.
- Lobachev, K. S., Stenger, J. E., Kozyreva, O. G., Jurka, J., Gordenin, D. A., and Resnick, M. A. (2000). Inverted Alu repeats unstable in yeast are excluded from the human genome. *Embo J* *19*, 3822-3830.
- Storici, F., Lewis, L. K., and Resnick, M. A. (2001). *In vivo* site-directed mutagenesis using oligonucleotides. *Nat Biotechnol* *19*, 773-776.
- Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast* *10*, 1793-1808.
- Wang, P., Kim, Y., Pollack, J., Narasimhan, B., and Tibshirani, R. (2005). A method for calling gains and losses in array CGH data. *Biostatistics* *6*, 45-58.

2.7.7. Supplementary Table 2.S1. Induction of GCR events by homologous and homeologous inverted *Alus* in wild type and *Δmre11* strains.

Insertion in <i>LYS2</i>	Homology %	GCR rate (x 10 ⁷)					
		Loss of <i>CAN1</i> region ^b		Gene Amplification			
		wild type	<i>Δmre11</i>	Extrachromosomal ^c		Intrachromosomal ^d	
		wild type	<i>Δmre11</i>	wild type	<i>Δmre11</i>	wild type	<i>Δmre11</i>
Inverted <i>Alus</i>	100	371 (294-421) ^a	1808 (1643-2138)	229 (185-316)	1031 (795-1250)	64 (59-93)	256 (124-423)
	94	27 (17-41)	495 (380-656)	40 (29-49)	303 (249-329)	3 (2-4)	44 (36-64)
	86	7 (6-8)	26 (23-31)	6 (5-9)	30 (25-33)	0.57 (0.35-0.87)	1 (0.45-1.4)
Direct <i>Alus</i>	100	0.015 (0.01-0.02)	6 (5-9)	0.02 (0.01-0.03)	2 (1-3)	ND	ND

^a Numbers in parentheses correspond to the 95% confidence interval.

^b The rates of the arm loss presented in the table were measured using TP strains. Similar rates were calculated for the TD strains (data not shown)

^c The rate of extrachromosomal amplification is approximate since 1/12 of Cu^RFh^R isolates are due to non-disjunction of chromosome V. The rates were estimated using TP strains.

^d The rates of intrachromosomal amplification were measured in TD strains. ND –not determined.

2.7.8. Supplementary Table 2.S2. Can^RAde⁻ isolates of strains with direct and inverted *Alu* repeats

Isolate	GCR class	Phenotype	Rearrangement (based on microarray)	Amplicon structure and total size	Amplicon breakpoint	FISH analysis	Est. size, kbs	Size on CHEF gel, kbs
IA-13	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	4x2x, 100kb	YELdelta4-6	Yes	932	~900
IA-14	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	5x4x, 100kb	YELdelta4-6		1116	~1100
IA-15	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta4-6		1062	~1400
IA-16	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta4-6	Yes	1062	~1100
IA-17	IA	Can ^R Ade ⁻	del.+amp. + telomere addition	4x3x, 100kb	YELdelta4-6	Yes	874	~950
IA-18	II	Can ^R Ade ⁻	del.+amp. + telomere addition	4x, 42kb	GDA1		712	~1400
IA-19	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta4-6		1062	~1100
IA-20	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	4x3x, 100kb	YELdelta4-6	Yes	1002	~1000
IA-21	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	5x3x, 100kb	YELdelta4-6		1032	1100
IA-22	II	Can ^R Ade ⁻	del.+amp. + telomere addition	4x, 42kb	GDA1	Yes	712	~1100
IA-23	IA	Can ^R Ade ⁻	del.+amp. + telomere addition	3x2x, 100kb	YELdelta4-6		774	~750
IA-24	III	Can ^S Ade ⁺	amp.	4x3x, 100kb	YELdelta4-6	Yes	916	~900

2.7.9. Supplementary Table 2.S3. Cu^RFh^R isolates containing intrachromosomal amplicons

Isolate	GCR class	Phenotype	Rearrangement (based on microarray)	Amplicon structure and total size	Amplicon breakpoint	FISH analysis	Est. size, kbs	Size on CHEF gel, kbs
IA-13	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	4x2x, 100kb	YELdelta 4-6	Yes	932	~900
IA-14	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	5x4x, 100kb	YELdelta 4-6		1116	~1100
IA-15	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta 4-6		1062	~1400
IA-16	IB	Can ^R Ade ⁻	del.+amp.+BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta 4-6	Yes	1062	~1100
IA-17	IA	Can ^R Ade ⁻	del.+amp. + telomere addition	4x3x, 100kb	YELdelta 4-6	Yes	874	~950
IA-18	II	Can ^R Ade ⁻	del.+amp. + telomere addition	4x, 42kb	GDA1		712	~1400
IA-19	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	6x3x, 100kb	YELdelta 4-6		1062	~1100
IA-20	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	4x3x, 100kb	YELdelta 4-6	Yes	1002	~1000
IA-21	IB	Can ^R Ade ⁻	del.+amp. + BIR (right arm, Chr.V)	5x3x, 100kb	YELdelta 4-6		1032	1100
IA-22	II	Can ^R Ade ⁻	del.+amp. + telomere addition	4x, 42kb	GDA1	Yes	712	~1100
IA-23	IA	Can ^R Ade ⁻	del.+amp. + telomere addition	3x2x, 100kb	YELdelta 4-6		774	~750
IA-24	III	Can ^S Ade ⁺	amp.	4x3x, 100kb	YELdelta 4-6	Yes	916	~900

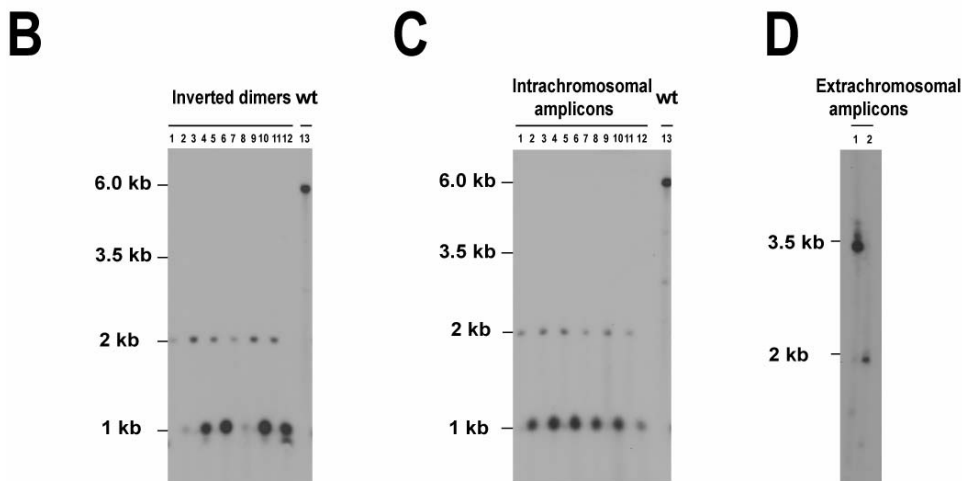
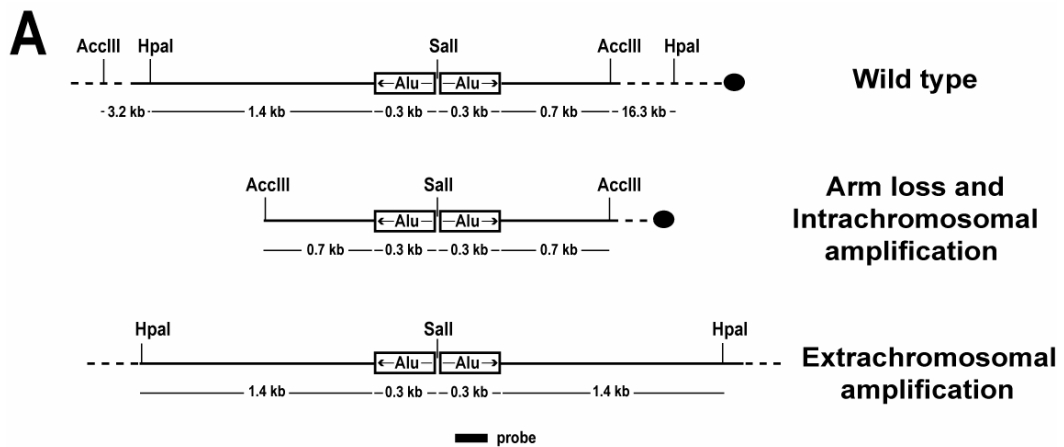


Figure 2.S1. *Alu* quasi-palindromes are located at the center of symmetry of the inverted duplications in strains with terminal deletions, and in strains with extrachromosomal and intrachromosomal amplicons.

(A) The restriction maps of the chromosomal regions containing *Alu* insertions (arrows inside the open boxes) in the wild type and rearranged chromosome Vs. The solid rectangle indicates the *Alu*-specific probe used in Southern blot hybridization.

(B) Southern analysis of inverted duplications in strains with terminal deletions. Genomic DNAs from six independent $\text{Can}^{\text{R}}\text{Ade}^{-}$ isolates (I-5, I-6, I-7, I-8, I-10 and I-12) were digested with *AccIII* (lanes with odd numbers) and *AccIII-SalI* (lanes with even numbers). Genomic DNA from wild type was digested with *AccIII* (lane 13). Digestions of I-5 to I-12 DNAs were loaded from left to right consecutively.

(C) Southern analysis of intrachromosomal amplicons. Genomic DNAs from the six independent $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates from TD strains (IA-13, IA-16, IA-17, IA-20, IA-22 and IA-24) were digested with *AccIII* (lanes with odd numbers) and *AccIII-SalI* (even numbers). Genomic DNA from wild type was digested with *AccIII* (lane 13). Digestions of IA-13 to IA-24 DNAs were loaded from left to right consecutively.

(D) Southern analysis of extrachromosomal amplicons. The genomic DNA of the $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates from TP strains was embedded into agarose plugs and separated on low-melt agarose gel. The gel was stained with ethidium bromide and the band corresponding to the extrachromosomal amplicon was extracted from the gel. The recovered DNA was digested with *HpaI* (lane 1) and *HpaI-SalI* (lane 2).

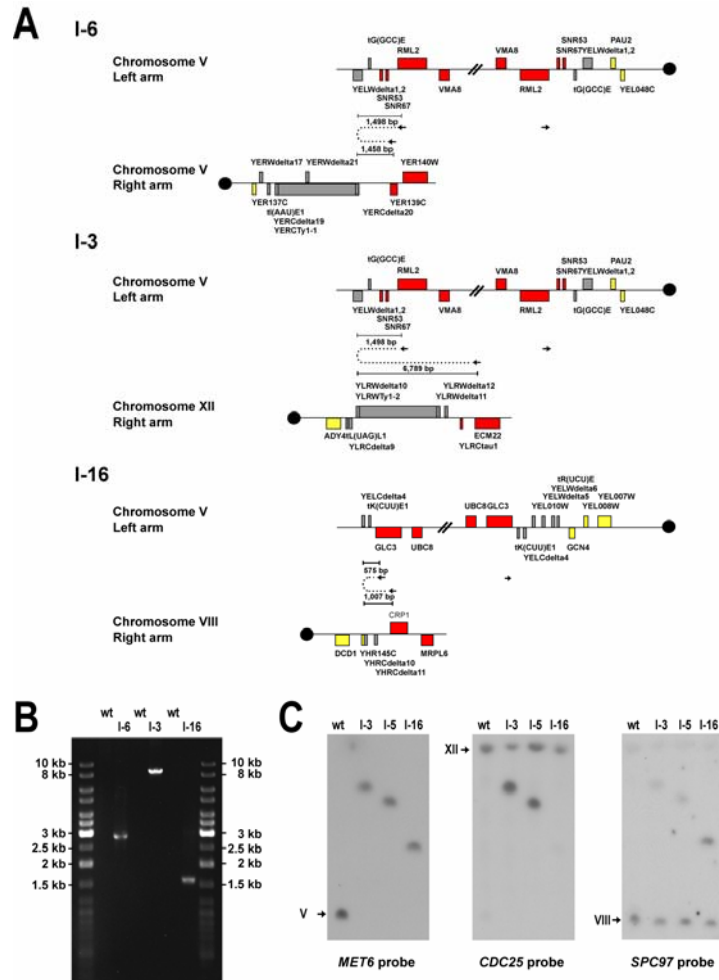


Figure 2.S2. Analysis of non-reciprocal translocations in *Can^RAde⁻* isolates. (A) Primer locations for PCR analysis of BIR-mediated translocations. Maps of the regions involved into the BIR events for the donor and recipient chromosomes based on microarray analysis are shown. Genes and repetitive elements are shown as solid boxes. Color coding is as follows: gray – repetitive elements, yellow – single copy sequences, and red – duplications. Centromeres are shown as solid circles. I-3, I-6 and I-12 are *Can^RAde⁻* isolates containing inverted duplication that (based on the microarray data) acquired telomeres by BIR with the right arms of chromosomes XII, V and VIII, respectively. Small arrows indicate the position of the primers used in PCR (panel B). The distances from the location of the primers to the deduced breakpoint are indicated. BIR events involving recombining sequences are depicted by dashed lines. (B) PCR analysis of translocations. PCR was performed with genomic DNA extracted from the wild type and I-3, I-6 and I-16 strains (ethidium bromide-stained gel). The sizes of the amplified bands correspond to the predicted sizes. (C) Southern analysis of translocations using chromosome-specific probes. Chromosomes from wt and translocation-containing isolates were separated on a CHEF gel. Based on microarray analysis, I-3 and I-5 both contain translocations involving chromosomes V and XII, and I-16 has a V-VIII translocation. The left panel shows hybridization to a *MET6*-specific probe (right arm of chromosome V; left panel). The membrane was stripped and hybridized to a *CDC25*-specific probe (right arm of chromosome XII; middle panel). The membrane was stripped again and hybridized to a *SPC97*-specific probe (right arm of chromosome VIII; right panel). Bands corresponding to the wild type chromosomes V, VIII and XII are indicated by arrows

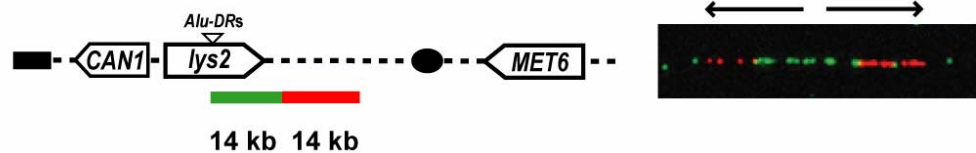
A**B**

Figure 2.S3. Structural analysis of arm loss events in TP strains containing direct *Alus*

(A) Microarray analysis of GCRs. The color coding is as in Figure 2. Each light grey panel represents a different type of chromosome rearrangement. The specific isolates, from top to bottom, are D-1, D-7, D-6, and D-4, respectively. The ratios and percentages of each class are indicated.

(B) The duplicated region adjacent to the terminal deletion is organized as an inverted repeat in isolate D-6; similar results were obtained with D-4 and D-11. The left panel shows the location of the regions for which fluorescent probes were made for FISH analysis. The right panel is an example of an inverted duplication as visualized by molecular combing. Arrowheads above the panel depict repeat units in the duplication.

2.8. References

- Albertson, D. G., Collins, C., McCormick, F., and Gray, J. W. (2003). Chromosome aberrations in solid tumors. *Nat. Genet.* *34*, 369-376.
- Albrecht, E. B., Hunyady, A. B., Stark, G. R., and Patterson, T. E. (2000). Mechanisms of *sod2* gene amplification in *Schizosaccharomyces pombe*. *Mol. Biol. Cell* *11*, 873-886.
- Chan, S. R., and Blackburn, E. H. (2004). Telomeres and telomerase. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *359*, 109-121.
- Chen, C., and Kolodner, R. D. (1999). Gross chromosomal rearrangements in *Saccharomyces cerevisiae* replication and recombination defective mutants. *Nat. Genet.* *23*, 81-85.
- Coquelle, A., Pipiras, E., Toledo, F., Buttin, G., and Debatisse, M. (1997). Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* *89*, 215-225.
- Coquelle, A., Rozier, L., Dutrillaux, B., and Debatisse, M. (2002). Induction of multiple double-strand breaks within an *hsr* by meganucleaseI-SceI expression or fragile site activation leads to formation of double minutes and other chromosomal rearrangements. *Oncogene* *21*, 7671-7679.
- Debatisse, M. a. M., B (2005). Gene amplification mechanisms, In *Genome Instability in Cancer Development*, E. A. Nigg, ed. (Netherlands: Springer), pp. 343-361.
- Dorsey, M., Peterson, C., Bray, K., and Paquin, C. E. (1992). Spontaneous amplification of the *ADH4* gene in *Saccharomyces cerevisiae*. *Genetics* *132*, 943-950.
- Dunham, M. J., Badrane, H., Ferea, T., Adams, J., Brown, P. O., Rosenzweig, F., and Botstein, D. (2002). Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* *99*, 16144-16149.
- Fletcher, J. A. (2005). Cytogenetics of Solid Tumors, In *The Principles of Clinical Cytogenetics*, S. L. a. K. Gersen, M.B., ed. (Totowa, New Jersey: Humana Press), pp. 421-451.
- Fried, M., Feo, S., and Heard, E. (1991). The role of inverted duplication in the generation of gene amplification in mammalian cells. *Biochim. Biophys. Acta* *1090*, 143-155.
- Haber, J. E., Thorburn, P. C., and Rogers, D. (1984). Meiotic and mitotic behavior of dicentric chromosomes in *Saccharomyces cerevisiae*. *Genetics* *106*, 185-205.

- Kobayashi, T., Horiuchi, T., Tongaonkar, P., Vu, L., and Nomura, M. (2004). SIR2 regulates recombination between different rDNA repeats, but not recombination within individual rRNA genes in yeast. *Cell* *117*, 441-453.
- Koszul, R., Caburet, S., Dujon, B., and Fischer, G. (2004). Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* *23*, 234-243.
- Kramer, K. M., Brock, J. A., Bloom, K., Moore, J. K., and Haber, J. E. (1994). Two different types of double-strand breaks in *Saccharomyces cerevisiae* are repaired by similar RAD52-independent, nonhomologous recombination events. *Mol. Cell. Biol.* *14*, 1293-1301.
- Kramer, K. M., and Haber, J. E. (1993). New telomeres in yeast are initiated with a highly selected subset of TG1-3 repeats. *Genes Dev.* *7*, 2345-2356.
- Kraus, E., Leung, W. Y., and Haber, J. E. (2001). Break-induced replication: a review and an example in budding yeast. *Proc Natl. Acad. Sci. USA* *98*, 8255-8262.
- Krogh, B. O., and Symington, L. S. (2004). Recombination proteins in yeast. *Annu. Rev. Genet.* *38*, 233-271.
- Kuo, M. T., Vyas, R. C., Jiang, L. X., and Hittelman, W. N. (1994). Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. *Mol. Cell. Biol.* *14*, 5202-5211.
- Lemoine, F. J., Degtyareva, N. P., Lobachev, K., and Petes, T. D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* *120*, 587-598.
- Lobachev, K. S., Gordenin, D. A., and Resnick, M. A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* *108*, 183-193.
- Lobachev, K. S., Stenger, J. E., Kozyreva, O. G., Jurka, J., Gordenin, D. A., and Resnick, M. A. (2000). Inverted Alu repeats unstable in yeast are excluded from the human genome. *EMBO J.* *19*, 3822-3830.
- Ma, C., Martin, S., Trask, B., and Hamlin, J. L. (1993). Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. *Genes Dev.* *7*, 605-620.
- Malkova, A., Ivanov, E. L., and Haber, J. E. (1996). Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proc. Natl. Acad. Sci USA* *93*, 7131-7136.
- Maringele, L., and Lydall, D. (2004). Telomerase- and recombination-independent immortalization of budding yeast. *Genes Dev.* *18*, 2663-2675.

- McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* 26, 234-282.
- Mondello, C., Rebuzzini, P., Dolzan, M., Edmonson, S., Taccioli, G. E., and Giulotto, E. (2001). Increased gene amplification in immortal rodent cells deficient for the DNA-dependent protein kinase catalytic subunit. *Cancer Res.* 61, 4520-4525.
- Moore, I. K., Martin, M. P., Dorsey, M. J., and Paquin, C. E. (2000). Formation of circular amplifications in *Saccharomyces cerevisiae* by a breakage-fusion-bridge mechanism. *Environ. Mol. Mutagen.* 36, 113-120.
- Myung, K., and Kolodner, R. D. (2003). Induction of genome instability by DNA damage in *Saccharomyces cerevisiae*. *DNA Repair (Amst.)* 2, 243-258.
- Naeem, R. C. (2005). Cytogenetics of Hematologic Neoplasms, In *The Principles of Clinical Cytogenetics*, S. L. a. K. Gersen, M.B., ed. (Totowa, New Jersey: Humana Press), pp. 365-420.
- Paulson, T. G., Almasan, A., Brody, L. L., and Wahl, G. M. (1998). Gene amplification in a p53-deficient cell line requires cell cycle progression under conditions that generate DNA breakage. *Mol. Cell. Biol.* 18, 3089-3100.
- Pennaneach, V., Putnam, C. D., and Kolodner, R. D. (2006). Chromosome healing by de novo telomere addition in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 59, 1357-1368.
- Pipiras, E., Coquelle, A., Bieth, A., and Debatisse, M. (1998). Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. *EMBO J.* 17, 325-333.
- Poupon, M. F., Smith, K. A., Chernova, O. B., Gilbert, C., and Stark, G. R. (1996). Inefficient growth arrest in response to dNTP starvation stimulates gene amplification through bridge-breakage-fusion cycles. *Mol. Biol. Cell* 7, 345-354.
- Putnam, C. D., Pennaneach, V., and Kolodner, R. D. (2005). *Saccharomyces cerevisiae* as a model system to define the chromosomal instability phenotype. *Mol. Cell. Biol.* 25, 7226-7238.
- Rachidi, N., Martinez, M. J., Barre, P., and Blondin, B. (2000). *Saccharomyces cerevisiae* PAU genes are induced by anaerobiosis. *Mol. Microbiol.* 35, 1421-1430.
- Rattray, A. J., Shafer, B. K., Neelam, B., and Strathern, J. N. (2005). A mechanism of palindromic gene amplification in *Saccharomyces cerevisiae*. *Genes Dev.* 19, 1390-1399.
- Resnick, M. A., Westmoreland, J., and Bloom, K. (1990). Heterogeneity and maintenance of centromere plasmid copy number in *Saccharomyces cerevisiae*. *Chromosoma* 99, 281-288.

- Shimizu, N., Shingaki, K., Kaneko-Sasaguri, Y., Hashizume, T., and Kanda, T. (2005). When, where and how the bridge breaks: anaphase bridge breakage plays a crucial role in gene amplification and HSR generation. *Exp. Cell. Res.* *302*, 233-243.
- Stark, G. R., Debatisse, M., Giulotto, E., and Wahl, G. M. (1989). Recent progress in understanding mechanisms of mammalian DNA amplification. *Cell* *57*, 901-908.
- Stenger, J. E., Lobachev, K. S., Gordenin, D., Darden, T. A., Jurka, J., and Resnick, M. A. (2001). Biased distribution of inverted and direct Alus in the human genome: implications for insertion, exclusion, and genome stability. *Genome Res.* *11*, 12-27.
- Sutherland, G. R. (2003). Rare fragile sites. *Cytogenet Genome Res.* *100*, 77-84.
- Tanaka, H., Bergstrom, D. A., Yao, M. C., and Tapscott, S. J. (2005). Widespread and nonrandom distribution of DNA palindromes in cancer cells provides a structural platform for subsequent gene amplification. *Nat. Genet.* *37*, 320-327.
- Toledo, F., Buttin, G., and Debatisse, M. (1993). The origin of chromosome rearrangements at early stages of AMPD2 gene amplification in Chinese hamster cells. *Curr. Biol.* *3*, 255-264.
- van den Berg, M. A., and Steensma, H. Y. (1997). Expression cassettes for formaldehyde and fluoroacetate resistance, two dominant markers in *Saccharomyces cerevisiae*. *Yeast* *13*, 551-559.
- Wang, P., Kim, Y., Pollack, J., Narasimhan, B., and Tibshirani, R. (2005). A method for calling gains and losses in array CGH data. *Biostatistics* *6*, 45-58.
- Watanabe, T., and Horiuchi, T. (2005). A novel gene amplification system in yeast based on double rolling-circle replication. *EMBO J.* *24*, 190-198.
- Whittaker, S. G., Rockmill, B. M., Blechl, A. E., Maloney, D. H., Resnick, M. A., and Fogel, S. (1988). The detection of mitotic and meiotic aneuploidy in yeast using a gene dosage selection system. *Mol. Gen. Genet.* *215*, 10-18.
- Windle, B. E., and Wahl, G. M. (1992). Molecular dissection of mammalian gene amplification: new mechanistic insights revealed by analyses of very early events. *Mutat. Res.* *276*, 199-224.
- Yunis, J. J., Soreng, A. L., and Bowe, A. E. (1987). Fragile sites are targets of diverse mutagens and carcinogens. *Oncogene* *1*, 59-69.
- Zhou, Z. H., Akgun, E., and Jasin, M. (2001). Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proc. Natl. Acad. Sci. USA* *98*, 8326-8333.

2.9. Contribution to the publication

The presented study resulted in a first author publication in *CELL* journal. Majority of the experimental work was done by the first author. The paper was also written primarily by the author with the assistance of the academic advisor.

CHAPTER 3

Intrachromosomal gene amplification triggered by hairpin-capped breaks requires homologous recombination and is independent of non-homologous end joining

3.1. Summary

Gene amplification is one of the major mechanisms of acquisition of drug resistance and activation of oncogenes in tumors. In mammalian cells, amplified chromosomal regions are manifested cytogenetically as extrachromosomal double minutes (DMs) and chromosomal homogeneously staining regions (HSRs). We recently demonstrated using yeast model system that hairpin-capped double strand breaks (DSBs) generated at the location of human *Alu*-quasipalindromes can trigger both types of gene amplification. Specifically, the dicentric chromosomes arising from replication of hairpin-capped molecules can be precursors for intrachromosomal amplicons. The formation of HSRs can be accounted for either by breakage-fusion-bridge (BFB) cycle which necessitates nonhomologous end-joining pathway (NHEJ) or by the repair event involving homologous recombination (HR). In this study, we report that intrachromosomal gene amplification mediated by hairpin-capped DSBs is independent of NHEJ machinery, however requires the functions of Rad52 and Rad51 proteins. Based on our observations, we propose a HR-dependent mechanism to explain how the breakage of dicentric chromosomes can lead to the formation of HSRs.

3.2. Introduction

Increase in gene dosage via gene amplification plays a pivotal role in the acquisition and maintenance of malignant phenotype. In addition, the emergence of resistance to chemotherapeutic agents in tumor cells is mainly caused by amplification of genes involved in metabolism or detoxification of drugs¹⁻³. Cytogenetic studies in human and animal cancer cells have identified two types of topographical organization of the amplified DNA segments, namely, extrachromosomal double minutes (DM) that are either circular or rod shaped elements and intrachromosomal homogeneously staining regions (HSR). Intrachromosomal amplicons in metaphase spreads of human cancer cells stain with alternate dark and light bands throughout their length. Molecular analysis of HSRs revealed the common structural organization of the amplicons: inverted ladders comprising of regular units of head-to-head repeats⁴⁻⁶. Chromosomes carrying HSRs are also often characterized by deletion of telomere-proximal regions⁴.

The mechanisms underlying the generation of intrachromosomal amplicons are not well understood. However, based on the observation that fused sister chromatids and anaphase bridges occur in cell populations undergoing amplification⁷⁻¹³, the breakage-fusion-bridge (BFB) cycle, proposed by B. McClintock in 1941¹⁴, was adopted as the primary mechanism for HSRs formation^{4,15}. The triggering step in this model is the formation of a dicentric chromosomal intermediate which is a prerequisite for the amplification. Studies in yeast and mammalian cells have demonstrated that dicentrics can arise in multiple ways. First, dicentric chromosomes can result from fusion between sister chromatids following replication of the broken DNA in accordance with the BFB mechanism. It has been shown that dicentrics can be generated following I-SceI

endonuclease targeted double-strand breaks (DSBs) in mammalian chromosomes¹⁶⁻¹⁸. Studies involving analysis of amplicon junctions revealed the presence of micro-homologies, implicating the role of NHEJ pathway in the generation of dicentrics¹⁹. Second, in cells with compromised or defective telomere functions, recognition of unprotected termini as DSBs resulted in dicentrics via terminal chromosomal fusions that requires NHEJ proteins^{15,20}. Third, it has been demonstrated that resection of HO endonuclease-mediated DSB can lead to large dicentric inverted dimers in yeast. Specifically, such sister chromatid fusions were a consequence of intermolecular single-strand annealing between two inverted repeats including mobile elements such as Ty retrotransposons²¹. Fourth, in mammalian and yeast cells, spontaneous or induced DSB adjacent to inverted repeats can promote the formation of large dicentric palindromes. This process involves foldback priming via intrastrand annealing between inverted repeats and subsequent synthesis²²⁻²⁷. Finally, dicentric formation can be an outcome of DSBs triggered by inverted repeats. A quasipalindrome comprised of two human *Alu* repeats induces hairpin-capped DSBs in yeast *Saccharomyces cerevisiae*²⁸. The unprocessed hairpin-capped broken intermediate can be converted to dicentric molecule following replication^{28,29}.

At anaphase, the DNA bridging the two centromeres can break owing to repelling of the centromeres in opposite directions. Two broken molecules will be produced, one of which will contain the duplicated regions of DNA organized in inverted orientation at the broken end. The duplicated fragment accounts for a single repeat unit that is characteristic of early stages of amplification. However, HSRs found in advanced cancer cells are comprised of a tandem array of such repeat units, indicating the existence of

subsequent processes^{1,4}. The multiple copies of duplicated DNA from the initial broken dicentric can be obtained via BFB model that involves reiterated formation of dicentric molecules by fusions between broken sister chromatids⁴. Repetitive fusions could be potentially mediated by NHEJ machinery¹⁵. The BFB model has been accepted as sufficient to explain intrachromosomal gene amplification. However recent data have called this assumption into question. DNA-PKcs-deficient mouse embryo fibroblast cells have a high frequency of gene amplification leading to HSRs³⁰. In addition, gene amplification is readily observed in pro-B lymphomas in mice with compromised NHEJ³¹. These indicate that NHEJ might not be the only mechanism for generating intrachromosomal amplicons.

We have developed an experimental system in yeast to follow the fate of the dicentric molecules derived from replication of palindrome-mediated hairpin-capped DSBs. We recently showed that a quasipalindrome comprised of two human *Alu* repeats triggers intrachromosomal gene amplification in yeast²⁹. Structural analysis of intrachromosomal amplicons revealed the presence of delta elements bordering the amplified regions and the repeated units within the amplicons. Based on this observation we proposed that intrachromosomal amplification could result from repair of the broken dicentric mediated by homologous recombination between delta elements. In this study, we delineate the contributions of NHEJ and HR in the formation of intrachromosomal amplicons. We demonstrate that intrachromosomal gene amplification triggered by inverted *Alu* repeats is *DNL4*-, *HDF1*- and *HDF2*-independent, however, is greatly compromised in *RAD51*- and *RAD52*-deficient strains. We propose a model to explain how homologous recombination can contribute towards the generation of HSRs.

3.3. Results and discussion

3.3.1. Experimental system to study intrachromosomal gene amplification mediated by hairpin-capped dsbs

Intrachromosomal gene amplification events were analyzed using TD strains described in detail in Narayanan et al 2006. Briefly, two gene dosage markers *CUPI* (encoding copper binding metallothionein) and *SFAI* (encoding formaldehyde dehydrogenase) were placed between the *Alu*-IRs, the site of hairpin-capped DSBs, and the centromere on chromosome V left arm in haploid yeast strain (**Fig. 3.1**). Isolates with intrachromosomal amplification of the centromere-proximal regions including *CUPI* and *SFAI* can be selected on medium containing high concentration of copper and formaldehyde ($\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$). This experimental system selects for greater than two copies of the intrachromosomal amplicons. Two additional markers, *CANI* and *ADE2* placed adjacent to *Alu*-IRs allow us to verify the presence of telomere-proximal regions in the intrachromosomal amplicons. $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates carrying the 43 kb terminal deletion of left arm of chrV will also be canavanine-resistant and red in color ($\text{Can}^{\text{S}}\text{Ade}^-$).

We recently demonstrated that hairpin-capped DSBs occurring at the site of inverted repeats in wild type strains greatly stimulate intrachromosomal gene amplification²⁹. In the majority of the $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ isolates (94%) obtained from wild type strains, intrachromosomal amplification was accompanied by the loss of telomere proximal region ($\text{Can}^{\text{S}}\text{Ade}^-$). Structural analysis of intrachromosomal amplicons revealed that they were organized as a tandem array of inverted repeats present in four or more copies. Based on the genetic and structural analyses of intrachromosomal gene amplification events, the following sequence of events was deduced. *Alu*-IRs extrude

into cruciform structures that are processed into hairpin-capped breaks by a cruciform-resolving enzyme. Unprocessed centromere-containing fragments are replicated to yield dicentric molecule.

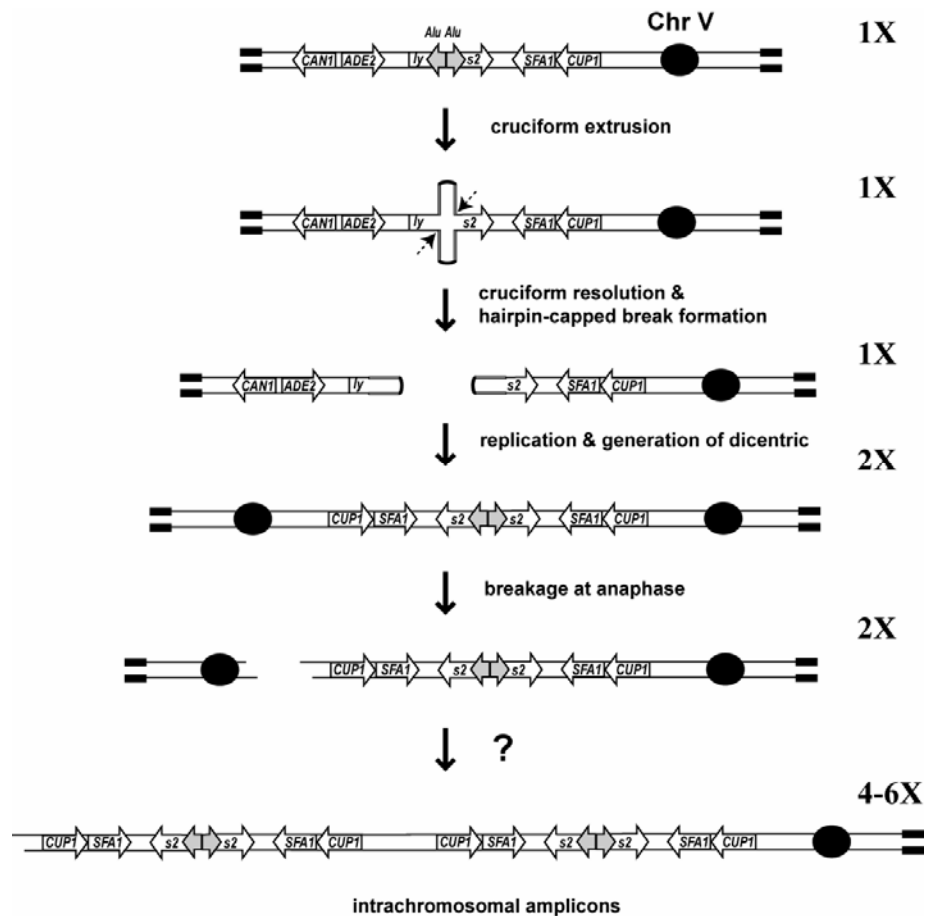


Figure 3.1. Experimental system to study intrachromosomal amplification resulting from hairpin-capped DSBs and the formation of dicentric intermediate. Selection for copper- and formaldehyde-resistant derivatives ($\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$) results in 4- to 6-fold intrachromosomal amplification of the regions adjacent to the *Alu* repeats. Majority of $\text{Cu}^{\text{R}}\text{Fh}^{\text{R}}$ clones are also canavanine-resistant and Ade^- auxotroph carrying 42 kb telomeric deletion. The mechanism of rearrangements is based on the structural analysis of the intermediates and final rearrangements²⁹. Open arrows depict relevant genes. Solid gray arrows correspond to *Alu* repeats. Filled squares and circles indicate telomeres and centromeres, respectively. Change in a copy number of *CUP1* and *SFA1* during rearrangements is shown.

The unstable dicentric breaks at anaphase and subsequently triggers the generation of intrachromosomal amplicons (**Fig.3.1**). In order to discriminate between the roles of NHEJ and HR machinery in mediating intrachromosomal gene amplification following dicentric breakage, we introduced mutations in key players in both pathways. We determined the rate of formation and the amplicon structure of Cu^RFh^R isolates in the mutant strains containing inverted *Alu* repeats.

3.3.2. *DNL4*, *HDF1* and *HDF2* are not required for inverted repeat-induced intrachromosomal gene amplification

We estimated the rates of intrachromosomal gene amplification induced by inverted repeats in strains containing disruptions of *DNL4*, *HDF1* and *HDF2*. In the NHEJ repair pathway, the ring shaped Hdf1/Hdf2 heterodimer recognizes and binds to DSB ends. The bound complex then recruits Dnl4 and Lif1 that carry out the ligation step following processing of the DSB termini³². In yeast, mutation in these proteins, severely compromise the repair of DSB ends via NHEJ pathway.

If the repair of the broken dicentric requires NHEJ, then we would expect to see a reduction in the rate of intrachromosomal amplification in NHEJ mutants. As shown in **Table 3.1**, the rates of intrachromosomal gene amplification in NHEJ mutants were similar to wild-type strain. Majority of the Cu^RFh^R clones obtained in these mutants had a terminal deletion of 43 kb region between *LYS2* region and left telomere. The structure of the amplicons, determined for Cu^RFh^R clones from $\Delta hdf2$ strains were also similar to the nature of amplicon arrangement detected in wild type strains. Physical

characterization of intrachromosomal amplification events observed in these Cu^RFh^R isolates using CHEF analysis and Southern blot hybridization (**Fig.3.2**) revealed chromosome V to be heavier than the wild type chromosome suggesting the presence of intrachromosomally amplified regions.

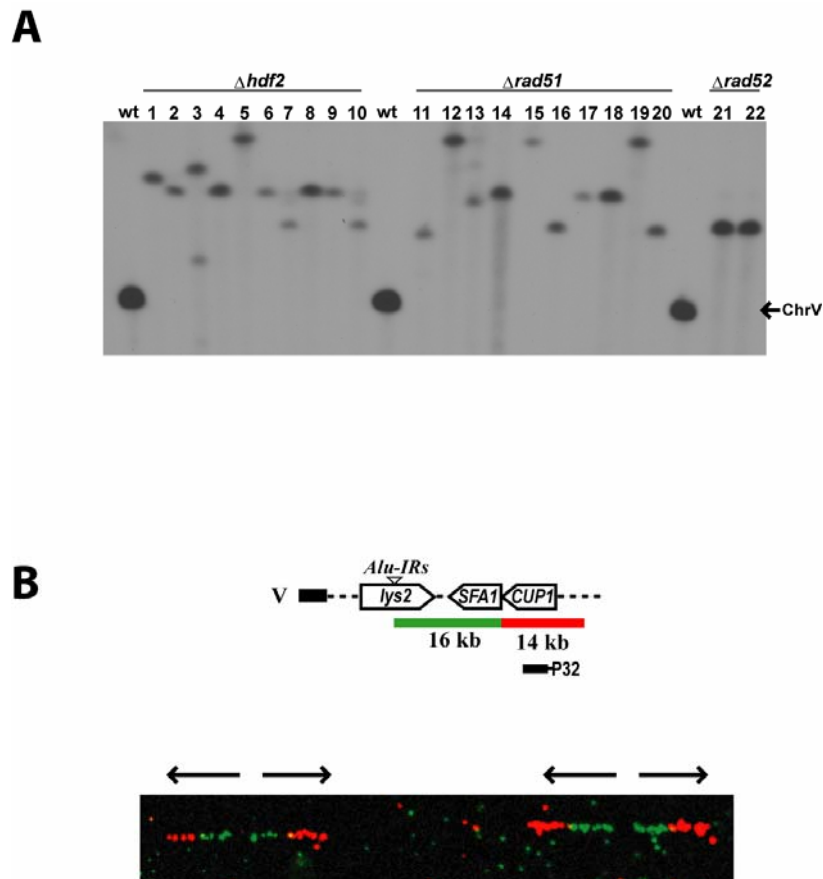


Figure 3.2. Analysis of intrachromosomal amplification resulting from the breakage of the dicentric in $\Delta hdf2$, $\Delta rad51$ and $\Delta rad52$ mutants.

(A) Southern analysis of a CHEF gel containing DNA from Cu^RFh^R isolates of $\Delta hdf2$ (lanes 1 to 10), $\Delta rad51$ (lanes 11 to 20), and $\Delta rad52$ (lanes 21 and 22) mutants. The lanes labeled “wt” contain DNA from TD strain. The gel was hybridized with a *CUP1*-specific probe.

(B) FISH/molecular combing analyses of Cu^RFh^R isolates from $\Delta hdf2$ strain. The upper panel shows the regions on chromosome V that are fluorescently-labeled for FISH and ³²P-labeled for Southern analyses. The bottom panel is an example of an amplicon visualized by molecular combing and FISH (isolate#1). This isolate contains 4X inverted array. Similar amplification pattern was detected for four randomly chosen independent isolates from $\Delta hdf2$ strain (#2, #5, #6 and #9) and for three isolates from $\Delta rad51$ strain (#11, #14 and #15). In addition, intrachromosomal amplicons carrying 6 copies of the amplified segments arranged as an inverted ladder were also detected in isolates #12 and #19 from $\Delta rad51$ mutant.

Molecular combing and dual-color fluorescent in situ hybridization (FISH)³³ performed on genomic DNA from four Cu^RFh^R variants derived from *Δhdf2* strains demonstrated the inverted repeat-like organization of the amplicons. Hence we conclude that hairpin-capped DSB induced intrachromosomal gene amplification in yeast does not require NHEJ proteins.

3.3.3. Intrachromosomal gene amplification depends on homologous recombination

Previously, we have demonstrated using CGH on microarrays that *Alu-IR*-induced intrachromosomal amplicons have delta elements at their borders²⁹. This observation suggested the role of homologous recombination in the generation of HSRs. We assessed the intrachromosomal amplification potential in *Δrad51* and *Δrad52* mutants. Rad52 binds to single-stranded DNA and facilitates the displacement of replication protein-A by the recombinase, Rad51. Rad51 is required for the strand invasion and exchange reaction. In addition to gene conversion, Rad52 also participates in single-strand annealing pathway of DSB repair. Consequently, both *Δrad51* and *Δrad52* mutants are defective in gene conversion while, *Δrad52* mutants also exhibit impaired SSA³⁴. Break-induced replication (BIR), another type of DSB repair, is also severely compromised in *Δrad52* mutants^{35,36} but, at least in some cases can occur in the absence of Rad51^{37,38}.

Disruption of both Rad51 and Rad52 led to a decrease in the rate of intrachromosomal amplicon generation. The rates of amplification in *Δrad51* and *Δrad52* strains were 53-fold and 160-fold lower than in wild type strains (**Table 3.1**).

Intrachromosomally amplified regions in Cu^RFh^R isolates of *Δrad51* mutants were

arranged as inverted ladders similar to the amplicon structure observed in wild type and *Ahd2* strains (**Fig. 3.2**). These results establish that in *Arad51* mutants, intrachromosomal amplicons are generated by similar pathways that function in wild type cells, albeit with lower efficiency.

Interestingly, the rare Cu^RFh^R isolates derived from *Arad52* strains were canavanine sensitive Ade⁺ prototrophs, indicating the presence of telomere proximal regions of chromosome V left arm. This conclusion was supported by Southern analysis (data not shown). CHEF and subsequent Southern blot hybridization with a probe specific for chromosome V also unveiled larger sized chromosomes carrying intrachromosomal amplifications (**Fig. 3.2**). The existence of telomere-proximal *CANI* region in the novel chromosomes carrying amplified segments negate the formation of intrachromosomal amplicons by mechanisms that operate in wild type involving the repair of the broken dicentric. We propose that the intrachromosomal amplification, in *Arad52* mutants, can be mediated by alternative mechanisms, such as interstitial DNA synthesis.

Overall, these data firmly support the role of homologous recombination in the generation of HSR-like amplicons in yeast. We propose that after breakage of the dicentric, the broken DNA could invade a homologous repeat on the same chromosome arm and thereby initiate BIR producing tandem head-to-head repeats (**Fig.3.3**). Ty or delta elements can provide homology for the recombination event. This is supported by our observation that delta elements border the amplified regions and the repeated units within the amplicons²⁹. Depending on the orientation of the recombining delta elements and their position relative to the amplified genes as well as centromere, additional

amplicon units will be achieved in two ways. If BIR is initiated in a direction opposing the centromere and towards the amplified genes, a “rolling” circle replication intermediate would be set up giving rise to at least four copies of the amplicon.

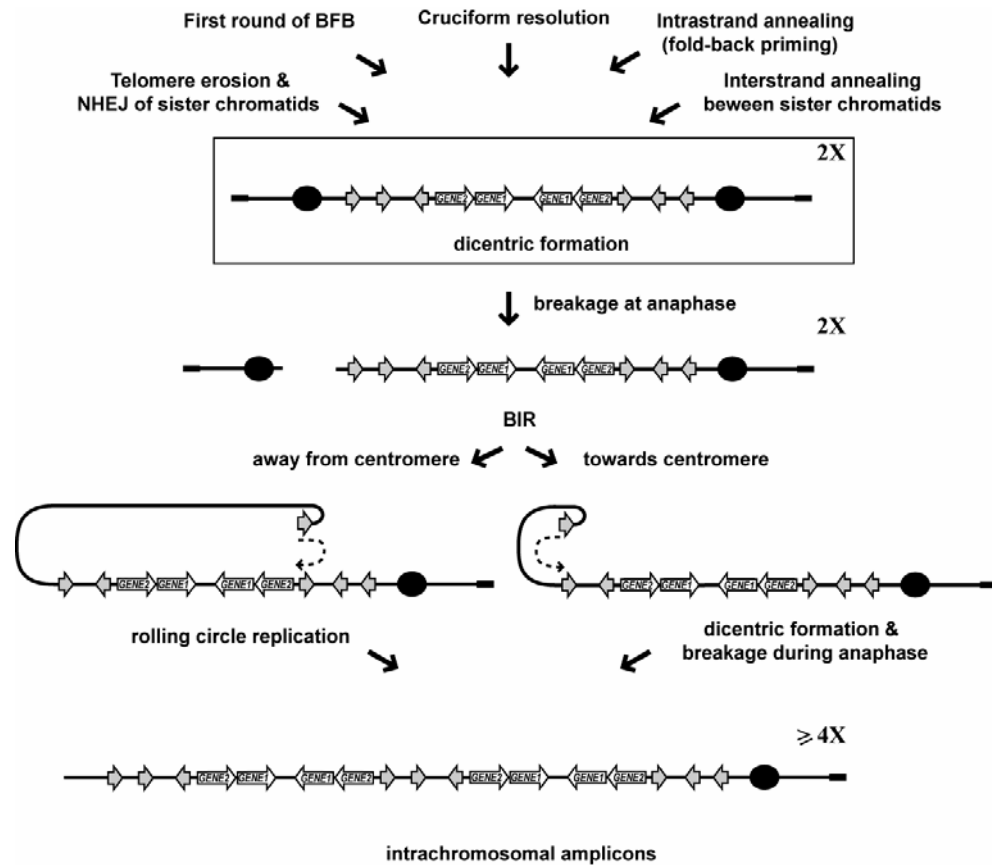


Figure 3.3. Model for BIR-mediated intrachromosomal amplification. Amplified genes are shown as open arrows. Solid gray arrows indicate the orientation and location of the repetitive sequences on a dicentric chromosome. A detailed description is presented in the text.

Alternatively, the broken DNA could invade a homologous repeat in a direct orientation, commencing BIR towards centromere thereby generating a new dicentric. The breakage during anaphase might lead to gain of additional copies. In both cases, the resulting intrachromosomal amplicons will be arranged as inverted ladders. Consistent with the proposed model, inhibition of BIR in *Δrad52* and *Δrad51* mutants impairs the

HSR formation. The stronger effect of the *Arad52* vs *Arad51* mutation on the amplification potential can be explained by the fact that BIR only partly depends on Rad51^{37,38}. It is important to note, in the proposed mechanism, the increase in a copy number is acquired in a single cellular division.

We find that HR promotes intrachromosomal gene amplification in the case of cruciform resolution and the formation of hairpin-capped molecules in yeast. It is possible that HR can also play a role in gene amplification when dicentrics are formed due to other mechanisms such as telomere-telomere fusions, fold-back priming or interstrand annealing, and this might apply to higher eukaryotes as well. The contribution of HR to HSR formation in mammalian cells is yet to be investigated in detail. It is conceivable that NHEJ is the primary mechanism in intrachromosomal gene amplification since it is robust in mammalian cells³². However, under certain conditions, such as, when NHEJ is compromised, it is plausible that the HSRs can be generated via HR-dependent pathways. The mammalian genomes are enriched with repetitive elements that can provide platform for BIR following dicentric breakage. We therefore propose that HR might be an alternative pathway to NHEJ in amplicon generation.

3.4. Table 3.1. Intrachromosomal gene amplification in wild type, NHEJ and HR mutants.

Strains	Rate of gene amplification (x 10 ⁸)	Fold decrease relative to wild type
Wild type	8 (7-9)*	1.0
<i>Δhdf2</i>	6 (6-9)	1.3
<i>Δhdf1</i>	6 (5-7)	1.3
<i>Δdnl4</i>	8 (7-10)	1.0
<i>Δrad51</i>	0.15 (0.08-0.25)	53.3
<i>Δrad52</i>	0.05 (0.01-0.15)	160

* Numbers in parentheses correspond to the 95% confidence interval.

3.5. References

1. Albertson DG. Gene amplification in cancer. *Trends Genet* 2006; 22:447-55.
2. Fletcher JA. Cytogenetics of Solid Tumors. In: Gersen SLaK, M.B., ed. *The Principles of Clinical Cytogenetics*. Totowa, New Jersey: Humana Press, 2005:421-51.
3. Naeem RC. Cytogenetics of Hematologic Neoplasms. In: Gersen SLaK, M.B., ed. *The Principles of Clinical Cytogenetics*. Totowa, New Jersey: Humana Press, 2005:365-420.
4. Debatisse MaM, B. Gene amplification mechanisms. In: Nigg EA, ed. *Genome Instability in Cancer Development*. Netherlands: Springer, 2005:343-61.
5. Stark GR, Debatisse M, Giulotto E, Wahl GM. Recent progress in understanding mechanisms of mammalian DNA amplification. *Cell* 1989; 57:901-8.
6. Windle BE, Wahl GM. Molecular dissection of mammalian gene amplification: new mechanistic insights revealed by analyses of very early events. *Mutat Res* 1992; 276:199-224.
7. Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* 1997; 89:215-25.
8. Fouladi B, Sabatier L, Miller D, Pottier G, Murnane JP. The relationship between spontaneous telomere loss and chromosome instability in a human tumor cell line. *Neoplasia* (New York, NY 2000; 2:540-54.

9. Lo AW, Sabatier L, Fouladi B, Pottier G, Ricoul M, Murnane JP. DNA amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line. *Neoplasia* (New York, NY 2002; 4:531-8.
10. Ma C, Martin S, Trask B, Hamlin JL. Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. *Genes Dev* 1993; 7:605-20.
11. Sabatier L, Ricoul M, Pottier G, Murnane JP. The loss of a single telomere can result in instability of multiple chromosomes in a human tumor cell line. *Mol Cancer Res* 2005; 3:139-50.
12. Toledo F, Buttin G, Debatisse M. The origin of chromosome rearrangements at early stages of AMPD2 gene amplification in Chinese hamster cells. *Curr Biol* 1993; 3:255-64.
13. Toledo F, Smith KA, Buttin G, Debatisse M. The evolution of the amplified adenylate deaminase 2 domains in Chinese hamster cells suggests the sequential operation of different mechanisms of DNA amplification. *Mutat Res* 1992; 276:261-73.
14. McClintock B. The stability of broken ends of chromosomes in *Zea mays*. *Genetics* 1941; 26:234-82.
15. Murnane JP. Telomeres and chromosome instability. *DNA repair* 2006; 5:1082-92.
16. Coquelle A, Rozier L, Dutrillaux B, Debatisse M. Induction of multiple double-strand breaks within an hsr by meganucleaseI-SceI expression or fragile site activation leads to formation of double minutes and other chromosomal rearrangements. *Oncogene* 2002; 21:7671-9.

17. Lo AW, Sprung CN, Fouladi B, Pedram M, Sabatier L, Ricoul M, Reynolds GE, Murnane JP. Chromosome instability as a result of double-strand breaks near telomeres in mouse embryonic stem cells. *Molecular and cellular biology* 2002; 22:4836-50.
18. Pipiras E, Coquelle A, Bieth A, Debatisse M. Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. *Embo J* 1998; 17:325-33.
19. Okuno Y, Hahn PJ, Gilbert DM. Structure of a palindromic amplicon junction implicates microhomology-mediated end joining as a mechanism of sister chromatid fusion during gene amplification. *Nucleic acids research* 2004; 32:749-56.
20. Williams B, Lustig AJ. The paradoxical relationship between NHEJ and telomeric fusion. *Molecular cell* 2003; 11:1125-6.
21. VanHulle K, Lemoine FJ, Narayanan V, Downing B, Hull K, McCullough C, Bellinger M, Lobachev K, Petes TD, Malkova A. Inverted DNA repeats channel repair of distant double-strand breaks into chromatid fusions and chromosomal rearrangements. *Molecular and cellular biology* 2007; 27:2601-14.
22. Albrecht EB, Hunyady AB, Stark GR, Patterson TE. Mechanisms of *sod2* gene amplification in *Schizosaccharomyces pombe*. *Mol Biol Cell* 2000; 11:873-86.
23. Butler DK, Yasuda LE, Yao MC. Induction of large DNA palindrome formation in yeast: implications for gene amplification and genome stability in eukaryotes. *Cell* 1996; 87:1115-22.
24. Maringele L, Lydall D. Telomerase- and recombination-independent immortalization of budding yeast. *Genes Dev* 2004; 18:2663-75.

25. Rattray AJ, Shafer BK, Neelam B, Strathern JN. A mechanism of palindromic gene amplification in *Saccharomyces cerevisiae*. *Genes Dev* 2005; 19:1390-9.
26. Tanaka H, Cao Y, Bergstrom DA, Kooperberg C, Tapscott SJ, Yao MC. Intrastrand annealing leads to the formation of a large DNA palindrome and determines the boundaries of genomic amplification in human cancer. *Molecular and cellular biology* 2007; 27:1993-2002.
27. Tanaka H, Tapscott SJ, Trask BJ, Yao MC. Short inverted repeats initiate gene amplification through the formation of a large DNA palindrome in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 2002; 99:8772-7.
28. Lobachev KS, Gordenin DA, Resnick MA. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* 2002; 108:183-93.
29. Narayanan V, Mieczkowski PA, Kim HM, Petes TD, Lobachev KS. The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. *Cell* 2006; 125:1283-96.
30. Mondello C, Rebuzzini P, Dolzan M, Edmonson S, Taccioli GE, Giulotto E. Increased gene amplification in immortal rodent cells deficient for the DNA-dependent protein kinase catalytic subunit. *Cancer Res* 2001; 61:4520-5.
31. Zhu C, Mills KD, Ferguson DO, Lee C, Manis J, Fleming J, Gao Y, Morton CC, Alt FW. Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* 2002; 109:811-21.

32. Burma S, Chen BP, Chen DJ. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA repair* 2006; 5:1042-8.
33. Conti C, Caburet, S., Schurra, C., Bensimon, A. *Molecular Combing. Current Protocols in Cytometry*. New York: John Wiley & Sons, Inc, 2001:8.10.1-8.23.
34. Krogh BO, Symington LS. Recombination proteins in yeast. *Annu Rev Genet* 2004; 38:233-71.
35. Davis AP, Symington LS. RAD51-dependent break-induced replication in yeast. *Molecular and cellular biology* 2004; 24:2344-51.
36. Malkova A, Ivanov EL, Haber JE. Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* 1996; 93:7131-6.
37. Malkova A, Signon L, Schaefer CB, Naylor ML, Theis JF, Newlon CS, Haber JE. RAD51-independent break-induced replication to repair a broken chromosome depends on a distant enhancer site. *Genes Dev* 2001; 15:1055-60.
38. Signon L, Malkova A, Naylor ML, Klein H, Haber JE. Genetic requirements for RAD51- and RAD54-independent break-induced replication repair of a chromosomal double-strand break. *Molecular and cellular biology* 2001; 21:2048-56.

3.6. Contribution to the publication

The presented study resulted in a first author publication in *CELL CYCLE* journal. All of the experimental work was done by the first author. The paper was also written primarily by the author with the assistance of the academic advisor.

CHAPTER 4

Post-replication repair mediates the inverted repeat fragility under conditions of compromised replication.

4.1. Summary

Palindromic sequences that adopt hairpin and cruciform structures are potent inducers of chromosomal fragility and genome rearrangements. In replication deficient strains of yeast *Saccharomyces cerevisiae*, we observe elevated levels of replication fork stalling and hairpin-capped double-strand break formation at the site of *Alu*-quasi palindromes. We show that the increased instability observed in replication mutants is dependent on post-replicative repair pathway that requires the Rad5 ATPase activity and proliferating cell nuclear antigen modification at lysine 107 residue. We have found that the yeast replication-pausing checkpoint proteins, Tof1-Csm3 and Mrc1, signal for template switching or double-strand generation when replication fork arrest is imposed by palindromes. We suggest that the requirements for replication-associated fragility at inverted repeats are different from those that operate during replication arrest at the protein-bound barriers, DNA lesions and other secondary structure-forming repeats.

4.2. Introduction

Fidelity of DNA replication is crucial for the maintenance of the genome stability. The replication fork movement is compromised at several genomic locations such as protein-bound replication fork barriers (Calzada et al., 2005), replication slow down zones (Cha and Kleckner, 2002), tRNA genes (Deshpande and Newlon, 1996), bidirectional barriers formed at the centromeres (Greenfeder and Newlon, 1992) and at secondary structure forming repetitive sequences (Krasilnikova and Mirkin, 2004; Voineagu et al., 2008; Zhang and Freudenreich, 2007). Fork stalling can also result from defects in the replisome. Paucity in the deoxyribonucleotide pools, discoordination between the DNA synthesis and the unwinding of duplex DNA ahead of the fork, uncoupling between leading and lagging strand DNA synthesis and obstructing the replicative helicase progression can also impede replication fork movement (reviewed in Branzei and Foiani, 2007a). Replication arrest can culminate in chromosomal rearrangements as a consequence of recombination events at the collapsed forks (Labib and Hodgson, 2007).

Eukaryotic cells have evolved multiple pathways that govern the stability of the replication fork and ensure unperturbed fork progression. Non-replicative DNA helicases aid the replication machinery to evade fork barriers. Rrm3 is required for displacing the fork binding Fob1 that recognizes and binds directly to replication fork barriers (RFB) and surrounding sites in the rDNA in yeast *Saccharomyces cerevisiae* (Kobayashi, 2003). A recent study demonstrated that Rrm3 helicase also aids in normal fork progression through more than 1000 discrete sites scattered throughout the yeast genome (Azvolinsky et al., 2006). Studies in budding yeast revealed that Srs2, a 3' to 5' helicase is recruited to the arrested replication forks by SUMOylated proliferating cell nuclear antigen

(PCNA). Srs2 plays an antirecombinagenic role by displacing Rad51 presynaptic filament and thereby suppressing homologous recombination at the stall sites (Sung and Klein, 2006). RecQ family of helicases have also been implicated to assist in the efficient traverse of the replication fork. Similar to Srs2, Sgs1 helicase in a complex with Top3 was shown to be an antagonist of crossovers that occurred spontaneously or those that were induced by DSBs (Ira et al., 2003). Moreover, Sgs1-Top3 complex was proposed to stimulate the processing of hemicatene-like structures generated at the sister chromatid junctions during template switching or as consequence of collision between two replication forks (Branzei and Foiani, 2007b). A similar activity was demonstrated *in vitro* for BLM, the orthologue of Sgs1 in humans (Wu and Hickson, 2003).

Besides recruiting helicases that ensure efficient restart of the halted fork and suppress unstable recombination events, the eukaryotic cells also possess mechanisms of lesion bypass that allow the completion of DNA replication in the presence of unrepaired damage. The protein complex system that coordinates the lesion bypass are the *RAD6* group of genes (Barbour and Xiao, 2003). It has been demonstrated that the post replication repair (PRR) is required to bypass lesions that are caused by UV or chemical agents in damage tolerance pathway (Klein, 2007; Smirnova and Klein, 2003; Torres-Ramos et al., 2002). The enzymatic system includes the members of ubiquitin conjugation system that cooperate in the modification of the eukaryotic processivity clamp for DNA polymerases, PCNA. Upon stalling of replication forks or DNA damage, the PCNA is mono-ubiquitinated at the highly conserved lysine (K) residue 164 by the ubiquitin-conjugating enzyme (E2) Rad6 and the ubiquitin ligase (E3) Rad18 (Hoege et al., 2002). Interesting, K164 residue is also targeted for SUMOylation in an Ubc9- and

Siz1-dependent manner during S-phase (Branzei et al., 2006). SUMOylated PCNA helps in the activation of Srs2 that suppresses recombination at collapsed forks (Papouli et al., 2005). PCNA is also SUMOylated at a second less conserved lysine 127 residue (Branzei et al., 2004). In addition to being monoubiquitinated, PCNA is polyubiquitinated at K164, a process that requires RING-finger ubiquitin ligase, Rad5 and the ubiquitin conjugating complex, Ubc13 and Mms2 (Ulrich and Jentsch, 2000). Ubc13 and Mms2 form a heterodimeric E2 enzyme that catalyzes specifically K63-linked polyubiquitin chains to K164 residue of PCNA. Monoubiquitination of PCNA facilitates the recruitment of translesion polymerases that perform an error-prone bypass of the lesion. Studies from yeast and humans demonstrated that translesion polymerases such as Pol η , Pol ζ and Pol ι interact tightly with monoubiquitylated PCNA (reviewed in (Andersen et al., 2008)). Polyubiquitination of the PCNA is required for the error-free damage bypass via the post replication repair (PRR) (Haracska et al., 2004; Hoegge et al., 2002) pathway. The Rad5 protein has the RING-finger domain responsible for E3 activity and an embedded helicase-like domain of the SWI/SNF family of helicases (Chen et al., 2005). Both the ubiquitin ligase activity and helicase activity were shown to be vital for template switching (Blastyak et al., 2007; Gangavarapu et al., 2006).

One of the earliest responses to replication stress is the activation of checkpoint kinases that are an integral part of the complex checkpoint surveillance machinery. The checkpoint response ensures the stability of the replication fork by modifying and recruiting proteins that help in the bypass of replication barriers, suppress recombination events and promote efficient fork restart. In *S. cerevisiae*, there are two key protein kinases that are involved in the replication stress response. These include, the sensor

protein kinase Mec1 (homolog of human ATR) that acts together with Ddc2 and phosphorylates the effector kinase Rad53p (homolog of human Chk1) (Branzei and Foiani, 2007a). The signal is transduced from the sensor to the effector via the mediator proteins. Mrc1 (mediator of checkpoint response), the orthologue of Claspin in humans, transduces checkpoint signal from Mec1 sensor protein to the Rad53 effector kinase. The phosphorylated Rad53 subsequently controls downstream events that maintain the normal rate of progression of DNA replication forks and prevents the firing from late-replicating origins (Tourriere et al., 2005). A functionally related protein complex is the Tof1-Csm3 complex. This protein complex is imperative for the site specific stalling at the rDNA locus by antagonizing Rrm3 helicase (Mohanty et al., 2006), albeit both Mrc1 and Tof1 are needed for pausing at tRNA genes (Calzada et al., 2005). Mrc1 and Tof1 are also required to alleviate the replication arrest by hairpin-structures on a plasmid (Voineagu et al., 2008). Besides functioning as mediator proteins for the checkpoint cascade, recent data have implicated Mrc1 and Tof1 as necessary factors for normal replication progression (Tourriere et al., 2005). This structural role is attributed to the regulation of Cdc45-MCM functions that coordinate the DNA polymerase and helicase activities.

Secondary structure-forming DNA repeats can hamper normal replication progression (Zhang and Freudenreich, 2007; Voineagu et al., 2008). Hairpin- and cruciform-forming inverted repeats stall replication fork on a plasmid in both prokaryotes and eukaryotes. In this study we demonstrate that palindromic sequences act as potent chromosomal replication barriers in yeast strains with compromised replisome. We demonstrate that the elevated levels of DSB formation at the location of the *Alu*-quasipalindrome in replication mutants requires the checkpoint activity of Mrc1 and

Tof1-Csm3 protein complex and the template switch mechanism mediated by PCNA modification at K107 residue and Rad5 helicase.

4.3. Results

4.3.1. Experimental system to study gross chromosomal rearrangements and chromosomal fragility resulting from inverted repeats.

We employed the previously developed experimental system based on the loss of *CAN1* and *ADE2* genes located on chromosome V (**Figure 4.1**) to monitor the hairpin-capped break formation and the resulting GCRs (Narayanan et al., 2006).

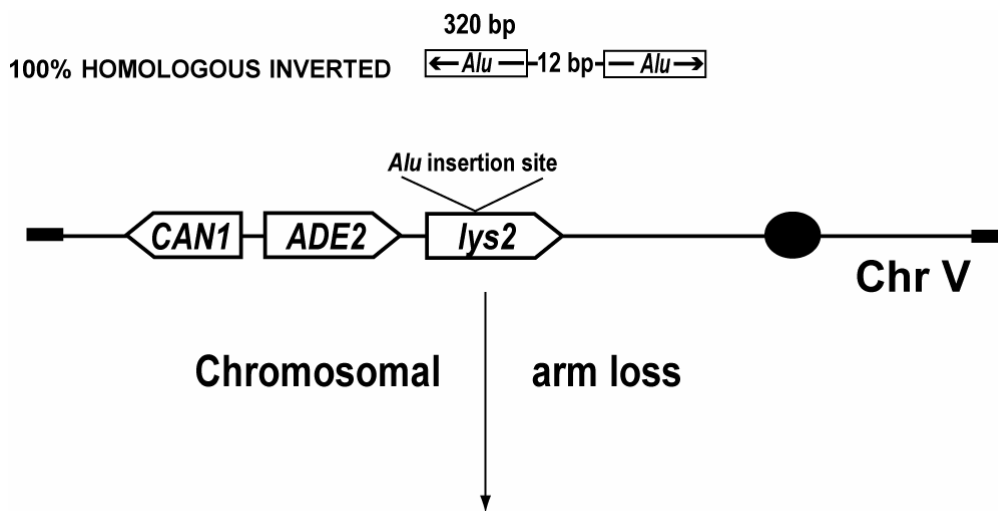


Figure 4.1. Experimental systems to study chromosomal fragility induced by *Alu*-quasipalindromes. 100 % identical inverted *Alu* repeats (open boxes with arrows inside) were inserted into a *Bam*HI site of the *LYS2* gene. The breakage at the location of *Alu*-IRs can lead to 43 kb telomere-proximal deletion resulting in Can^RAde⁻ clones.

Briefly, haploid yeast strains were constructed where the left arm of chromosome V in the region of *CAN1* gene was modified. A *LYS2* cassette containing 100% homologous

inverted repeats (*Alu*-IRs), that are 320 bp long and separated by a 12 bp spacer, were placed centromere-proximal to *CAN1*, such that the region between *LYS2* and the telomere does not contain essential genes and can be lost. The *ADE2* gene was moved telomere-distal to *CAN1*. A hairpin-capped break at the location secondary structure-forming *Alu*-IRs can cause deletion of the chromosome V region including *CAN1* and *ADE2* resulting in canavanine-resistant red colonies (Can^RAde⁻).

4.3.2. Defects in DNA replication leads to elevated levels of chromosomal fragility and GCR formation.

In a screen to identify mutants that had increased levels of hairpin-capped DSB formation, *pol3-P664L* mutation was discovered. The *pol3-P664L* strains exhibited elevated rates of *CAN1* loss (**Table 4.1**). There was a 32-fold induction of GCRs over wild type levels. A similar observation was made when the expression of *POL3* was down regulated using repressible *tetO* promoter, indicating that the *pol3-P664L* mutation might reflect the reduced levels of functional polymerase δ . To test if a comparable result can be obtained by down regulating other key replication proteins involved in lagging strand synthesis, we suppressed the expression of *POL1* (encoding polymerase α) and *POL30* (encoding PCNA) using tetracycline repressible promoters. Down regulation of polymerase α and PCNA were synergistic with *Alu*-IR fragility (**Table 4.1**).

Surprisingly, either mutation in catalytic domain of polymerase ϵ (responsible for leading strand synthesis (Pursell et al., 2007)), *pol2-M644I*, or reduced expression of *POL2* gene also led to an increase in arm loss events (**Table 4.1**). These data suggest that problems

in either leading or lagging strand synthesis are synergistic with inverted repeats-induced GCRs.

4.3.3. Replication fork stalling at *Alu*-IRs in replication mutants triggers elevated hairpin-capped DSB formation.

The increase in *CANI* loss in replication mutants likely reflects the consequence of stalled fork at the site of secondary structure and DSB formation. To assess this directly, we carried out two dimensional (2D) gel analyses of replication fork intermediates in wild type and *pol3-P664L* strains. Replication across *Alu*-quasipalindrome in wild type strains was not altered (**Figure 4.2A**, left panel).

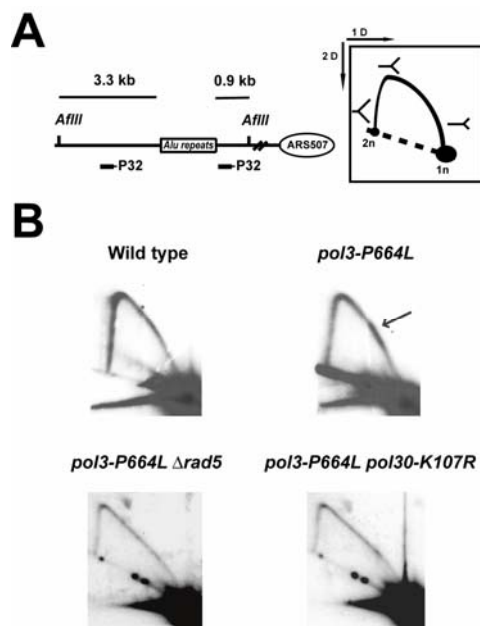


Figure 4.2. 2D gel analysis of replication fork progression in wild type and replication deficient strains with *Alu*-IRs.

A. Neutral/neutral 2-D electrophoresis was used to resolve unreplicated molecules and Y-like structures (Brewer and Fangman, 1987). Replication initiated at *ARS507*, proceeds from right to left through the region containing the repeats. Cleavage with *AflIII*, positions the inverted *Alu* repeats on the long shoulder of the Y-arc. The 4 kb *AflIII* digested *LYS2* fragment was used as a probe in Southern blot hybridization

B. Accumulation of the replication intermediates leads to the appearance of bulges on the replication arc indicated by arrow

However, we were able to detect replication slow down region coinciding with *Alu*-IRs in *pol3-P664L* mutants (Figure 4.2A, right panel).

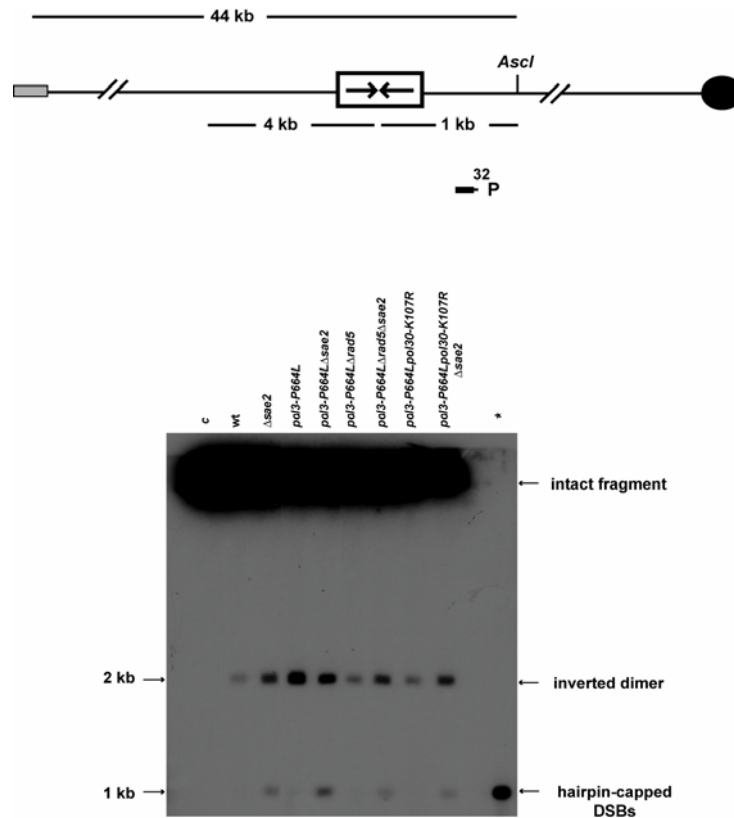


Figure 4.3. Analysis of hairpin-capped DSB formation at the site of *Alu*-IRs using in-plug digestions with rarely-cutting restriction enzymes. The positions of rarely cutting *AscI* restriction enzyme sites relative to *Alu*-IRs are shown. The centromere and the telomere are shown by solid black rectangle and solid grey circle, respectively. The chromosomal DNA from strains containing inverted *Alu* repeats was embedded in agarose plugs, digested with *AscI*, electrophoresed into the agarose gel to release the broken fragments. Southern blot hybridization was performed using *LYS2*-specific probes indicated by black solid rectangles. The DSB and the asterisk next to the gel panels correspond to broken and rearranged molecules, respectively. Horizontal arrows indicate position and the sizes of the bands corresponding to the DSB and dimer intermediates. The lane marked “C” depicts the *SalI-AscI* digestion of the genomic DNA prepared from the wild-type strain containing *Alu*-IRs.

Consistently, using CHEF gel electrophoresis and Southern blot hybridization, we found that the *pol3-P664L* mutants had enriched levels of DSB and dimer intermediates

(**data not shown**). Moreover, down-regulation of the leading and lagging strand components encoded by *POL1*, *POL2*, *POL3* and *POL30* also led to elevated levels of hairpin-capped break formation at the location of *Alu*-IRs (data not shown). Using in-plug digestion with rarely cutting restriction enzymes, we could detect DSB intermediates and the corresponding dimers on either sides of *Alu*-IRs (**Figure 4.3** and data not shown) indicating that the mechanism of cruciform resolution and hairpin-capped break formation also operates in replication defective mutants as in wild type strains, albeit with higher efficiency. It is important to note that in replication mutants, Mre11/Rad50/Xrs2 complex and Sae2 are not responsible for DSB generation but rather are involved in DSB processing (**Figure 4.3**) similar to the case when the replication machinery is not affected (Lobachev et al, 2002).

4.3.4. Chromosome fragility due to compromised replication requires some components of PRR

Defect in leading or lagging strand synthesis can generate long single stranded regions that provide optimal conditions for hairpin formation, which can in turn block the progression of the replication fork. It has been demonstrated that PRR is required to bypass lesions that are caused by UV or chemical agents in damage tolerance pathway (Klein, 2007; Smirnova and Klein, 2003; Torres-Ramos et al., 2002). To evaluate if the replication arrest imposed by lesion-like hairpin might also activate PRR ultimately leading to the fragility we disrupted the key players in the PRR pathway.

Disruption of *RAD5* gene, encoding for ubiquitin ligase and DNA helicase (Blastyak et al., 2007), in *pol3-P664L* mutants resulted in a significant decrease in the

level of GCRs: the rate of arm loss events in the double mutant was comparable to that observed in the wild type strain (**Table 4.2**). We also examined the effect of ubiquitin ligase deficient *rad5-I916A* and helicase dead *rad5(GAA)* mutations on replication-associated hairpin-capped DSB induction. The helicase dead allele and not the ubiquitin ligase defective mutation, similar to $\Delta rad5$, caused a reduction in the frequencies of the *Alu*-IR-induced GCRs in *pol3* mutants (**Table 4.2**). Similar decrease in GCR levels was detected in $\Delta rad6pol3-P664L$ and $\Delta rad18pol3-P664L$ strains. Consistent with the genetic data, we have found that breakage and replication arrest in $\Delta rad5pol3-P664L$ double mutants are also compromised (**Figure 4.2B**, left panel and **Figure 4.3**). However, interestingly, disruption of *UBC13* and *MMS2* did not alter the GCR levels in wild type or *pol3-P664L* strains (**Table 4.S1**, see Supplementary information provided this article online and data not shown). These data suggest that PRR is responsible for *Alu*-IR fragility only when replication machinery is defective. Specifically, the helicase activity of Rad5 is needed to unwind the nascent strand at the stall site to undergo template switching.

PCNA ubiquitination at K 164 or SUMOylation at K 127 is not required for the fragility (**Table 4.S1**), indicating that problems put forth by secondary structure are sensed differently by PRR from UV- or chemical-induced lesions. Moreover, we have found that the *pol30-K107R* mutation abolishes the hyper-GCRs observed in *pol3-P664L* strains (**Table 4.1**). PCNA is monoubiquitinated at K107 in strains defective in lagging strand synthesis that lack DNA Ligase I (Dr. A-K. Bielinsky, personal communication). As anticipated, there was a marked decrease in the accumulation of replication slow down zone and DSB intermediates in *pol30-K107Rpol3-P664L* double

mutants in comparison with *pol3-P664L* strains (**Figure 4.2B**, right panel and **Figure 4.3**).

These data demonstrate that PRR involvement and requirements for PCNA modification when replication fork encounters secondary structure are different from situations involving other types of DNA lesions and are similar to scenario when maturation of Okazaki fragments is affected.

4.3.5. Mrc1 and Tof1-Csm3 are required for the hyper fragility observed in *pol3-P664L* strains

Yeast activate checkpoint kinases in response to exposed single stranded DNA that might arise as a consequence of defective replisome (Branzei and Foiani, 2007a). We assessed if the palindrome fragility requires the role of Mec1 and the replication-pausing checkpoint proteins Tof1-Csm3 and Mrc1 that have been previously demonstrated to aid the replication machinery to evade DNA damage lesions, protein-bound barriers and assist in the global maintenance of fork integrity. We have found that while the kinase dead *mec1-21* mutation did not affect the fragility in *pol3-P664L* strains, the rate of arm loss in $\Delta csm3 pol3-P664L$, $\Delta tof1 pol3-P664L$ and $\Delta mrc1 pol3-P664L$ double mutants was approximately 10 fold lower than in *pol3-P664L* strain (**Table 4.1**, **Table 4.S1** and data not shown). The replication arrest in $\Delta tof1 pol3-P664L$ and $\Delta mrc1 pol3-P664L$ double mutants, however, was not different from that observed in *pol3* mutants (**Figure 4.S1**).

These data indicate that Tof1-Csm3 and Mrc1 proteins might signal for template switching or DSB generation when replication fork stalling is imposed by palindromes.

Hence, the requirements for replication-associated fragility at *Alu*-IRs are different from those that operate during replication arrest and DSB formation at the protein-bound barriers, DNA lesions or expanded CTG repeats (Cordon-Preciado et al., 2006; Freudenreich and Lahiri, 2004; Hodgson et al., 2007; Lahiri et al., 2004; Lopes et al., 2001; Tourriere et al., 2005).

4.4. Discussion

Hairpin- and cruciform-forming inverted repeats are widespread in eukaryotic genomes, including the human (Bureau et al., 1996; Consortium, 2001; Cox and Mirkin, 1997; Hancock, 2002; Karlin and Burge, 1996; Katti et al., 2001; Lisnic et al., 2005; Schroth and Ho, 1995; Stenger et al., 2001). Palindromic sequences are potent inducers of chromosomal fragility wherein the DSB ends are hairpin-capped (Lobachev et al., 2002). In addition, long *Alu*-quasipalindromes are replication stall sites on a plasmid in bacteria, yeast and primate cells (Voineagu et al., 2008). Although there is an evident connection to replication, the exact cause for DSB formation by unstable repeats in mitotically dividing eukaryotic cells is unknown. In this study we have unraveled a novel mechanism of hairpin-capped DSB formation that is dependent on replication. We have found that in yeast strains with compromised replication, checkpoint response and post replication repair machinery channel replication arrest at the site of inverted *Alus* to fragility.

4.4.1. Replication-dependent mechanism of *Alu*-IR fragility requires PRR pathway

Mutations in leading and lagging strand proteins are synergistic with *Alu*-IR-induced hairpin-capped DSBs and GCRs (**Figure 4.3**, **Table 4.1** and data not shown). This data correlates with a previous observation where chromosomal DSBs were also found at the site of two inverted Ty1 elements upon the down regulation of polymerase α (Lemoine et al., 2005). Impairment of lagging or leading strand synthesis might create long single-stranded regions where the formation of hairpins is thermodynamically favored. The scenario when hairpin is formed on the lagging strand template is shown as an example in the **Figure 4.4**. The lesion-like hairpin secondary structure can obstruct the progression of replication fork leading to stalling (**Figure 4.2A**, **right panel**). Following replication arrest, the components of PRR pathway are recruited for damage bypass. Rad5 ATPase activity can promote template switching of the stalled nascent strand to the newly synthesized leading sister strand. Absence of Rad5 helicase activity eliminates *Alu*-IR-induced replication arrest in *pol3* mutant strains (**Figure 4.2B**, left panel). It is quite possible that PRR aids in stabilization of fork stalling by hairpin-structure. Consistently, *pol30-K107R* mutation eradicates the profound replication arrest and DSB formation in replication mutants (**Figure 4.2B**, right panel, **Figure 4.3** and data not shown). PRR pathway might be the primary pathway of restoring the stalled replication fork at the hairpin structure. In the absence of template switch repair, the cell can recruit alternate mechanisms to enable the bypass of the secondary structure or restart of the replication fork.

The unwinding of the duplex during template switching can promote cruciform formation. Alternatively, synthesis on the lagging strand past the inverted repeats and reannealing back with the lagging strand template containing hairpin can also lead to

cruciform structure. The cruciform structure is subsequently targeted for processing, by a yet to identified nuclease to yield hairpin-capped DSBs (**Figure 4.3** and data not shown).

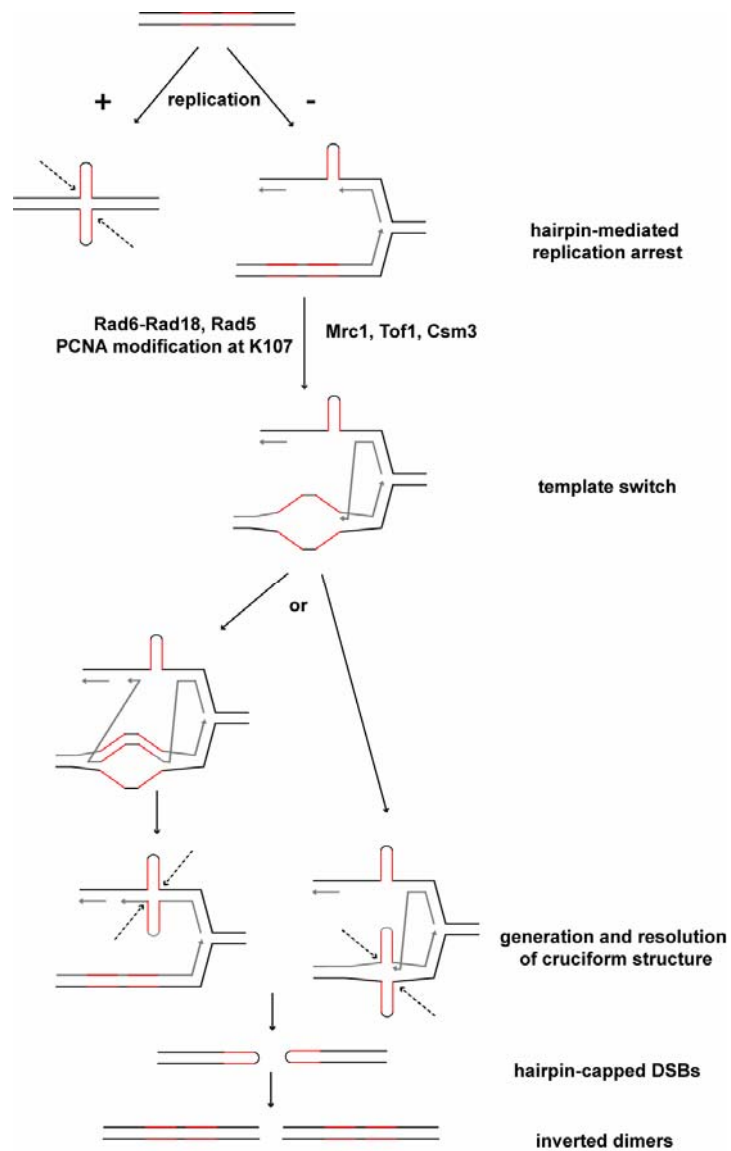


Figure 4.4. Replication-dependent and –independent mechanisms of cruciform-mediated DSBs. “+” and “-“ denote robust and compromised replication, respectively. Involvement of Checkpoint and PRR proteins in replication-dependent pathway are shown. Cruciforms are substrates for resolution leading to hairpin-capped DSBs. Dashed arrows indicate the nuclease attack of the cruciform structure. Unprocessed hairpin-capped DSBs are converted to linear inverted dimers following DNA replication.

4.4.2. *Alu*-IR-induced replication arrest activates Tof1-Csm3 and Mrc1 proteins

We have found that the checkpoint requirements under conditions of robust and damaged replisome are different. Mec1 is actively involved under conditions of uncompromised replication (**Table 4.S1**). However, mutations in Tof1, Csm3 or Mrc1 eliminate the hyper GCRs observed in replication deficient strains (**Table 4.2**). Mrc1 and Tof1 have been implicated in aiding the movement of replication fork across various kinds of lesions including tRNA genes, CTG/CAG triplet repeats, protein bound barriers and UV-induced damages. Moreover, these proteins have been shown to be required for normal replication progression in S-phase (Tourriere et al., 2005). Our data suggests that either the template switching or cruciform resolution in the PRR pathway requires function of replication-pausing checkpoint proteins Tof1-Csm3 and Mrc1.

Disruptions of the fork stabilizing proteins do not alter breakage in strains with uncompromised replication (data not shown). This clearly demonstrates that the checkpoint response to replication stalling by the secondary structure is unique. Plasmid-based studies indicated that Mrc1 and Tof1 counteract replication arrest by hairpins in yeast (Voineagu et al., 2008). Disruption of Mrc1 and Tof1 does not alter the replication progression in both wild type and *pol3-P664L* strains (data not shown and **Figure 4.S1**). The differential role of the checkpoint proteins can be explained by the structural differences in the chromatin organization or the disparity between plasmid and chromosomal replication. Furthermore, *Alu*-quasipalindromes hinder replication progression on the chromosome only in replication mutants unlike on a plasmid template where arrest is observed in wild type strains (**Figure 4.2A** and Voineagu et al., 2008).

Our data suggests that the arrested fork might be responsible for activating the checkpoint cascade that coordinates the generation of hairpin-capped DSBs.

In conclusion, this study identifies the PRR pathway as a novel mechanism for the generation of hairpin-capped DSBs in replication deficient strains. Our data indicates that there are two pathways, namely, replication-independent and replication-dependent mechanisms of inverted repeat fragility. We suggest that the requirements for replication-associated fragility at the inverted repeats are different from those that operate during replication arrest at the protein-bound barriers, DNA lesions and other secondary structure-forming repeats.

4.5. Materials and Methods

4.5.1. Strains and Plasmids

All strains in this study were isogenic to TP strains described in Narayanan et al., 2006. *RAD5*, *RAD6*, *RAD18*, *MRC1*, *TOF1* and *CSM3* genes were disrupted with the *kanMX* cassette. *pol30-K107R* allele was introduced using *dellito perfetto* technique. Nucleotide sequences of the primers used for integrations and disruptions are available upon request.

4.5.2. Genetic Techniques

The rates and 95% confidence intervals of the arm loss and recombination between *lys2* alleles were estimated in fluctuation tests using at least 14 independent cultures. The canavanine-containing media was made with a low concentration of adenine (5mg/L) to allow color detection; strains with an *ade2* mutations form red colonies in medium with low levels of adenine.

4.5.3. Structural Analysis of the Genome Rearrangements

For DSB detection, chromosomal DNA was embedded into agarose plugs using the CHEF Genomic DNA plug Kit from Bio-Rad (~6 x 10⁸ cells/1ml of plug). Gels were run in 0.5X TBE at 14°C using the Bio-Rad CHEF Mapper XA for 19.44 hours with switch times of 2.08s-42.91s. Southern blot hybridization was performed using 350 bp probe homologous to *DSF1* gene as previously described (Lobachev et al, 2002).

For in-plug digestion experiments, approximately 7µg of genomic DNA embedded in agarose was digested with *NotI* or *AscI* (Fermentas). Prior to loading in the gel, the restriction enzyme was removed and plugs were incubated for 30 min at room temperature in electrophoresis buffer (1X TBE). The chromosomal DNA was separated in 0.8 % agarose gel in 1XTBE at 0.8 V/cm for 20 hr for *AscI* digested samples and 36 hours for *NotI* digested samples. Southern blot hybridization was performed with 350 bp and 300 bp ³²P-labeled probes that are homologous to the left and right side of *Alu* insertion in *LYS2*. Nucleotide sequences of the oligonucleotides used for PCR amplification will be provided upon request.

4.5.4. 2D Analysis of Replication Fork Intermediates

Cells were arrested in G1 with α -factor and released synchronously into S-phase. At 40 min after release, chromosomal DNA was extracted and neutral/neutral 2-D analysis was carried out in according to ⁴⁶. The *AflIII*-digested DNA was separated in the first dimension on a 0.4% gel without ethidium bromide in 1X TBE buffer at 1V/cm for 38 hours in first dimension. The second dimension gel was run at 6 V/cm in 1X TBE buffer

containing 0.3 $\mu\text{g/ml}$ ethidium bromide for 13 hours. Southern blot hybridization using 4 kb *Afl*III-digested *LYS2* fragment was performed to highlight replication intermediates

4.6. Table 4.1. Effect of compromised replication on *Alu*-IR-induced fragility

Genetic background	Arm loss ($\times 10^{-5}$)	Fold increase over WT
Wild type	3.7 (2.9-4.2) ^a	1.0
<i>pol3-P664L</i>	119 (98-320)	32
<i>pol2-M644I</i>	98 (93-100)	26
Tet- <i>POL1</i> ^b	285 (239-305)	77
Tet- <i>POL2</i> ^b	119 (88-149)	32
Tet- <i>POL3</i> ^b	223 (196-307)	60
Tet- <i>POL30</i> ^b	175 (123-199)	47

^a Numbers in parentheses correspond to the 95% confidence interval.

^b Promoters of the indicated ORFs have been substituted for repressible *tetO* promoter.

Gene expression was controlled by growing yeast in YPD medium containing 2 $\mu\text{g/ml}$ doxycycline.

4.7. Table 4.2. PRR and checkpoint requirements for fragility in wild type and *pol3-P664L* strains

Genetic background	Arm loss (x 10 ⁻⁵)
Wild type	3.7 (2.9-4.2) ^a
<i>pol3-P664L</i>	119 (98-320)
<i>Δrad5</i>	4.3 (3.6-5.2)
<i>Δpad6</i>	3 (2.9-4.2)
<i>Δrad18</i>	3.9 (3.7-4.3)
<i>pol30-K107</i>	4.2 (3.4-5.2)
<i>Δmrc1</i>	3.9 (3.3-4.5)
<i>Δtof1</i>	3.1 (2.9-4.7)
<i>Δrad5pol3-P664L</i>	4.5 (3.9-5.1)
<i>rad5-I916Apol3-P664L</i>	125 (114-129)
<i>rad5(GAA)pol3-P664L</i>	6.5 (3.2-7.5)
<i>Δrad6pol3-P664L</i>	10.4 (8.6-11.3)
<i>Δrad18pol3-P664L</i>	8.2 (7.6-8.6)
<i>pol30-K107Rpol3-P664L</i>	4.5 (3.9-5.1)
<i>Δmrc1pol3-P664L</i>	11.9 (11.8-12)
<i>Δtof1pol3-P664L</i>	13.5 (10.9-16.7)

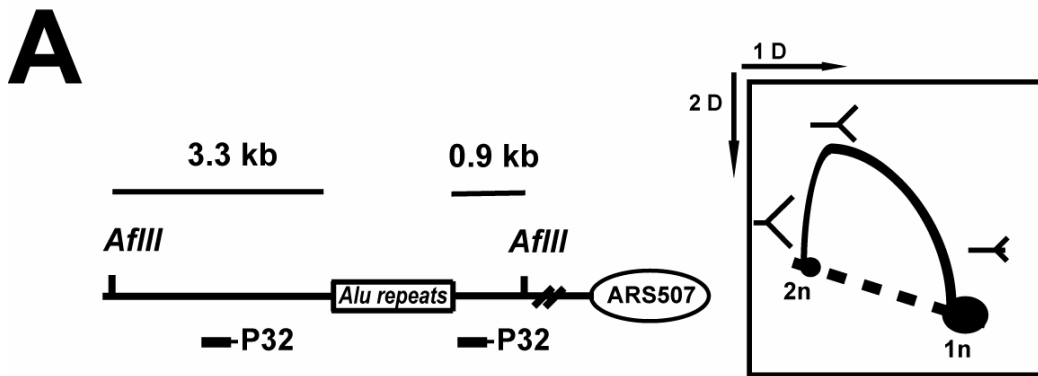
^a Numbers in parentheses correspond to the 95% confidence interval.

4.8. Supplementary Information

Table 4.S1. Length- and orientation-dependent size variations in expanded GAA and TTC repeat tracts.

Genetic background	Arm loss (x 10 ⁻⁵)
Wild type	3.7 (2.9-4.2) ^a
<i>pol3-P664L</i>	119 (98-320)
<i>mec1-21</i>	0.5 (0.2-0.7)
<i>Δubc13pol3-P664L</i>	132 (124-148)
<i>Δmms2pol3-P664L</i>	127 (115-132)
<i>pol30-K164Rpol3-P664L</i>	118 (113-127)
<i>pol30-K127Rpol3-P664L</i>	125 (104-145)
<i>mec1-21pol3-P664L</i>	104 (95-115)

^a Numbers in parentheses correspond to the 95% confidence interval.



B

pol3-P664L Δcsm3 *pol3-P664L Δtof1*

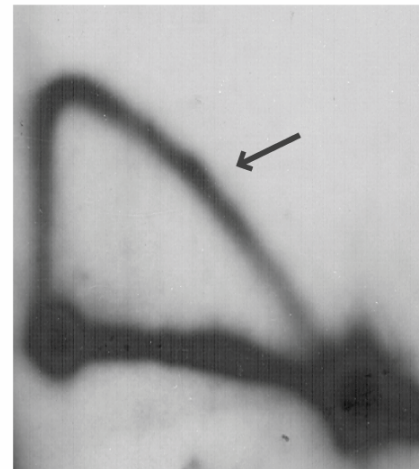
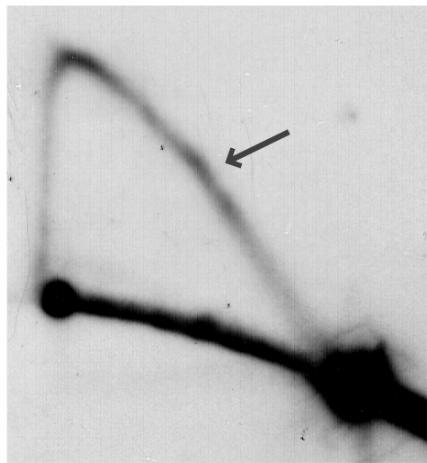


Figure 4.S1. 2D gel analysis of replication fork progression in *pol3-P664LΔmrc1* and *pol3-P664LΔcsm3* mutant strains. 2D analysis was performed as described in Figure 4.2. The replication stall zone across the *Alu*-IRs is indicated by arrow.

4.9. References

Andersen, P.L., Xu, F., and Xiao, W. (2008). Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. *Cell Res* 18, 162-173.

Azvolinsky, A., Dunaway, S., Torres, J.Z., Bessler, J.B., and Zakian, V.A. (2006). The *S. cerevisiae* Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. *Genes Dev* 20, 3104-3116.

Barbour, L., and Xiao, W. (2003). Regulation of alternative replication bypass pathways at stalled replication forks and its effects on genome stability: a yeast model. *Mutat Res* 532, 137-155.

Blastyak, A., Pinter, L., Unk, I., Prakash, L., Prakash, S., and Haracska, L. (2007). Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. *Mol Cell* 28, 167-175.

Branzei, D., and Foiani, M. (2007a). Interplay of replication checkpoints and repair proteins at stalled replication forks. *DNA Repair (Amst)* 6, 994-1003.

Branzei, D., and Foiani, M. (2007b). RecQ helicases queuing with Srs2 to disrupt Rad51 filaments and suppress recombination. *Genes Dev* 21, 3019-3026.

Branzei, D., Seki, M., and Enomoto, T. (2004). Rad18/Rad5/Mms2-mediated polyubiquitination of PCNA is implicated in replication completion during replication stress. *Genes Cells* 9, 1031-1042.

Branzei, D., Sollier, J., Liberi, G., Zhao, X., Maeda, D., Seki, M., Enomoto, T., Ohta, K., and Foiani, M. (2006). Ubc9- and mms21-mediated sumoylation counteracts recombinogenic events at damaged replication forks. *Cell* 127, 509-522.

Bureau, T.E., Ronald, P.C., and Wessler, S.R. (1996). A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. *Proc Natl Acad Sci U S A* 93, 8524-8529.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. (2005). Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes Dev* 19, 1905-1919.

Cha, R.S., and Kleckner, N. (2002). ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science* 297, 602-606.

- Chen, S., Davies, A.A., Sagan, D., and Ulrich, H.D. (2005). The RING finger ATPase Rad5p of *Saccharomyces cerevisiae* contributes to DNA double-strand break repair in a ubiquitin-independent manner. *Nucleic Acids Res* 33, 5878-5886.
- Consortium, I.H.G.S. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Cordon-Preciado, V., Ufano, S., and Bueno, A. (2006). Limiting amounts of budding yeast Rad53 S-phase checkpoint activity results in increased resistance to DNA alkylation damage. *Nucleic Acids Res* 34, 5852-5862.
- Cox, R., and Mirkin, S.M. (1997). Characteristic enrichment of DNA repeats in different genomes. *Proc Natl Acad Sci U S A* 94, 5237-5242.
- Deshpande, A.M., and Newlon, C.S. (1996). DNA replication fork pause sites dependent on transcription. *Science* 272, 1030-1033.
- Freudenreich, C.H., and Lahiri, M. (2004). Structure-forming CAG/CTG repeat sequences are sensitive to breakage in the absence of Mrc1 checkpoint function and S-phase checkpoint signaling: implications for trinucleotide repeat expansion diseases. *Cell Cycle* 3, 1370-1374.
- Gangavarapu, V., Haracska, L., Unk, I., Johnson, R.E., Prakash, S., and Prakash, L. (2006). Mms2-Ubc13-dependent and -independent roles of Rad5 ubiquitin ligase in postreplication repair and translesion DNA synthesis in *Saccharomyces cerevisiae*. *Mol Cell Biol* 26, 7783-7790.
- Greenfeder, S.A., and Newlon, C.S. (1992). Replication forks pause at yeast centromeres. *Mol Cell Biol* 12, 4056-4066.
- Hancock, J.M. (2002). Genome size and the accumulation of simple sequence repeats: implications of new data from genome sequencing projects. *Genetica* 115, 93-103.
- Haracska, L., Torres-Ramos, C.A., Johnson, R.E., Prakash, S., and Prakash, L. (2004). Opposing effects of ubiquitin conjugation and SUMO modification of PCNA on replicational bypass of DNA lesions in *Saccharomyces cerevisiae*. *Mol Cell Biol* 24, 4267-4274.
- Hodgson, B., Calzada, A., and Labib, K. (2007). Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase. *Molecular biology of the cell* 18, 3894-3902.
- Hoegge, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419, 135-141.

- Ira, G., Malkova, A., Liberi, G., Foiani, M., and Haber, J.E. (2003). Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in yeast. *Cell* *115*, 401-411.
- Karlin, S., and Burge, C. (1996). Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proc Natl Acad Sci U S A* *93*, 1560-1565.
- Katti, M.V., Ranjekar, P.K., and Gupta, V.S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol Biol Evol* *18*, 1161-1167.
- Klein, H.L. (2007). Reversal of fortune: Rad5 to the rescue. *Mol Cell* *28*, 181-183.
- Kobayashi, T. (2003). The replication fork barrier site forms a unique structure with Fob1p and inhibits the replication fork. *Mol Cell Biol* *23*, 9178-9188.
- Krasilnikova, M.M., and Mirkin, S.M. (2004). Replication stalling at Friedreich's ataxia (GAA)_n repeats in vivo. *Mol Cell Biol* *24*, 2286-2295.
- Labib, K., and Hodgson, B. (2007). Replication fork barriers: pausing for a break or stalling for time? *EMBO Rep* *8*, 346-353.
- Lahiri, M., Gustafson, T.L., Majors, E.R., and Freudenreich, C.H. (2004). Expanded CAG repeats activate the DNA damage checkpoint pathway. *Mol Cell* *15*, 287-293.
- Lemoine, F.J., Degtyareva, N.P., Lobachev, K., and Petes, T.D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* *120*, 587-598.
- Lisnic, B., Svetec, I.K., Saric, H., Nikolic, I., and Zgaga, Z. (2005). Palindrome content of the yeast *Saccharomyces cerevisiae* genome. *Curr Genet* *47*, 289-297.
- Lobachev, K.S., Gordenin, D.A., Resnick, M.A., Stenger, J.E., Gordenin, D., Darden, T.A., Jurka, J., Kozyreva, O.G., Shor, B.M., Tran, H.T., *et al.* (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* *108*, 183-193.
- Lopes, M., Cotta-Ramusino, C., Pelliccioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C.S., and Foiani, M. (2001). The DNA replication checkpoint response stabilizes stalled replication forks. *Nature* *412*, 557-561.
- Mohanty, B.K., Bairwa, N.K., and Bastia, D. (2006). The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* *103*, 897-902.
- Narayanan, V., Mieczkowski, P.A., Kim, H.M., Petes, T.D., and Lobachev, K.S. (2006). The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. *Cell* *125*, 1283-1296.

- Papouli, E., Chen, S., Davies, A.A., Huttner, D., Krejci, L., Sung, P., and Ulrich, H.D. (2005). Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol Cell* 19, 123-133.
- Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E., and Kunkel, T.A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317, 127-130.
- Schroth, G.P., and Ho, P.S. (1995). Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA. *Nucleic Acids Res* 23, 1977-1983.
- Smirnova, M., and Klein, H.L. (2003). Role of the error-free damage bypass postreplication repair pathway in the maintenance of genomic stability. *Mutat Res* 532, 117-135.
- Stenger, J.E., Lobachev, K.S., Gordenin, D., Darden, T.A., Jurka, J., and Resnick, M.A. (2001). Biased distribution of inverted and direct Alus in the human genome: implications for insertion, exclusion, and genome stability. *Genome Res* 11, 12-27.
- Sung, P., and Klein, H. (2006). Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* 7, 739-750.
- Torres-Ramos, C.A., Prakash, S., and Prakash, L. (2002). Requirement of RAD5 and MMS2 for postreplication repair of UV-damaged DNA in *Saccharomyces cerevisiae*. *Mol Cell Biol* 22, 2419-2426.
- Tourriere, H., Versini, G., Cordon-Preciado, V., Alabert, C., and Pasero, P. (2005). Mrc1 and Tof1 promote replication fork progression and recovery independently of Rad53. *Mol Cell* 19, 699-706.
- Ulrich, H.D., and Jentsch, S. (2000). Two RING finger proteins mediate cooperation between ubiquitin-conjugating enzymes in DNA repair. *EMBO J* 19, 3388-3397.
- Voineagu, I, Narayanan, V, Lobachev, K.S. and Mirkin, S (2008). Replication stalling at unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins. *PNAS*, in press.
- Wu, L., and Hickson, I.D. (2003). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426, 870-874.
- Zhang, H., and Freudenreich, C.H. (2007). An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae*. *Mol Cell* 27, 367-379.

4.10. Contribution to the publication

The presented study is in preparation. All the experimental work was done by the author.

The chapter was also written by the author.

CHAPTER 5

Estimation of mutation rates

The rates and 95% confidence intervals of the arm loss and gene amplification were estimated in fluctuation tests using at least 14 independent cultures. The cells from each culture were suspended in a defined volume of water. Yeast cells were then plated on selective medium: canavanine-containing media was made with a low concentration of adenine (5mg/L) to allow color detection and Copper plates were prepared from SD complete media with final concentration of 700 μ M CuSO₄ solution. Cells from every culture were also plated on YPD medium with appropriate dilutions. To select for amplification events, we replica plated the Cu^R colonies to freshly-made 2mM formaldehyde plates. After two days of incubation, these plates were replica plated again to formaldehyde-containing medium to verify the growth. The colonies formed by the mutant cells were counted on the selective medium and the total number of viable cells in culture was determined based on the colony count on YPD medium.

Mutation rate (μ) for each culture was calculated based on the mutant frequency ($f=m/N$)

Where, $m=C_{sel} \cdot V_{tot} \cdot D_{sel} / V_{sel}$, total number of mutants in culture

$N=C_{com} \cdot V_{tot} \cdot D_{com} / V_{com}$, total number of cells in culture

V_{com}- volume plated from YPD;

V_{sel}- volume plated from selective medium;

V_{tot}-total volume of culture;

C_{sel}-number of colonies on all selective plates;

C_{com}-number of colonies on all YPD plates;

D_{sel}-dilution factor, selective plates;

D_{com}-dilution factor, YPD plates

Mutation rate (μ) was calculated from the transcendental equation $\mu = f / \ln(N\mu)$, derived in (Drake 1991). The median rate is used as a measure of mutability. The 95% confidence intervals for a median rate were determined as Dixon and Massey, 1969.

Alternative rate $\mu(\text{small})$ calculation was employed for low rates, especially when majority of cultures have no mutants.

$$\mu(\text{small}) = \sum_{i-k} C_{\text{sel}} / \sum_{i-k} C_{\text{tot}}$$

where, $C_{\text{tot}} = C_{\text{com}} * V_{\text{sel}} * D_{\text{com}} / (V_{\text{com}} * D_{\text{sel}})$,
total number of cells plated on all selective plates in a given culture.

References

Drake, J.W., *PNAS* 1991 Aug 15;88(16):7160-4.

Dixon, W. J., and F. J. Massey, Jr., 1969 *Introduction to Statistical Analysis*, pg. 349. McGraw-Hill, New York.

CHAPTER 6

CONCLUSION

Palindromic sequences, closely spaced inverted repeats and quasipalindromes can be found in a variety of organisms such as viruses, bacteria and eukaryotes. In many of these genomes these secondary structure forming sequences have a functional role. Very short perfect palindromes (less than 8 bp) are known to be recognition sites for type II restriction-modification systems that play a critical role in ecology and evolution of bacterial genomes (Gelfand and Koonin, 1997). In prokaryotes, IRs are often found in the binding sites for regulatory proteins. IRs also serve as protein binding sites in eukaryotes. In yeast, sequences containing palindromes with 1 bp spacer coordinate cellular response to unfolded or misfolded proteins in the endoplasmic reticulum (Mori et al., 1998). In humans, the promoter region of the *erbB-2* gene contains two palindromic sequences that bind a heterodimeric protein complex (Chen and Gill, 1996). In mouse B lymphoma cells, palindromic sequences were identified as break points during class switch recombination (Tashiro et al., 2001). The hairpin-forming ability of inverted repeats is exploited in genetic elements involved into basic cellular processes. These include transcription terminators, attenuators and in the replication origin sites of both prokaryotes and eukaryotes.

Despite their functional versatility, long inverted repeats were shown to be very unstable in different organisms, from bacteria to mammalian cells. There is accumulating evidence suggesting that fragility at the location of secondary structure is an important factor that compromises the genome integrity (discussed in introduction). Using yeast model system, it has been demonstrated that inverted repeats are potent inducers of gross

chromosomal rearrangements observed in cancer cells (Freudenreich and Lahiri, 2004; Lahiri et al., 2004; Lemoine et al., 2005; Narayanan et al., 2006). Palindrome-mediated hairpin-capped DSBs can result in chromosomal arm loss, extrachromosomal and intrachromosomal gene amplification events (Narayanan et al., 2006). Both arm loss and intrachromosomal gene amplification were frequently accompanied by nonreciprocal translocations. This spectrum of rearrangements was governed by the applied selection, the nature of the break, and the chromosomal location of the amplified gene relative to the site of the hairpin-capped DSB. Interestingly, the majority of the breakpoints of the resulting GCRs did not co-localize with the initial hairpin-capped break site and the sequence that triggered GCR was still present at the center of the duplicated or amplified regions. This observation might explain why some recurrent aberrations in cancer cells do not have structure prone sequence motifs at the rearrangement breakpoints. The yeast experiments indicated that it was possible to model events leading to cancer associated abnormalities, identify rules dictating specific patterns of rearrangements, predict the location of the causative secondary structure-forming sequences and uncover susceptible phenotypes.

The mechanisms of palindrome-mediated instability defined in yeast makes specific predictions about the genetic parameters and structural organization of the gross chromosomal aberrations. If the mechanism is conserved among eukaryotes, the predictions can be used in cancer genomic studies to unveil the origin and nature of the tumorigenic rearrangements. Moreover, analyses of genetic composition can aid in evaluating the predisposition of individuals to chromosomal instability and rearrangements.

The following are the conclusions summarized from the graduate work:

1. Perfectly homologous *Alu*-quasipalindromes are potent inducers of chromosomal fragility and gross chromosomal rearrangements including deletions, translocations and gene amplification.
2. Generation of intrachromosomal amplicons by hairpin-capped breaks is dependent on homologous recombination proteins and does not require nonhomologous end-joining machinery.
3. DSB induction at a large distance from inverted repeats on the chromosome leads to formation of large dicentric inverted dimers via interaction between inverted Tys.
4. Homologous and homeologous inverted *Alu* repeats are strong pause sites for the DNA replication on a plasmid in three model systems (*E.coli*, *S. cerevisiae* and COS monkey cells)
5. There are two distinct mechanisms of inverted repeat induced fragility, namely, replication dependent and replication independent pathways. The requirements of replication-induced fragility at the location of secondary structure are different from that of protein-bound barriers or DNA lesions.

6.1. References

- Chen, Y., and Gill, G.N. (1996). A heterodimeric nuclear protein complex binds two palindromic sequences in the proximal enhancer of the human erbB-2 gene. *J Biol Chem* *271*, 5183-5188.
- Cunningham, L.A., Cote, A.G., Cam-Ozdemir, C., and Lewis, S.M. (2003). Rapid, stabilizing palindrome rearrangements in somatic cells by the center-break mechanism. *Mol Cell Biol* *23*, 8740-8750.
- de los Santos, T., Hunter, N., Lee, C., Larkin, B., Loidl, J., and Hollingsworth, N.M. (2003). The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. *Genetics* *164*, 81-94.
- Freudenreich, C.H., and Lahiri, M. (2004). Structure-forming CAG/CTG repeat sequences are sensitive to breakage in the absence of Mrc1 checkpoint function and S-phase checkpoint signaling: implications for trinucleotide repeat expansion diseases. *Cell Cycle* *3*, 1370-1374.
- Gelfand, M.S., and Koonin, E.V. (1997). Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes. *Nucleic acids research* *25*, 2430-2439.
- Gordenin, D.A., Lobachev, K.S., Degtyareva, N.P., Malkova, A.L., Perkins, E., and Resnick, M.A. (1993). Inverted DNA repeats: a source of eukaryotic genomic instability. *Mol Cell Biol* *13*, 5315-5322.
- Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P.A., Lilley, D.M., and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell* *103*, 1133-1142.
- Lahiri, M., Gustafson, T.L., Majors, E.R., and Freudenreich, C.H. (2004). Expanded CAG repeats activate the DNA damage checkpoint pathway. *Molecular cell* *15*, 287-293.
- Lemoine, F.J., Degtyareva, N.P., Lobachev, K., and Petes, T.D. (2005). Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* *120*, 587-598.
- Lobachev, K.S., Gordenin, D.A., and Resnick, M.A. (2002). The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* *108*, 183-193.

Mori, K., Ogawa, N., Kawahara, T., Yanagi, H., and Yura, T. (1998). Palindrome with spacer of one nucleotide is characteristic of the cis-acting unfolded protein response element in *Saccharomyces cerevisiae*. *J Biol Chem* 273, 9912-9920.

Narayanan, V., Mieczkowski, P.A., Kim, H.M., Petes, T.D., and Lobachev, K.S. (2006). The pattern of gene amplification is determined by the chromosomal location of hairpin-capped breaks. *Cell* 125, 1283-1296.

Tashiro, J., Kinoshita, K., and Honjo, T. (2001). Palindromic but not G-rich sequences are targets of class switch recombination. *Int Immunol* 13, 495-505.

CHAPTER 7

LIST OF PUBLICATIONS

1. Narayanan V, Mieczkowski P., Kim H-M, Petes T.D., Lobachev K.S “The pattern of gene amplification is determined by chromosomal location of hairpin-capped breaks”, *Cell* 125:1283-1296, 2006.
2. VanHulle K., Lemoine F. J., Narayanan V., Downing B., Hull K., McCullough C., Bellinger M., Lobachev K., Petes T. D., Malkova A. “Inverted DNA repeats channel repair of distant double-strand breaks into chromatid fusions and chromosomal rearrangements”, *Molecular and Cellular Biology*, 27:2601-2614, 2007 .
3. Narayanan V., and Lobachev K.S. “Intrachromosomal gene amplification triggered by hairpin-capped breaks requires homologous recombination and is independent of non-homologous end joining”, *Cell Cycle*,6:1814-8., 2007
4. Lobachev K.S., Rattray A., Narayanan V. “Hairpin- and cruciform-mediated chromosome breakage: causes and consequences in eukaryotic cells”, *Frontiers in Bioscience*, 2007, 12:4208-4220, 2007
5. Voineagu I., Narayanan V., Lobachev K.S., Mirkin S.M., Replication stalling at unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins”, *PNAS*, in press.
6. Kim H-M., Narayanan V., Mieczkowski P, Petes T.D, Krasilnikova M.M., Mirkin S.M., Lobachev K.S., “Chromosome fragility at the expanded

GAA/TTC tracts in yeast depends on repeat orientation and requires the mismatch repair system”, *submitted*.