

THE UNIVERSITY OF CALIFORNIA, SAN DIEGO

Epigenetic regulation of centromere and neocentromere activity
*in *Drosophila melanogaster**

A dissertation submitted in partial satisfaction of the requirements for the

Doctor of Philosophy in Biology

by

Keith Andrew Maggert

Committee in charge:

Gary Karpen, chair

Lawrence Goldstein

Dan Lindsley

John Newport

Barbara Wakimoto

2000

Copyright
Keith Andrew Maggert, 2000.
All rights reserved.

The dissertation of Keith Andrew Maggert is approved, and it
is acceptable in quality and form:

2000

To my grandmother, Hermina M. Shane,

and my friend, David F. Hessler,

for their guiding memories

and to my wife, Suzanne L. Eckert,

for her active exaltation

".. chromosomes are occasionally broken in such a way that the fracture takes its path through the kinetochore [centromere]. [Centromere] fragments thus produced evidently may be functional at times, for the resulting chromosome fragments continue to participate regularly in the mitotic process."

— Franz Schrader, 1939.

Table of Contents

Signature page.....	iii
Dedication.....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Figures.....	ix
List of Tables.....	xi
List of Formulæ.....	xii
Preface.....	xiii
Curriculum vitæ.....	xiv
Acknowledgements.....	xviii
Abstract of the Dissertation.....	xxiv
CHAPTER ONE	1
Introduction	
The Role of the Centromere.....	3
The Structure of the Centromere.....	9
Epigenetics: Descriptions and Explanations.....	17
Preludes To This Work.....	23
Prevailing Models for Centromere Activation: tests and predictions.....	25
CHAPTER TWO	31
Minichromosome Derivatives of <i>Drosophila</i>	
Structure of <i>Dp8-23</i>	33
Structure of <i>Dpγ238</i>	38
Structure of <i>Tγ1337</i>	41
Summary.....	47
CHAPTER THREE	49
The γ-Screen	
Logic of the Screen.....	50
Mechanics of the Screen.....	51

Results of the Screen.....	55
Frequencies of Offspring	58

CHAPTER FOUR **60**

Structural Analyses of Minichromosome Derivatives

Methodology for Determining Derivative Structure.....	62
Structures of the <i>Dpγ238 yellow</i> ⁺ Derivatives	63
Structures of the <i>Dp8-23 yellow</i> ⁺ Derivatives	64
Comparison of Class 3A Euchromatic Breakpoints of <i>Dpγ238</i> and <i>Dp8-23</i>	66
Structures of <i>Tγ1337 rosy</i> ⁻ Derivatives.....	69
Structures of <i>Dp8-23</i> and <i>Tγ1337 rosy</i> ⁺ Derivatives.....	73
Structures of <i>Dpγ238 rosy</i> ⁺ Derivatives.....	74
Analysis of Structurally Acentric Chromosomes	78
Analysis of Breaks within the <i>Dpγ238</i> Test Fragment	81
Summary.....	84

CHAPTER FIVE **86**

Genetic Analyses of Minichromosome Derivatives

Transmission Tests: methods and results.....	87
Brooding Assays: methods and results	91
Acentrism and Size-Dependency: a treatment of chromosome loss.....	93
Selection for Stability	95
Summary.....	99

CHAPTER SIX **100**

Synthesis and Discussion

Centromeres and Neocentromeres: Sufficiency of Mark.....	102
Centromeres and Neocentromeres: Sufficiency of Size	105
Commonalities Among Centromeres.....	108
Activating and Deactivating Centromeres	110
Conclusion	118

APPENDICES **120**

Appendix A: Materials and Methodology

Transmission tests	121
Brooding Assays	121
Chromosome <i>I</i> Background	122
Plasmid DNA Preparation.....	122
Adult Fly DNA Preparation	123
Genomic DNA Inserts.....	124
Pulse Field Gel Electrophoresis and Molecular Weight Determination.....	124
Southern Hybridization.....	125
Statistical Analyses	126
Polymerase Chain Reaction	126

Appendix B: Data

List of Clones	128
Primers and Oligonucleotides	128
Maps of Chromosomes	130
Map of yellow	131
Transmission Data	132
Brooding data.....	134
Selection Data	135
Dpg238 Breakpoint.....	136
<i>Ty1337</i> Breakpoints	136

Bibliography

List of Figures

Figure 1: The Centromere	3
Figure 2: Fate of Acentric Chromosome	7
Figure 3: Structure of the <i>Saccharomyces cerevisiae</i> Centromere	10
Figure 4: Structure of the <i>Schizosaccharomyces pombe</i> Centromere.....	12
Figure 5: Epigenetic Regulation in the <i>Schizosaccharomyces pombe</i> Centromere.....	13
Figure 6: Structure of a <i>Drosophila melanogaster</i> Centromere	15
Figure 7: Structure of the Mammalian Centromere.....	16
Figure 8: Energetics of Epigenetics	20
Figure 9: Two Models for Establishing Centromere Identity	27
Figure 10: Implications of the Models for Centromere Maintenance.....	29
Figure 11: The Ontogeny of <i>Dp8-23</i> , <i>Dpγ238</i> , and <i>Tγ1337</i>	34
Figure 12: <i>Dp8-23</i> Structure and Landmarks.....	36
Figure 13: <i>Dpγ238</i> Structure and Landmarks.....	40
Figure 14: <i>Tγ1337</i> Cytology and Structure.....	42
Figure 15: <i>Tγ1337</i> Structure and Landmarks.....	44
Figure 16: Segregation Products of <i>Tγ1337</i>	46
Figure 17: Screen to Recover γ^- -Induced Derivatives of <i>Dpγ238</i> and <i>Dp8-23</i>	52
Figure 18: Screen to Recover γ^- -Induced Derivatives of <i>Tγ1337</i>	53
Figure 19: Scheme to Recover and Balance <i>Tγ1337</i> , γ^- Derivatives.....	54
Figure 20: Classes of Chromosomes Recovered from γ^- -Screens	56
Figure 21: Distribution of Class 3A (<i>yellow</i> ⁺) Breaks of <i>Dpγ238</i>	63
Figure 22: Normal Distribution of Breakpoints of <i>Dpγ238</i> Class 3A Derivatives.....	64
Figure 23: Distribution of Class 3A (<i>yellow</i> ⁺) Breaks of <i>Dp8-23</i>	65
Figure 24: Normal Distribution of Breakpoints of <i>Dp8-23</i> Class 3A Derivatives.....	66
Figure 25: Comparison of <i>Dpγ238</i> and <i>Dp8-23</i> Class 3A Derivative Breakpoints	67
Figure 26: Distribution of Class 4 (<i>rosy</i> ⁺) Breaks of <i>Dpγ238</i>	77
Figure 27: Mutability of Chromatin Types.....	77
Figure 28: Types of Divisions.....	88
Figure 29: Brooding Analysis of Derivative 26C Loss	93
Figure 30: Test for <i>in vialo</i> Evolution of 26C Stability	97
Figure 31: Energetics of Centromere Identity	111

Figure 32: Energetic Explanation of Centromere Spreading	116
Figure 34: Structural Maps of <i>Dp8-23</i> , <i>Dpγ238</i> , and <i>Tγ1337</i>	131
Figure 35: Map of <i>yellow</i>	131

List of Tables

Table 1: Viability of $T\gamma 1337$ Segregation Products.....	47
Table 2: Chromosome Derivatives: Numbers.....	59
Table 3: Frequencies of euchromatic breaks from $Dp\gamma 238$, $Dp8-23$, and $T\gamma 1337$	70
Table 4: Chromosome Derivatives: Numbers and Classes.....	85
Table 5: Comparison of Coefficients of Variance from Transmission Tests	99

List of Formulae

<i>Eqn (1)</i>	$\mathbf{T} = \mathbf{f}^{\mathbf{s}}$	90
<i>Eqn (2)</i>	$\mathbf{T} = \mathbf{f}_{\text{mt}}^{17} \cdot \mathbf{f}_{\text{me}}^2$	90
<i>Eqn (3)</i>	$\mathbf{f} = \mathbf{h} \cdot \ln \mathbf{L}$	94

Preface

This dissertation encompasses the thesis research that I have completed between March 1996 and April 2000, under the tutelage of Dr. Gary H. Karpen at the Salk Institute for Biological Studies. It is primarily concerned with an investigation of a neocentromere in *Drosophila melanogaster*. The research has involved an investigation of the role of chromosome architecture as causative or permissive for neocentromere activation, the equitability of neocentromere and centromere, and the structure of centromeric chromatin. I will present molecular, genetic, and cytological techniques that I have used in order to investigate these problems. Additionally, I will present a theoretical treatment of chromosome stability and centromere identity, heterochromatin function, and evolutionary considerations of centromere identity that extends from my work, as well as that of others in the lab.

Curriculum vitae

Personal

Date of birth: March 10, 1970

Place of birth: San Diego, California, United States of America

Education

1992 – 2000 University of California, San Diego, Biology Ph.D.

1988 – 1992 University of California, Santa Cruz, Biochemistry and Molecular Biology B.Sc.

Research Experience

1990 – 1992 Structure and regulation of the vacuolar H⁺-pumping ATP hydrolase in *Saccharomyces cerevisiae*; Dr. Lincoln Taiz, University of California, Santa Cruz

1992 – 1996 Genetic regulation of mesoderm determination and differentiation in *Drosophila melanogaster*; Dr. Michael Levine, University of California, San Diego

1996 – 2000 Epigenetic regulation of centromere and neocentromere activity in *Drosophila melanogaster*; Dr. Gary Karpen, The Salk Institute for Biological Studies and University of California, San Diego

Teaching and Outreach

High School Science Fair Judge — Greater San Diego Science and Engineering Fair, 1989 - 2000

Introduction to Biology — Teaching Assistant, University of California, Santa Cruz,

1990 - 1991

Introduction to Biology — Teaching Assistant, University of California, Santa Cruz, 1991 - 1992

Biochemical Techniques Lab — Teaching Assistant, University of California, San Diego, 1993

Genetics — Teaching Assistant, University of California, San Diego, 1994

Metabolic Biochemistry — Teaching Assistant, University of California, San Diego, 1995

Embryology Lab — Teaching Assistant, University of California, San Diego, 1995

Embryology Lab — Teaching Assistant, Woods Hole Oceanographic Institute, 1995

Independent Research Instruction — The Salk Institute for Biological Studies/University of California, San Diego, 1999 - 2000

High School Science Day — Association of Women in Science, 1998

Students' Science Day at the Salk Institute — The Salk Institute for Biological Studies, 1999

The Salk Institute Mobile Science Lab — The Salk Institute for Biological Studies, 1999 - 2000

Memberships and Awards

College Honors — University of California, Santa Cruz, 1992

Honors in the Major — University of California, Santa Cruz, 1992

Lucille P Markey Charitable Trust Fellowship — University of California, San Diego, 1992 - 1998

Excellence in Teaching Award, Department of Biology — University of California, San Diego, 1993

Excellence in Teaching Award, Department of Biology — University of California, San Diego, 1995

Basic Biochemical and Biological Mechanisms in Cancer — University of California, San Diego, 1994 – 1998

Chapman Charitable Trust Fellowship — The Salk Institute for Biological Studies, 1999 – 2000

Society for Research Fellows Presentation Award — The Salk Institute for Biological Studies, 1999

Publications

Taiz L, Nelson H, Maggert K, Morgan L, Yatabe B, Taiz SL, Rubinstein B, Nelson N. Functional analysis of conserved cysteine residues in the catalytic subunit of the yeast vacuolar H⁺-ATPase. *Biochimica et Biophysica Acta*, 1194(2):329-34, 1994.

Ip YT, Maggert K, Levine M. Uncoupling gastrulation and mesoderm differentiation in the *Drosophila* embryo. *EMBO Journal*, 13(24):5826-5834, 1994.

Maggert K, Levine M, Frasch M. The somatic-visceral subdivision of the embryonic mesoderm is initiated by dorsal gradient thresholds in *Drosophila*. *Development*, 121:2107-2116, 1995.

Dobie KW, Hari KL, Maggert KA, Karpen GH. Centromere proteins and chromosome inheritance: a complex affair. *Current Opinion in Genetics & Development*, 9: 206-217, 1999.

Maggert KA, Karpen GH. Acquisition and metastability of centromere identity and function: sequence analysis of a human neocentromere. *Genome Research*, 10: 725-728, 2000.

Maggert KA, Karpen GH. The Activation of a Neocentromere in *Drosophila* Requires Proximity to an Endogenous Centromere. *Nature Genetics*, submitted, 2000.

Acknowledgements

I have endeavored to thank you all along the way, but if I haven't, let me take the opportunity now.

I acknowledge my wife, Suzanne L. Eckert. The lives of graduate students are not easy, and she has tempered mine. She has kept me scientifically and personally honest, kept my excitement high, and forced me to grow in areas that I didn't want. Her contributions were professional (statistical understanding and clarity of prose) and personal (morale, support, and release). Stated simply, this endeavor would have failed without her.

I thank my parents, who have found it important not just to dismiss my quirks, but to see them for what they are, and support them. Being a scientist has been a goal my entire life, due to their example. I am honored to follow in their footsteps. I do not dedicate this work to them, because it seems to little for all that they've given me.

My grandmother was my most unashamed proponent, and I wish that she had made it to see me attain in reality what I always was in her mind's eye.

The entire Huckell clan has offered support that has been inspiring. For years I have been their intermittent pseudochild/sibling, an opportunity for which I am

indebted. I cannot begin to explain how grateful I am for their support and love.

You'd be so lucky to find such people.

I would like to acknowledge those who prefigured my career in science. The first must have been Frank White, my elementary school teacher, who fostered my interest just by having his own. Of tremendous importance was Ted Hammel, under whose auspices I did a high school science fair project. He guided me directly, as well as through my father.

I thank those who have helped me develop as a person doing science, instead of just a scientist. I must nod primarily to David Hessler. His influence was so strong that I am still adapting to his lead five years after his unjust death. Others include the sanity-bursts and support and love offered by George and Jen Rothrock, and Kim Kuney. Escape has been necessary for success, and for that I thank Jim and Deb Huntley, Jason Parks, Joe Stagg, Poseidon, Abrahadabra, 311, Shannon Hoon, Andrew K. Duff, Greggson Schachner, Tiffany Clark, Drew Waight and Jeff Hodson.

Each group of people has a unique collective identity and personality. And I've enjoyed being part of each one: the Levinites, the Karpenters, and (to some degree) the pomberanians. I have had epiphanies and excitements in the presence of these people, for which I blame them. I thank every one of them, although they are too numerous to mention. However, I wish to pay special attention to the

contributions of Mike McKeown (who taught me to stop worrying and love Science), Susan Forsburg (Science, she taught me, is an incessant goldmine), Jim Mason, and Harvey Karten. I cannot express my thanks sufficiently. Of particular importance were my interactions with peers and colleagues: David Arnosti, Mike Blower, Haini Cai, Mike Catlett, Kevin Cook, Joe Corbo, Ken Dobie, Susan Gray, Kumar Hari, Tony Ip, Cameron Kennedy, Hiep Le, Debbie Liang, Amy Lui, Tory McGrath, Terence Murphy, Jannette Rusch, Dan Sherman, Steve Small, Xiaoping Sun, Robin Truelove, Janice Wahlstrom, Bob Zeller, Jumin Zhou, and the members of the Chromoclub. But particularly Katy Donaldson. Despite our misunderstood coffee forays, I don't regret a thing; I don't know if I should thank her as a scientist, or a friend.

Grudgingly, I acknowledge the influence that Michael Levine has had on my science. I cannot say that my time in his lab was enjoyable, but it was perhaps important. I now understand treachery and the difference between success and happiness; pretty good lessons for a young man.

I thank Gerold Schubiger for his direction and advice. I could think of no better person to learn from; I am grateful that I have had the opportunity to watch him, learn from him, and share a friendship with him.

Excellent technical assistance was offered by Chris M. Yan, some of whose work is discussed in this dissertation. I have been lucky in finding such a dedicated

and talented scientist to work with, and I am pleased to claim interaction with him. I also thank Andrew Skora, whose excitement and attitude I wish I could emulate.

These experiments would not be possible without an incredibly open and giving community of *Drosophila* researchers. Although I thank each member in spirit, I acknowledge the materials or comments offered by specific individuals: Kent Golic, Pam Geyer, Lori Wallrath, John Tamkun, Pat O'Farrell, John Sedat, Aubrey de Grey, Kami Ahmad, Steve Henikoff, Mel Green, Manfred Frasch, Jim Posakony, Mike Ashburner, R. Scott Hawley, and all professors and students at the 1996 Advanced *Drosophila* Genetics course at Cold Spring Harbor. I return to Kent Golic and John Tamkun, and thank them also for laying a paved course for the next few years of my life in science: I wish this time to be successful and long-lived.

My work has been made better by the work of people whose work has allowed me to concentrate on chromosomes, and not the vagaries of life. Specifically, I think Tom Tomp, Gris Arrelano-Ramirez, Mark Lieberman, and Jamie Simon. I also thank Martha Stacklin and Gabrielle Wienhausen for helping me to become a better professor. They urged me and exorcised very specific demons, and showed me the way to fulfill a long-lived goal. I have a deep respect and gratitude for Ron McConnaughey, perhaps the best person-job match I have ever observed, whose manner is incomparable.

I thank each member of my committee. John Newport has been on both incarnations, and has served as a particularly insightful member in both cases. Barbara Wakimoto has been interested, despite the distance, and has been able to offer uniquely wise suggestions for both experiments and interpretations. I am grateful that Dan Lindsley accepted the request to be on my committee. Having him as a witness to my work has been rigorous, energizing, and inspiring. I thank Larry Goldstein for, once more, delving into chromosome biology. I am particularly indebted to Dan and Barbara for the time they gave and the depth of thought that they put into editing this dissertation. And I wish to thank, in particular, Gary Karpen. He accepted me into his lab under awkward conditions, and was supportive of my situation. I was made to feel welcomed and valued, and for that I am grateful. Gary taught me a whole new way to think about what I was doing, and why it was necessary. It was under his tutelage that I expanded my interests and learned to love research older than myself. He has forced me to pay attention to things that I hope will plague me: reviews, deadlines, managerial skills, excitement and attitude, and so much more. He has been a positive counterpoint to lessons I have under my belt.

Finally, and most frivolously, I thank my dogs, Phobos God-of-Fear, and Zoey (the honorary Deimos). They have no idea what I'm doing, but they're unabashedly passionate about their support.

Aspects of this dissertation have appeared in press. My work presented in Chapters 2 and 3 have been submitted for peer review and publication. Many models described in Chapter 6 appeared in a Genome Research Insight/Outlook article (Maggert and Karpen, 2000), a solicited commentary on Barry, *et al.*, 2000 (Barry, *et al.*, 2000).

Abstract of the Dissertation

Epigenetic regulation of centromere and neocentromere activity

in *Drosophila melanogaster*

by

Keith Andrew Maggert

Doctor of Philosophy in Biology

University of California San Diego, 2000

Gary H. Karpen, chair

The centromere is the region of a chromosome that is essential for proper segregation and inheritance of genetic information. Centromeres are regulated to occur exactly once per chromosome; failure to do so leads to whole-chromosome loss or chromosome damage, and loss of the linked genetic material. The mechanism for faithful regulation of centromere activity and number is not known. The presence of ectopic centromeres (neocentromeres) has allowed me to probe requirements and

characteristics of centromere activation, maintenance, and structure. I utilized chromosome derivatives that placed a 290 kilobase test segment in three different contexts within the *Drosophila melanogaster* genome – immediately adjacent to 1) centric heterochromatin, 2) centromeric chromatin, or 3) euchromatin. Using irradiation mutagenesis, I freed this test segment from the source chromosome, and genetically assayed whether the liberated test fragment exhibited centromeric activity. I observed that this test fragment behaves differently with respect to centromeric activity when liberated from these different contexts, despite an apparent sequence identity. Test segments juxtaposed to an active centromere produced fragments with neocentromeric activity, whereas test segments far from centromeres did not. Once established, this neocentromere activity was stable through thousands of cell divisions. Additionally, the imposition of neocentromeric activity on juxtaposed DNA showed that centromeric activity is capable of moving or spreading. Taken together, these observations indicate that the neocentromere, and by extension the centromere, of *Drosophila melanogaster* is regulated epigenetically.

CHAPTER ONE

Introduction

Introduction

The centromere was identified cytologically as the most pronounced constriction of a condensed metaphase chromosome by Walther Flemming in 1880 (Flemming, 1880). The centromere was called the first visible gene, and was easy to examine physically, so became the subject of many cytological and cytogenetic studies over the following 120 years. It is the site of interaction between the chromosome and the cellular machinery responsible for aligning, synapsing, and moving chromosomes through the mitotic and meiotic cell cycles (Rieder and Salmon, 1998). In this respect the behavior of the centromere is causative for familiar rules of Mendelian genetics. My work is concerned with two aspects of centromere activity. First, I investigate whether centromere identity is conferred by specific sequences found in the genome. Second, I describe arguments that bear on the stability of centromere activity and chromosome inheritance. Germane to these points are an understanding of the anatomy of the inheritance machinery, an understanding of sequence-based (genetic) and non-sequence-based (epigenetic) phenomena, and the interdependence of chromosome structure and chromosome function.

The Role of the Centromere

The metazoan centromere (Pluta, et al., 1995) corresponds to the chromatin that underlies the site of polar spindle fiber attachment, and is causative in faithful chromosome segregation in mitosis and meiosis. Microtubule – centromere interactions are mediated through a proteinaceous, trilaminar structure termed the kinetochore (Figure 1). The kinetochore houses distinct activities – microtubule capture, metaphase chromosome congression, and cell cycle checkpoint control (Rieder and Salmon, 1998; Dobie, et al., 1999). Similarly, the centromeric chromatin houses distinct behaviors – the reproducible location of the primary constriction

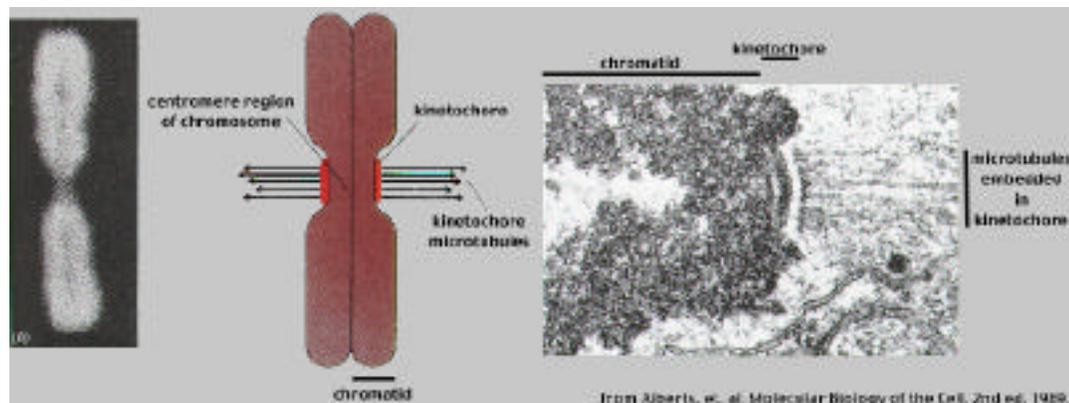


Figure 1: The Centromere

The centromere is seen as the largest constriction on a condensed metaphase chromosome. It is the site of kinetochore nucleation. Figure adapted from The Molecular Biology of the Cell (Alberts, et al., 1989).

and aspects of the pairing and association of replicated homologous chromosomes (sister chromatid cohesion). Through genetic, biochemical, and cytological investigation, we know much about the structure and behavior of the kinetochore and heterochromatin in the cell cycle. However, we know little about the mechanism of how centromere identity is specified.

Centromeres are limited to one per chromosome in almost all organisms. Most chromosomes carry a single centromere at a discrete and reproducible site. Exceptions exist in the ecdysozoa, specifically the *Homoptera* and *Heteroptera*, and possibly the *Lepidoptera* (White, 1954; Swanson, et al., 1967). Some representatives of these classes exhibit holocentric chromosomes, and behave as if the centromere is diffused along the entire chromosome. This is not unique to animals, as holocentric behavior is also known for the plant *Luzula purpurea* (Swanson, et al., 1967). Fragmentation of the genome does not generally result in loss of genetic material in organisms with holocentric chromosomes since each fragment is capable of nucleating a kinetochore and segregating to daughter cells (as in *Steinococcus* (White, 1954)). Some organisms contain alternate centromeres, where clearly different chromosomal loci behave as centromeres in different cell types. Specifically, germ cell mitotic centromeres in *Parascaris spp.* lie within euchromatin, while the meiotic centromere appears to be at the telomeres (Goday and Pimpinelli, 1989; Goday, et al., 1992).

Whether a chromosome has constant or alternative discrete centromeres, or a large diffuse centromere, genetic information is transferred faithfully. The location and behavior of centromeres may appear to be drastically different in different organisms or different cell types. In Chapter 6, I will return to the differences between centromere types, and attempt to show that they do not differ in an appreciable way from each other.

Deviation from exactly one active centromere per chromosome has drastic, often lethal, effects. A chromosome with no centromere is said to be acentric. Acentric chromosomes can be generated *via* recombination between a paracentric inversion and a chromosome of normal sequence when the crossover occurs within the inverted region (McClintock, 1987). Alternatively, an acentric chromosome may be generated by a double stranded break that separates a distal chromosome segment from the centromere (Heddle and Carrano, 1977; Schmid and Bauchinger, 1980; Oleincik, et al., 1994; Hahnfeldt, et al., 1995). Whether a chromosome may spontaneously become acentric is difficult to determine, as the product of such an event would be expected to be lost at the next cell division. However, centromeres are observed to be deactivated when linked to another active centromere (Sullivan and Willard, 1998). This inactivation occurs in the absence of overt structural alteration, and we may infer that the same event may occur on a structurally normal

chromosome, yielding an acentric chromosome. The mechanism of spontaneous centromere inactivation is unknown, but hypotheses borne from this and other work will lead me to models proposed in Chapter 6.

Several mechanisms can generate a chromosome with more than one centromere. For example, a recombination event that occurs within the inverted segment of a paracentric inversion and its normal homologue yields a dicentric chromosome (McClintock, 1987). A supernumerary centromere may also arise spontaneously or during the course of chromosome aberration. Such an ectopic centromere is called a neocentromere, in order to distinguish it from the pre-existing centromere (Rhoades, 1952). Like centromere inactivation, neocentromere activation is not understood, but hypotheses explaining their occurrence will be presented in Chapter 6 (Maggert and Karpen, 2000).

The cellular consequences of having an acentric or dicentric chromosome can be extreme. Owing to resultant aneuploidy, a common phenotype associated with chromosome aberrations of these types is cellular or organismal death. Phenotypes correlated with aneuploidy also include cancer, developmental disorders and mental retardation (Potter, 1991), and loosely-correlated phenomena such as aging (Koen, 1976), and Alzheimer's Disease (Potter, 1991). These phenotypes likely result from aneuploidy that results from loss of specific genic material during cell division (Figure

2, arrow) or to fragmentation of a dicentric chromosome that may undergo cycles of breakage-fusion-bridge (McClintock, 1987). These cycles may result in not only chromatin loss, but often cell cycle arrest or cell death

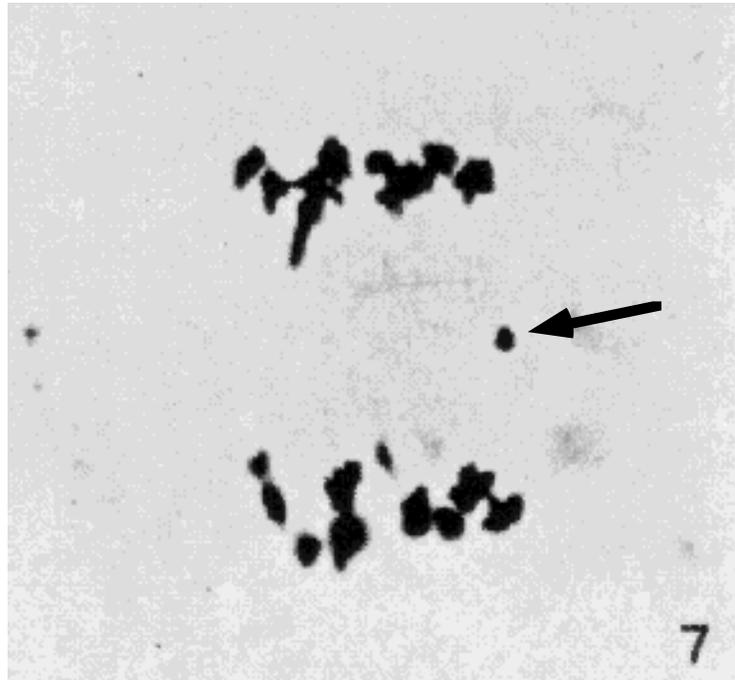


Figure 2: Fate of Acentric Chromosome

Chromosome spread during meiosis-I in *Zea mays* shows the product of a double-crossover (e.g. - 1,3 and 2,3 chromatid exchanges) between a paracentric inversion and its wild-type homologue. The acentric product is left at the meiotic plate during anaphase, and will be subsequently lost (figure adapted from McClintock (McClintock, 1987)).

(McClintock, 1987; Ahmad and Golic, 1999). Hence, the proper function of resident centromeres, and the ability of the cell to monitor and enforce a 'one-centromere-per-chromosome' situation is critical for proper genome integrity and cell viability.

As noted in the epigraphic quote (page v), Franz Schrader, and his contemporaries, noted that a chromosome can break within the centromere (Darlington, 1939; Schrader, 1939). If such a break occurred in prometaphase, after chromosome condensation and kinetochore nucleation on a metacentric chromosome, two telocentric chromosomes were generated. Each newly generated telocentric chromosome retained full function, as the resultant chromosomes neither elicited cell cycle arrest nor chromosome loss (McClintock, 1987). Additionally, the chromosomes were able to segregate for more than one cell division after breakage, suggesting that centromeric identity was intact, even during multiple rounds of DNA synthesis and kinetochore assembly and disassembly. These observations argue that the centromere may be internally functionally redundant. Such a model has been proposed by Brinkley (Wise and Brinkley, 1997), who noted an iterated, looped structure within stretched centromeres. The observations of Schrader, *et al.*, can be understood if the centromere is composed of loops of chromatin, and each loop is itself sufficient to imbue activity. Genetic and cytological data suggesting a repetition within the centromere are supported by molecular genetic studies in a variety of

organisms.

The Structure of the Centromere

Structural analyses of centromeres have most clearly elucidated the determinants of centromere identity in the eumycota, most notably *Saccharomyces cerevisiae*. Centromeres (CEN) in this yeast consist of a simple tripartite structure. CDEs I, II, and III correspond to 8 base pair, ~80 base pair, and 26 base pair, sequences with characterized protein-nucleic acid interactions (Figure 3) (Cumberledge and Carbon, 1987). Genetic and biochemical analyses have helped define DNA-protein and protein-protein interactions that are likely responsible for the activity and topology of the CEN (Murphy, et al., 1991; Sears, et al., 1995). The interactions between these three CDE-protein subcomplexes suggest a loop structure reminiscent of characteristics proposed for larger centromeres (Bechert, et al., 1999). Although little sequence identity exists between the *S. cerevisiae* CEN and the centromeres characterized in larger eukaryotes, functional homology may not be expected at the level of DNA sequence (Sunkel and Coelho, 1995; Dobie, et al., 1999). Still, it is compelling to note that some protein factors (*i.e.* – *Cse4p* and *CENP-A*) appear conserved among eukaryotes. Although the mechanism for directing these putative CEN proteins to the DNA may vary between organisms, once in place their

In most multicellular eukaryotes, the functional centromere is located within the heterochromatin, which contains large arrays of repetitive DNA (hundreds of kilobases to tens of megabases). These "regional centromeres" are so-named because centromere activity maps to tens to thousands of kilobases, and each kinetochore binds dozens of microtubules during metaphase and anaphase (Goldstein, 1981; Rieder, 1982; Hayden, et al., 1990). Because of their large size, our understanding of the molecular structure and composition of these regional centromeres has lagged behind molecular analysis of point centromeres. Thus, technical difficulties involved in sequencing, cloning, mutagenizing, and functionally analyzing the regional CEN has prevented an accurate comparison between regional and point centromeres at a molecular level. The possibility that a regional centromere may simply be an extended array of point centromeres is an attractive hypothesis that can only recently be investigated.

In the yeast *Schizosaccharomyces pombe*, the structure of the centromere is more similar to those found in animals. Specifically, the *S. pombe* centromere is a core of complex DNA (Figure 4, CC1 or cnt1), flanked by palindromically-arranged repeat blocks (B' or imr1R and L, and K" or otr1R and L) (Hahnenberger, et al., 1991; Partridge, et al., 2000). These centromeres are punctuated by tRNA genes (Figure 4, A, E, I). Clarke, Allshire, and co-workers

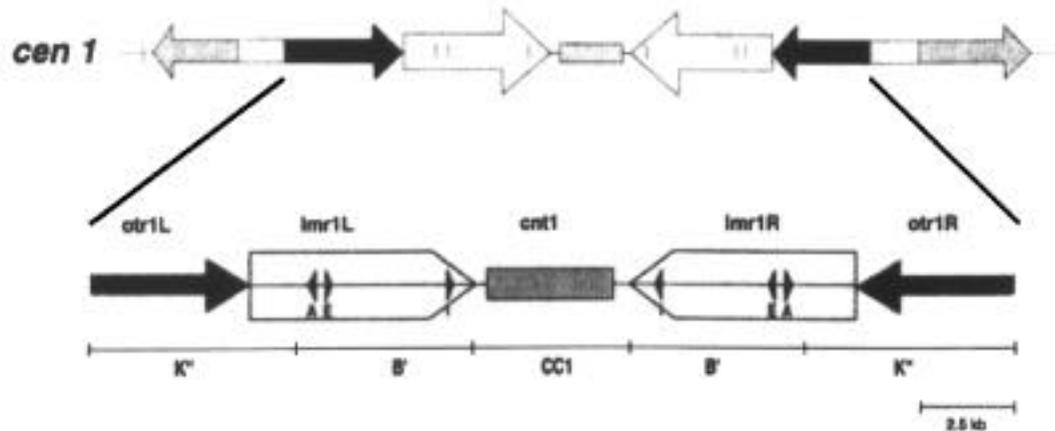


Figure 4: Structure of the *Schizosaccharomyces pombe* Centromere

Arrows indicate repeat units and directionality. The *cen1* locus is comprised of inner repeats (*lmrs*) and outer repeats (*otrs*) that flank a central core (*cnt*). *K*, *B*, and *cc* refer to older nomenclature (Hahnenberger, et al., 1991). *A*, *E*, and *I* are tRNA genes. Figure taken from Allshire, *et al* (Partridge, et al., 2000).

(Steiner and Clarke, 1994; Ekwall, et al., 1997) have shown that activity of the *S. pombe* CEN is clearly linked to this repeat-core-repeat locus, but it is also regulated epigenetically. The epigenetic characteristics have been investigated using minimal DNA constructs that contain centromeric DNA, which have been reintroduced to *S. pombe* cells and assayed for centromere function by observing episome loss rates. These episomal chromosomes can be recovered as either highly stable or unstable

populations. The two populations are identical in DNA sequence, but differ in their centromere function (Steiner and Clarke, 1994). A population of unstable chromosomes may give rise to populations of stable chromosomes, showing that the conversion of centromere activity is stochastic and itself maintained once established (Figure 5). Additionally, the activity of the episomal chromosome centromere is susceptible to alteration by chemical agents. Treatment of cells that contain a stable population of episomal chromosomes with trichostatin-A, a drug that affects acetylation equilibria within the cell, will destabilize the centromere and generate an

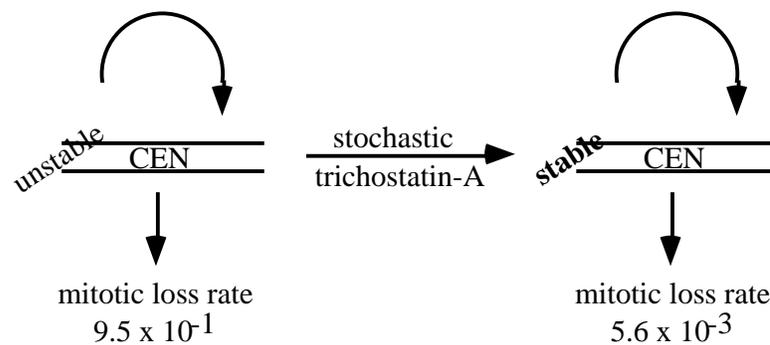


Figure 5: Epigenetic Regulation in the *Schizosaccharomyces pombe* Centromere

A centromere has two states, stable and unstable. Interconversion may occur randomly (Steiner and Clarke, 1994) or under induction by pharmacological agent (Ekwall, et al., 1997).

unstable centromere population (Ekwall, et al., 1997). This lowered stability is maintained, even after withdrawal of the drug. These two sets of experiments show that a conversion between stable and unstable centromere states can occur on identical sequence, and that the conversion can occur sporadically or by drug induction. This epigenetic alteration occurs on sequence predisposed to acting as a centromere, and likely represents a maturation of centromere activity through epigenetic means.

Karpen and co-workers have shown that the genetic activity of a *Drosophila melanogaster* centromere maps to a 420 kilobase region, comprised entirely of two repeat units (~235 kilobases of AATAT_n and ~150 kilobases of AAGAG_n), punctuated by single transposable elements, and bounded on one side by a 39 kilobase block of degenerate transposable elements (termed Maupiti, Figure 6) (Sun, et al., 1997). However, the sequences found at this characterized centromere are also found elsewhere in the genome, where they are not associated with centromere activity (Sun, et al., 1997). Hence, the repeat sequences seem insufficient to identify centromere activity. Moreover, Murphy, *et al.* have shown that some DNA sequences that are not present at normal centromeres can exhibit new centromeric, or neocentromeric, activity (Murphy and Karpen, 1995b; Williams, et al., 1998). As I will discuss below, these neocentromeres have raised the possibility that centromeres in *Drosophila melanogaster* are also regulated epigenetically.

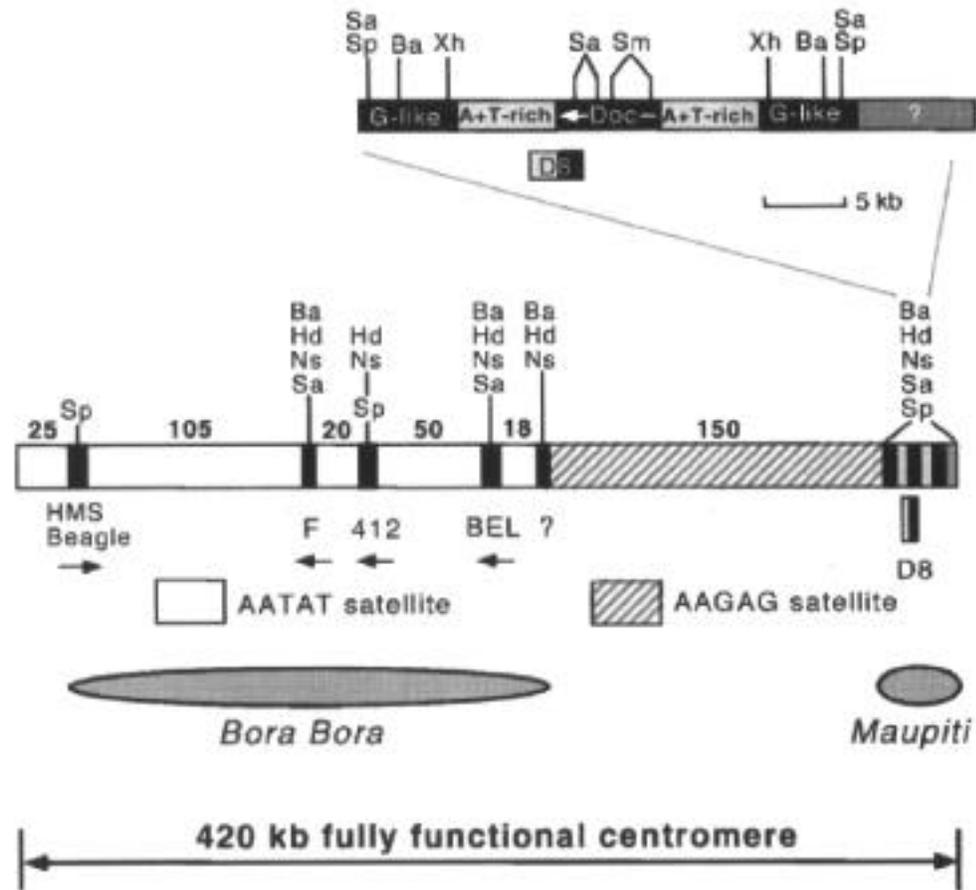


Figure 6: Structure of a *Drosophila melanogaster* Centromere

This centromere is derived from the X chromosome, and is 420 kilobases in size. Sequence analyses have shown that it is composed of simple satellites and transposable elements. Figure taken from Sun, *et al* (Sun, et al., 1997).

Comparatively, the *Homo sapiens* centromere shows the least amount of sequence complexity (Figure 7). The human centromere is located within a block of

apparently homogenous alphoid repeat (Jorgensen, 1997; de la Puente, et al., 1998; Marzais, et al., 1999). These blocks may encompass ten megabases of DNA, yet the centromere is a subset of this region, based on the location of the primary constriction, the localization of centromeric antigens, and the location of the kinetochore (Vig, et al., 1996). Humans also exhibit neocentromere activity in non-centromere

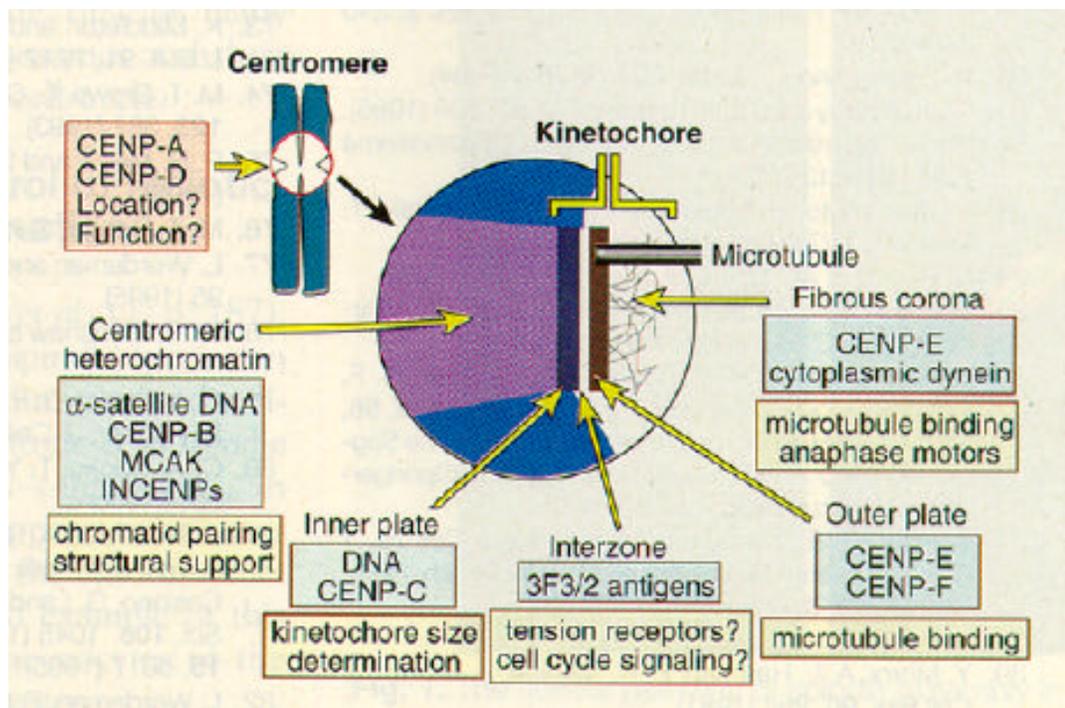


Figure 7: Structure of the Mammalian Centromere

The mammalian centromere occupies a subset of centric heterochromatin. It is not defined by sequence, but rather by a suite of centromere-specific factors. Figure from Pluta, *et al* (Pluta, et al., 1995).

regions (Barry, et al., 1999). Complete sequence has been obtained from a neocentromere, and from the parental source of this neocentromere that does not exhibit centromere activity (Barry, et al., 2000). The two sequences are identical, showing that neocentromere activity arose without structural alteration. Taken together, the unnecessary and insufficient contribution of specific sequence at active human centromeres has led many to believe that the centromeres of large metazoa may be commonly epigenetic.

Epigenetics clearly underlies some aspects of centromere identity – while *S. cerevisiae* may represent a case of a centromere with little or no epigenetic modification, *S. pombe* represents a case of overlap between sequence-determination and epigenetics. Thus, a continuum may exist among organisms with some having centromeres that are totally or largely sequence-based to those with significant epigenetic determination. It is possible that centromeres with no obvious consensus are mostly or entirely epigenetic in character. Before I can investigate or explain potential epigenetic regulation at the *Drosophila* centromere, I detail common and defining characteristics of epigenetic phenomena.

Epigenetics: Descriptions and Explanations

Epigenetic alterations have been known to plant breeders for centuries, but

their underlying causes are only now being widely appreciated. These phenomena have been known and studied for over 60 years in plants, animals, mycota, and eubacteria, and are exemplified by mitotically or meiotically heritable changes that are not concordant with alterations in DNA sequence. Epigenetics represents a violation of Mendel's observation that recessive characters may reappear unchanged in offspring of hybrids (heterozygotes) (Mendel, 1865). In these epigenetic cases, recessive characters are lost in the first post-hybrid generation, but may reappear in subsequent generations. These observations have two immediate explanations – allelic interaction (paramutation) and stochastic-yet-heritable change.

Paramutation describes templating interactions between homologous loci. In the case of *Zea mays*, paramutation is known to occur at no fewer than three loci (Chandler, et al., 1996). One canon of paramutation occurs at the *purple plant (pl)* locus (Chandler, et al., 1996). An allele, described as paramutagenic (*Pl-Rhoades*) is, by itself, dominant. However, like homozygous loss-of-function *pl* mutations, heterozygotes of *Pl-Rh/+* are purple. Upon outcross to a wild-type strain, Mendelian genetics predicts an offspring ratio of *purple:wild-type* at 1:1. What is observed, however, is all *purple* offspring. Subsequent generations of outcross will eventually produce revertants that are wild-type in coloration, and these revertants correspond to the original wild-type allele of *pl*. Hence, the *pl* locus is said to be paramutable.

Whatever the mechanism at work in *Pl-Rh/+* individuals, a change is conferred upon the paramutable locus by the paramutagenic homologue, and the change that is established confers a purple phenotype to the corn. The change is inherited mitotically and meiotically for many generations. Notably, the change is not permanent, being susceptible to spontaneous reversion.

The reversion to a pre-paramutated state is stochastic, and itself stable. This stochastic-yet-heritable change need not be limited to the erasure of a paramutated state. Stochastic changes may also manifest with the acquisition of a novel state without interaction with a previously altered or paramutable locus. Whether the reversion of paramutated alleles, the stability of *S. pombe* centromeres, or others, change may occur without detectable structural change. Frequencies of paramutation-induced change are universally higher than those that standard mutation would predict, and in fact no sequence polymorphisms are observed. Once established, paramutations are generally stable for many generations. The frequencies of changes may be affected by mutations elsewhere in the genome, by chemical agents, or by environmental effects. I generalize and state that epigenetic phenomena are subject to random fluctuation in state, much like any equilibrium condition. States, once established, are generally more commonly maintained than altered, and the energy of alteration may be affected by numerous factors (temperature, genetic background, *etc.*)

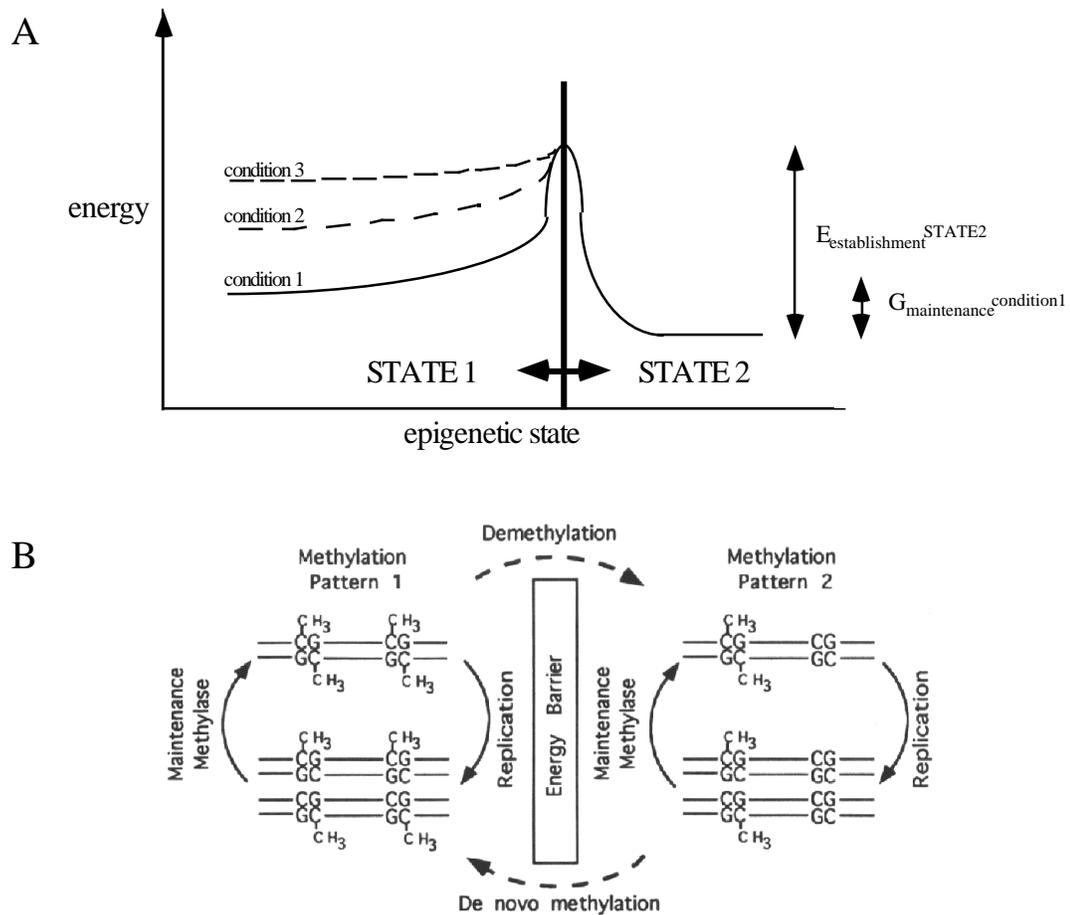


Figure 8: Energetics of Epigenetics

(A) A hypothetical model of the energy associated with epigenetic states and changes between them. Conditions 1, 2, and 3 could be environmental, genetic, *etc.*, conditions that affect the energy of one state. Differences in the energy levels of states will affect the equilibrium, and thus the frequency of spontaneous interconversion. Condition 1 will rarely convert to State 2, while Condition 3 will do so frequently.

$E_{\text{establishment}}$ will determine the frequency of conversion, and G will affect the ease of maintenance. (B) is a diagrammatic example taken from Riggs and Porter (Riggs and Porter, 1996), pertaining to methylation. In this case, maintenance methylases have a lower K_m (or higher V_{max}) with hemi-methylated DNA than with unmethylated or fully-methylated DNA (Lewin, 1990).

(Figure 8A).

Paramutation consists of two aspects that are separable in mechanism, yet are both required for the observation. First, paramutation involves the interaction between two sequences. In paramutation, this interaction occurs between homologous regions of separate chromosomes, and a paramutated state is established on the paramutated allele by the paramutagenic allele. Second, paramutation shows stability through cell division. The fact that a paramutation can be carried by a wild-type allele for numerous generations shows that alterations are self-templating during or after S-phase. Each of these two characteristics (establishment and maintenance) may be observed separately, and both are required for the phenomenon of paramutation, yet even in paramutation we can see that they are distinct. Paramutated alleles are themselves not paramutagenic, indicating that the first aspect of paramutation (interaction/establishment) is separable from the second (self-replicating). These characteristics are common to all epigenetic systems (*e.g.* – dosage compensation in mammals, imprinting in insects, mammals, and fungi, *etc.*) (Grewal and Klar, 1996; Russo, et al., 1996; Willard, 1996; Golic, et al., 1998). Universally, both establishment and maintenance must exist in epigenetic inheritance.

Paramutation likely increases the stochastic change between two epigenetic states. Two states may both be potential conditions for the same region of a gene, and

the rate of interconversion is low. Paramutation may not affect the mechanism by which similar DNA behaves differently, but may instead increase the frequency of establishment of one state or the other. This explains why states may convert spontaneously, yet changes are dramatically more frequent in paramutagenic conditions.

Non-sequence based mechanisms have been proposed for epigenetic inheritance. Each of half-a-dozen explanations have one feature in common: an energy barrier for *de novo* state change (the erasure or establishment of a new state) higher than a state maintenance. In conceptually the simplest model, methylation is causative for an alteration in gene expression (Figure 8B) (Riggs and Porter, 1996). Methylation is strongly correlated to epigenetic state in mammals, but its role as a causative agent for establishment is questionable (Bongiorni, et al., 1999; Jacobsen, 1999; Jakowitsch, et al., 1999; Jones and Wolffe, 1999; Ng and Bird, 1999).

As stated before, recent experiments have led investigators of metazoan centromeres to consider epigenetic regulation as a central mechanism for identity (Choo, 2000). Whether this is evidence from neocentromeres, in transformation and stability, in pharmacological alteration of centromere activity, the invocation of epigenetic contribution is inescapable.

Preludes To This Work

My work addresses the possibility that specification and regulation of the *Drosophila* centromere involves epigenetic aspects. My studies extend previous observations made by members of the Karpen lab, especially those of Dr. Terence Murphy. A pilot screen using a minichromosome, *Dp γ 238*, was first performed by Bonkovsky and Karpen, then followed by an extensive screen by Terence Murphy. *Dp γ 238* is a small chromosome that contains an active centromere and a structure that made it useful as a substrate for γ -irradiation (Murphy and Karpen, 1995b) (see Chapter 2 and Figure 11). Both noted the recovery of neocentromere-containing derivatives from *Dp γ 238*. These neocentromere-containing derivatives were devoid of any heterochromatin normally linked to centromeric activity, yet were presumed to have an active centromere based on their genetic stability through numerous generations. Murphy and Karpen have discussed the generation and stability of these neocentromere-containing fragments (Karpen and Allshire, 1997; Murphy and Karpen, 1998; Williams, et al., 1998). From these studies, Murphy and Karpen, *et al.*, proposed that neocentromeres had two characteristics. First, the neocentromere could be epigenetically regulated. This follows from the observation that, once activated, a neocentromere is stable despite the lack of centromeric chromatin. Epigenetic

regulation makes little claim as to the origin of activity, but predicts that activity is stable once established. The fact that derivatives from *Dpγ238* have been stable since their creation in 1994, having transited through thousands of cell divisions, argues for a stable mark that identified the neocentromere, much like the *S. pombe* cases described previously (Karpen and Allshire, 1997). Second, the neocentromere was proposed to 'spread' or 'ooze' onto outlying DNA from the endogenous centromere (Karpen and Allshire, 1997; Murphy and Karpen, 1998; Williams, et al., 1998; Maggert and Karpen, 2000). This model arose through the observation that the centromere of *Dpγ238* was juxtaposed to the sequence (termed "test segment" when attached to a centromere) that was later seen to exhibit neocentromere activity (termed "test fragment" after liberation and neocentromere manifestation).

This system, then, served as an ideal test for determining epigenetic contribution to neocentromere identity. In the one example tested (*Dpγ238*), two criteria were met. First, Murphy and Karpen demonstrated that establishment occurred on the test fragment derived from *Dpγ238*. This showed that a neocentromere could exist in *Drosophila*, and that the test fragment was sufficient to contain it, fulfilling the requirement of establishment. Second, whatever established neocentromere activity on the *Dpγ238*-derived test fragment, that activity was stable

once produced. This fulfilled the requirement of maintenance. What remained, then, was an investigation of what characteristics of *Dpγ238* allowed neocentromere establishment.

Prevailing Models for Centromere Activation: tests and predictions

The ubiquity of simple repeat DNA at regional centromeres in most metazoa has led to the belief that these sequences are causative in centromere identity (Harrington, et al., 1997). In this view, the primary sequence of the DNA is responsible for conferring the genetic activity of the centromere to the specific region of the chromosome, perhaps through sequence-specific DNA binding proteins. The study of CEN structure in *S. cerevisiae* has strongly supported this model (Hyman and Sorger, 1995). However, the sufficiency of repetitive DNA for centromeric activity has faced three strong criticisms. First, only a portion of the apparently homogenous blocks of centric repetitive DNA have centromere activity (Vig, et al., 1996). Second, centromeres may be turned off without apparent alteration in structure of underlying repetitive DNA (Page, et al., 1995; Page, et al., 1996). Third, centromeric activity may be found at sites in the genome that are devoid of the satellite sequences that are correlated with centromeres (Voullaire, et al., 1993; Barry, et al., 1999; Saffery, et al., 2000a; Warburton and al., 2000). Closely-related species, even individuals belonging

to the same species, may vary from each other in the constitution of repeat sequence at their centromeres (Talbert and Henikoff, 2000). Moreover, different chromosomes in the same individual show variation in the repetitive DNA arrays that contain centromeric activity (Sun, et al., 1997). Hence, most centromere have repetitive DNA, but the specific primary sequences are neither sufficient nor necessary to confer centromere activity.

It follows from these observations that multiple sequences may function as centromeres. The question, then, becomes one of understanding which sequences are chosen, and how other sequences are excluded.

Prevailing models can be divided to two types. Both models address how a chromosome might regulate centromere activity so only one centromere is active per chromosome, and how a neocentromere might arise.

The first model posits that a centromere is defined by a centromere consensus DNA sequence, or potential CEN (Figure 9A, pCEN) (Platero, et al., 1999; Koch, 2000). At each cell division, pCENs along a chromosome communicate with other sites, and one dominant centromere (CEN) is chosen. Activity at the dominant centromere prevents activity at other potential sites, hence polycentric behavior is prevented by active lateral inhibition. Should lateral inhibition be prevented (*e.g.* – by

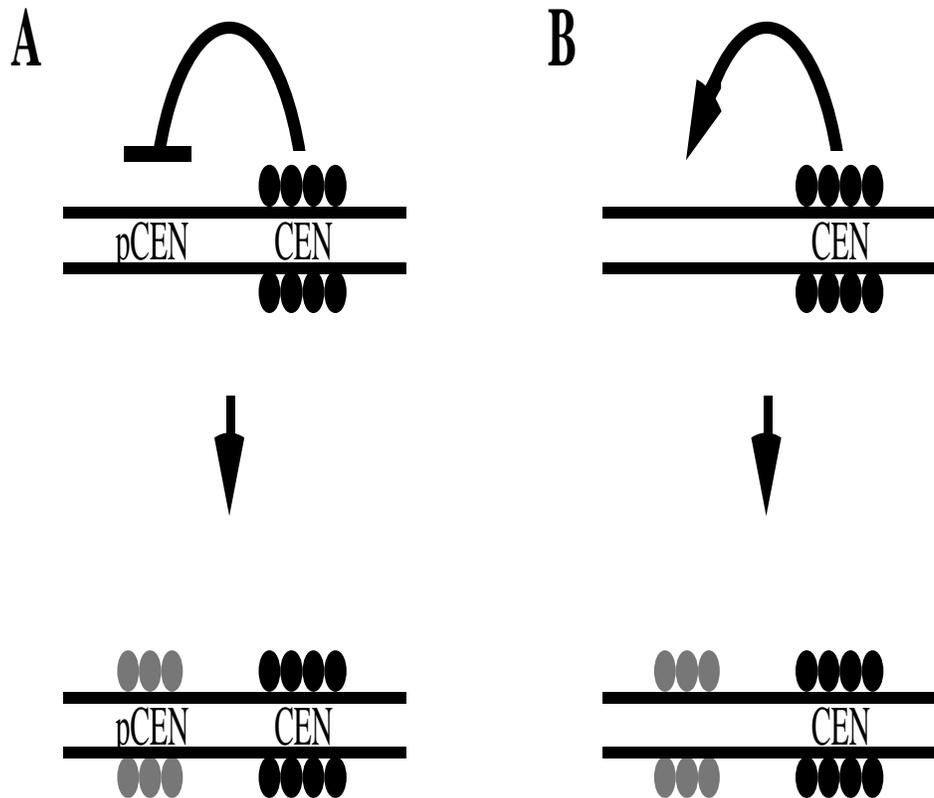


Figure 9: Two Models for Establishing Centromere Identity

In (A), a potential centromere (pCEN) is kept repressed by the activity (black circles) of a dominant centromere (CEN). After breakage (vertical arrow), the pCEN is capable of exhibiting centromere activity, and is observed as a neocentromere (gray circles). In (B), no region of the chromosome is inherently centromeric. Activity is transferred by an existing centromere to neighboring sequence, where it can be manifest as a neocentromere.

a breakage that separates a pCEN from a dominant and repressive CEN), then any sequence capable of acting as a centromere can then become active as a neocentromere (nCEN), since it is no longer *in cis* to a more dominant centromere.

The mechanism of centromere propagation is evident: sequences sufficient to act as centromeres are capable of interacting with the appropriate factors and establishing centromere activity *de novo* every cell division (Figure 10A).

The second model proposes that centromeres are epigenetically defined (Choo, 1997; Karpen and Allshire, 1997), and that neocentromeres observed from *Dp γ 238* are the product of the interaction of the extant centromere with outlying DNA (Figure 9B). Here, separating the neocentromere (nCEN) from the centromere does not allow activation, but merely allows the detection of neocentromere activity. In this case, sequence alone is insufficient to activate or maintain neocentromere activity, demanding that some characteristic of centromere activity remains associated with this locus through replication, and is retemplated onto newly synthesized DNA (Figure 10B).

These two models are not exclusive. The experiments in *S. pombe* suggest that some sequences have a higher propensity to form centromeres, and this inherent ability is further modified by epigenetic alteration (Steiner and Clarke, 1994; Ekwall, et al., 1997). *Dp γ 238*'s ability to liberate fragments that contain neocentromeres merely states that the test segment has the proper structure to allow neocentromeres to form. It does not state whether that ability is conferred by the presence of a potential centromere on the test segment, or whether the ability is conferred by some other

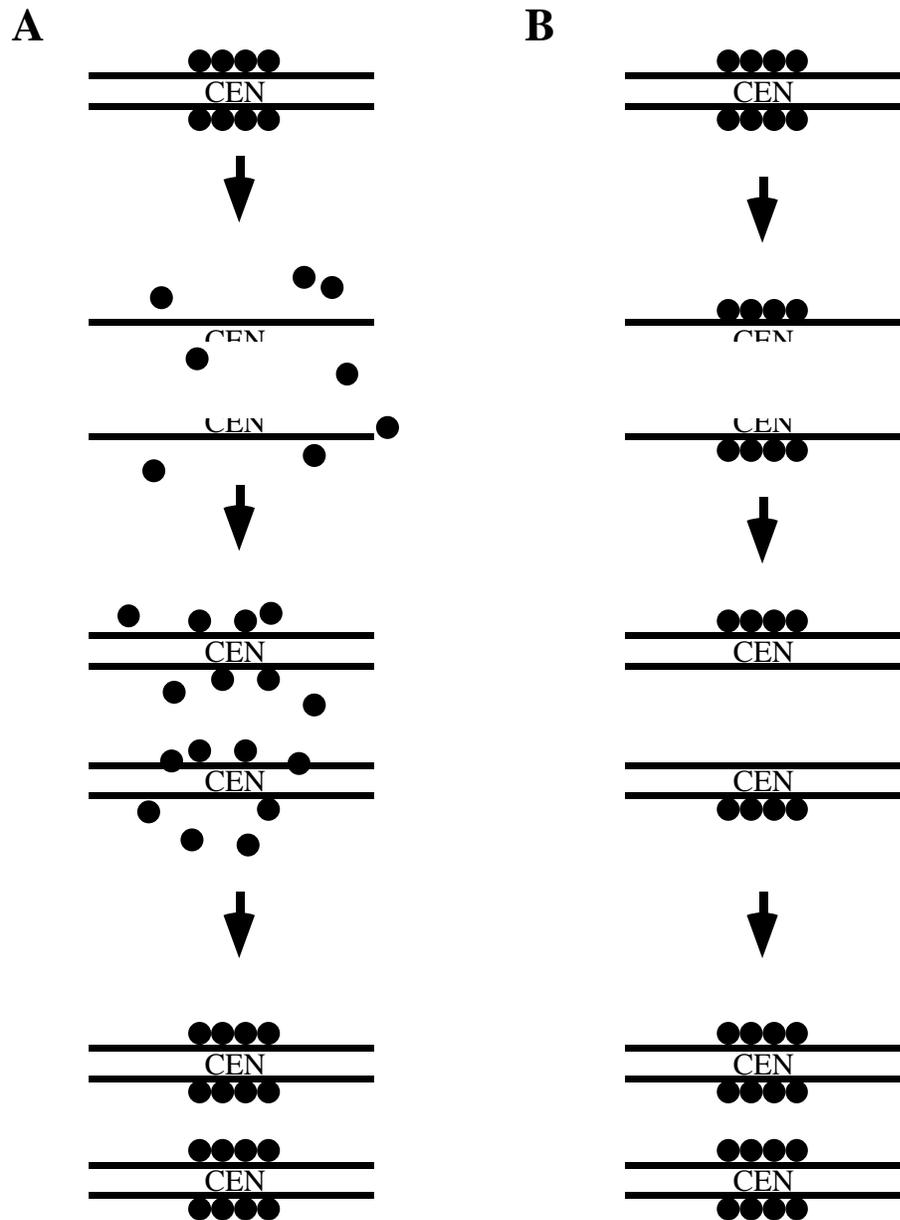


Figure 10: Implications of the Models for Centromere Maintenance

In (A), sequence is itself sufficient to reassociate centromeric activity (black circles) with centromere DNA (lines represent the two strands of each chromosome; the first and second arrows show denaturation and DNA synthesis, respectively). In (B), some factor causative for centromere activity remains associated through DNA synthesis, and recruits similar factors back to each newly-replicated chromosome.

aspect of the test segment's structure. By observing the ability of the test segment to manifest centromere activity when liberated from different sources, I assessed whether the sequence of the test segment is sufficient for neocentromere activity.

I tested the role of three different sequence contexts on the ability of the test segment to manifest neocentromeric activity using well-established approaches for generating and analyzing a new set of chromosome derivatives. If the context of the test segment played no role in the ability of neocentromeres to form, then I expected to see the liberated test segment to manifest neocentromere activity at the same rate from all sources. If context was important in specifying neocentromere activity, then I expected to see differences in the frequency of recovery of neocentromere-containing fragments from the three different source chromosomes.

The next chapter treats the ontogeny and structure of the chromosome derivatives that I will utilize during this study. Each one places the test segment in a different chromatin context, as defined by cytological, molecular, and genetic characteristics.

CHAPTER TWO

Minichromosome Derivatives of *Drosophila*

Minichromosome Derivatives of *Drosophila*

Drosophila melanogaster is amenable to cytogenetic analysis and manipulation. No experimental system has contributed more to our understanding of mitotic and meiotic chromosome behavior than *Drosophila*. This history of cytogenetics has given Drosophilists tools in the form of chromosome aberration stocks, specialized chromosomes, and detailed physical and cognate genetic and molecular maps of chromosomes. Regions of the genome can be moved, swapped, deleted, or duplicated within the bounds of chromosome stability and proper genetic constitution (Lindsley and Zimm, 1992). I have utilized a free duplication of the X chromosome and derivative chromosomes. These derivatives bear the benefit of being small, stable, and structurally characterized (Murphy and Karpen, 1995b). Hence, induced aberrations can be rapidly analyzed, and regions within these chromosome derivatives can be assessed for function.

I have investigated the mechanism of neocentromere activation by comparing the ability of a specific piece of DNA to display neocentromeric activity when placed in different chromosomal contexts. Specifically, I used chromosomal aberrations that placed an identical 290 kilobase “test segment” of DNA in three contexts. This test segment is originally derived from the tip of the X chromosome (1A1-1B1) and

contains approximately 290 kilobases of *X* telomeric chromatin (Lindsley and Zimm, 1992); this encompasses all of the *X* chromosome sequence that resides distal to the *yellow* locus. It also contains two 14.5 kilobase P-elements, each carrying a *rosy*⁺ gene. One of these elements is inserted in subtelomeric heterochromatin, about 40 kilobases from the chromosome end, and the other is inserted into euchromatin, approximately 105 kilobases from the chromosome end (Karpen and Spradling, 1992). This test segment, once liberated (denoted "test fragment") has been shown to exhibit neocentromeric activity (Murphy and Karpen, 1995b; Williams, et al., 1998). In this study, the 290 kilobase test segment was placed in three different chromosomal contexts: 1) juxtaposed to centric (but non-centromeric) heterochromatin in the *Dp8-23* chromosome, 2) juxtaposed to centromeric chromatin in *Dpγ238*, and 3) juxtaposed to euchromatin in the *Ty1337* translocation. Figure 11 summarizes the origins and overall structural features of the three chromosome aberrations used in this study. What follows is a detailed description of each chromosome.

Structure of *Dp8-23*

As summarized by Figure 11A – C, *Dp8-23* is a freely-segregating *X*-chromosome duplication. It is the product of several mutageneses, including two - induced chromosomal aberrations and two P-element transpositions.

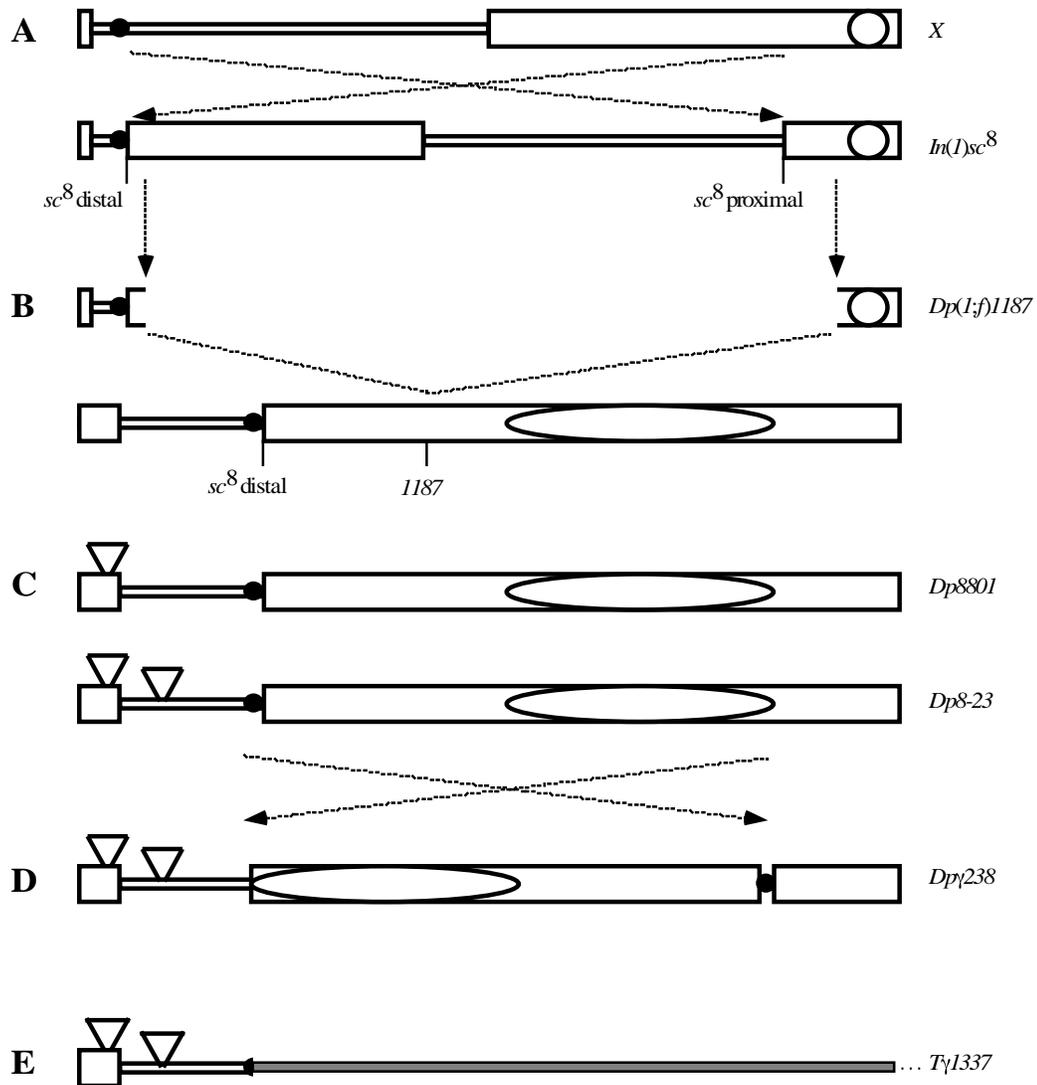


Figure 11: The Ontogeny of *Dp8-23*, *Dpγ238*, and *Tγ1337*

(A) Irradiation of *X* chromosome generated the paracentric inversion *In(I)sc⁸*. Thick bars are heterochromatin, thin bars are euchromatin, and the black circle is the *yellow* gene. The open circles and ovals indicate the location of the centromere. Dashed lines show inversions or deletions. (B) *In(I)sc⁸* was deleted internally (arrows), producing *Dp(1;f)1187* (*Dp1187*). The two chromosomes both represent *Dp1187*, but at different scale. (C) *Dp1187* was used as substrate for a *P*-element-mediated introduction of a *P{ry⁺}* element (triangle), generating *Dp8801*. *Dp8801* was used to make the *Dp8-23*. (D) *Dp8-23* was inverted, producing *Dpγ238*. (E) *Dp8-23* was also involved in a translocation with the telomere of 2L, producing *Tγ1337*.

The X chromosome was subjected to γ -irradiation in 1929 by Sidirov (Lindsley and Zimm, 1992). One of the products of these studies was the large inversion, *In(1)sc*⁸, whose breakpoints were in the *achaete-scute* complex, proximal to *yellow* (at approximately +30 kilobases, relative to *yellow* transcription start), and in the proximal heterochromatin element 1.688 (a 359-base pair repeat defined by its density of 1.688 g/cc, Figure 11A). Nearly 25 years later, Krivshenko and Cooper further irradiated *In(1)sc*⁸, generating a gross deficiency that was recovered as the free duplication, *Dp(1;f)1187* (*Dp1187*) (Lindsley and Zimm, 1992). *Dp1187* was a 2-break event that deleted the bulk of the chromatin in *In(1)sc*⁸. Breakpoints for *Dp1187* were approximately 260 kilobases into the inverted 1.688 repeat and approximately 740 kilobases from the right arm telomere (Figure 11B). This 1.29 megabase chromosome is the fusion of the left telomere, a block of 1.688, and the right arm, and retains all the elements of a normal chromosome – it has been stable mitotically and meiotically for over 40 years, and has been shown or inferred to contain telomeres, a centromere, and origins of replication. Karpen and Spradling used *Dp1187* as a substrate for the introduction of two *rosy*⁺-bearing P-elements, yielding *Dp8-23*, *y*⁺ *ry*⁺ (Figure 11C, Figure 12) (Karpen and Spradling, 1992).

For these studies, I have maintained *Dp8-23* in a *y*¹; *ry*⁵⁰⁶ background. This allows me to monitor the presence or absence of *Dp8-23* by body color (*yellow*⁺)

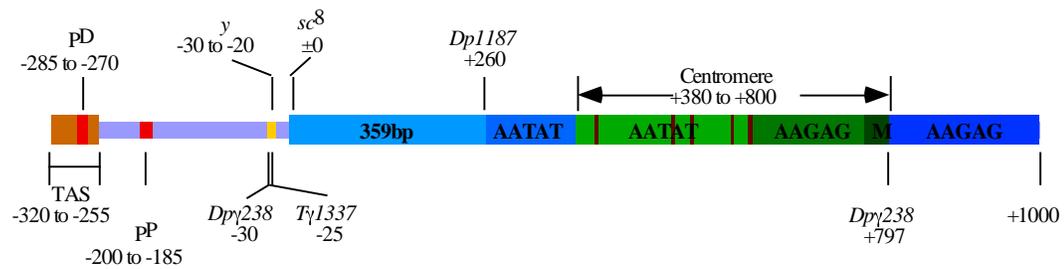


Figure 12: *Dp8-23* Structure and Landmarks

The entire sequence of *Dp8-23*. Color scheme is: orange, subtelomeric sequence; red, P-elements; light blue, euchromatin; yellow, *yellow* gene; dark blue, centric heterochromatin; green, centromeric chromatin; brown, transposable elements. Chromosome aberration breakpoints (*In(1)sc⁸* or *sc⁸*, and *Dp(1;f)1187* or *Dp1187*) and gene locations are marked relative to the *sc⁸* breakpoint at ± 0 kilobases. From here, heterochromatin coordinates are positive, and the euchromatic coordinates are negative.

phenotype or eye color (*rosy⁺*) phenotype. *Dp8-23* does not affect the viability of the fly, and so there is no preferential recovery of *Dp*-bearing or *Dp*-deficient progeny classes.

The *yellow* locus lies between approximately -30 to -20 kilobases, relative to the euchromatic/heterochromatic junction of *Dp8-23* (arbitrarily defined as ± 0 (Karpen and Spradling, 1990), Figure 12). PEV at *yellow* is mild on *Dp8-23*, and has been characterized elsewhere (Karpen and Spradling, 1990; Donaldson and Karpen,

1997). The euchromatin proximal to *yellow* abuts a block of the pericentric heterochromatin of the *X* chromosome. This heterochromatin is composed of 260 kilobases of the 359-bp repeat 1.688, and is in inverted-orientation relative to the *X* linked centric heterochromatin. This class of centric heterochromatic repeat is about two megabases from the cytological centromere in the *X* chromosome (Lohe, et al., 1993), and has never been seen to possess centromeric activity, either when linked to or separated from the *X* centromere.

Dp8-23 serves as one major substrate for the analyses that I performed. It has been genetically and structurally characterized by others in the lab. Its stability does not differ from that of *Dp1187*; it is mitotically stable and transits through meiosis with high fidelity for both segregation and disjunction. *Dp8-23* and derivatives thereof have been seen in copy numbers higher than two, indicating that nondisjunction may occur (Cook, et al., 1997). This seems to have little or no effect on viability, and copy numbers as high as six have been seen during routine cytological observations (unpublished observation). Although no constriction is visible cytologically, the stability of *Dp8-23* and the telomeric origin of the heterochromatic block argue that *Dp8-23* contains sequence identical to the *X* centromere. This centromere has been localized and analyzed. It corresponds to a 420 kilobases region of two classes of simple repeat ($(AATAT)_n$ and $(AAGAG)_n$) punctuated

by transposable elements. This minimal centromere, defined genetically through the use of a *Dp8-23*-derivative (Le, et al., 1995), is separated by 380 kilobases of 1.688 and AATAT_n from the test segment. This centromeric region is also capable of binding centromere-specific antigens, showing that it corresponds to a centromerically-active region of the chromosome (Williams, et al., 1998; Blower and Karpen, in preparation).

Dp8-23 contains diagnostic restriction sites that have been used in conjunction with pulsed-field gel electrophoresis (PFGE) to generate high-resolution aberration maps of derivatives of *Dp8-23* (Le, et al., 1995; Murphy and Karpen, 1995b; Sun, et al., 1997; Williams, et al., 1998). The *Dp8-23* region that contains the *sc*⁸ breakpoint is cloned, and serves as a unique marker for flies containing *Dp1187* or most of its derivatives.

Structure of *Dpγ238*

Dpγ238 is a free chromosome derivative that places the test segment next to an active centromere. It is derived from an inversion of *Dp8-23*, and possesses the same stability and genetic markers as *Dp8-23* (Le, et al., 1995).

This inversion was derived by irradiation of *Dp8-23*. Structural analysis showed that *Dpγ238* is a pericentric inversion that separated the *rosy*⁺ and *yellow*⁺

genes to opposite sides of the centromere (Figure 11D). The euchromatic inversion breakpoint of *Dp γ 238* was located near *yellow*, and the heterochromatic breakpoint was imprecisely mapped to the right arm of the *Dp8-23* (Le, et al., 1995). To exactly map the euchromatic and centric breakpoints of *Dp γ 238*, I used a strategy of scanning polymerase chain reaction (PCR) and inverse PCR (Spradling, et al., 1999). Together, I was able to map the *yellow*-linked breakpoint of the inversion to -5078 base pairs relative to *yellow* transcription. When the sequence of the centric breakpoint of *Dp γ 238* was compared to a fine-scale molecular analysis of the centromere from a related chromosome, *Dp γ 1230*, Janice Wahlstrom and I were able to determine that the heterochromatic *Dp γ 238* breakpoint was within the mapped centromere (Wahlstrom and Maggert, unpublished observation).

The right end of the centromere in *Dp8-23* is defined as the island Maupiti. Maupiti is 39 kilobases in size, and consists of copies of degenerated transposable elements and middle-repetitive motifs (Sun, et al., 1997) (Wahlstrom, personal communication). The *Dp γ 238* inversion breakpoint lies 3 kilobases from the right end of Maupiti, splitting the centromere into two pieces. The first piece is ~417 kilobases in size and juxtaposed directly to the test fragment. The other part of the centromere is 3 kilobases in size, and is juxtaposed to the *yellow*⁺ gene within a block of centric

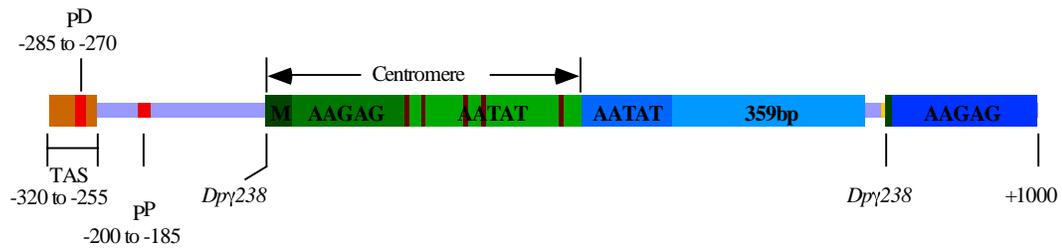


Figure 13: $Dp\gamma 238$ Structure and Landmarks

Symbols are the same as Figure 12.

heterochromatin (Figure 13). The sequence of the $Dp\gamma 238$ inversion breakpoints and a map of *yellow* showing the inversion breakpoint are reported in Appendix B.

The stability of $Dp\gamma 238$ is identical to its progenitor chromosome, $Dp8-23$ (Le, et al., 1995). Structurally, it appears to be a simple 2-break inversion. The overall size of $Dp\gamma 238$ is identical to the size of $Dp8-23$, and restriction patterns from $Dp\gamma 238$ match the predicted restriction pattern from a pericentric inversion of $Dp8-23$ (Le, et al., 1995). This, and higher-resolution restriction fragment comparison (Williams, et al., 1998), support the claim that $Dp\gamma 238$ is related to $Dp8-23$ by a single inversion event, and no sequences have been inserted or deleted.

Structure of *T γ 1337*

The final chromosome used as substrate in this study is the translocation *T γ 1337*. *T γ 1337* is a reciprocal translocation between the test segment on *Dp8-23* and the tip of chromosome 2L. This aberration places the test segment in a euchromatic context. Since this translocation is reciprocal, but lies proximal to haploinsufficient loci, the segregation products of *T γ 1337* include inviable combinations, which affect the predictions of the genetic screen discussed in Chapter 3.

When originally generated as a derivative of *Dp8-23*, *T γ 1337*, *ry⁺* was found to segregate away from chromosome 2 balancers (*CyO* and *SM1*). It was not resolvable on PFGE, and was thus larger than *Dp8-23* (Le, et al., 1995)(Murphy, unpublished observations). These data argued that it was a translocation with chromosome 2 (Figure 11E). The observation that it was *rosy⁺ yellow⁻* (Le, et al., 1995) suggested that it was either broken in the *yellow* region (and thus mutant for *yellow*) (Donaldson and Karpen, 1997) or was broken proximal to *yellow* and the *Dp^P* (Proximal, centromere-bearing) piece was lost while the *Dp^D* (Distal, telomeric) piece was appended to chromosome 2. This latter possibility demanded that the chromosome 2 breakpoint be telomeric to avoid aneuploidy (Lindsley, et al., 1972). The observation that *T γ 1337* could be made homozygous (data not shown) suggested that the

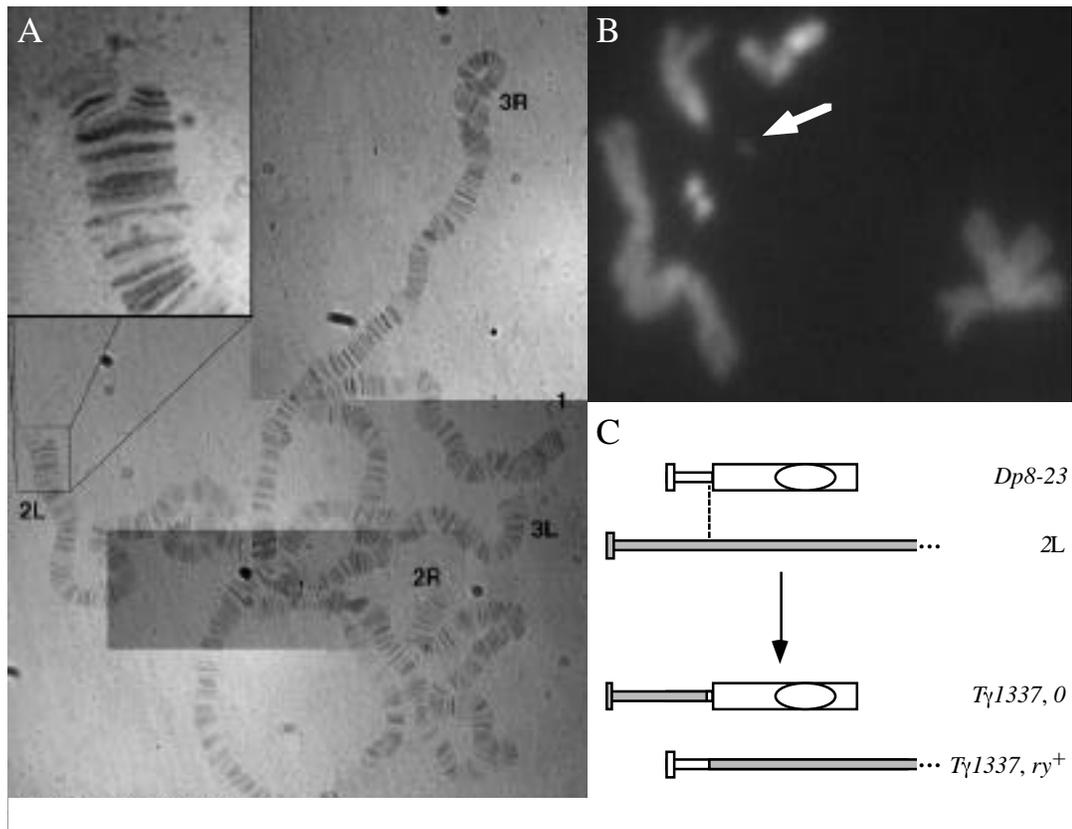


Figure 14: $T\gamma 1337$ Cytology and Structure

(A) Salivary gland chromosomes of a $T\gamma 1337/+$ individual. Inset shows close-up of 2L, and bifurcation indicates a translocation between distal 2L and $Dp8-23$. This chromosome is the $2^P Dp^D$ half. The absence of 2L banding in the chromocenter indicates the reciprocal half is too small to see. (B) Neuroblasts from a $T\gamma 1337$ individual. Minichromosomes ($Dp^P 2^D$) are present (arrows), showing that $T\gamma 1337$ is reciprocal. (C) Schematic of the translocation $T\gamma 1337$ and its products. The white chromosome represents $Dp8-23$, and the gray represents 2L. Upper set are original chromosomes, and the lower set shows that 2L and $Dp8-23$ have exchanged tips (producing the $Dp^P 2^D$ and $2^P Dp^D$, as shown).

translocation was reciprocal (indicating a break within the *Dp*-linked *yellow*), or the autosomal breakpoint was distal to any recessive lethal loci.

Analysis of salivary gland chromosomes from individuals heterozygous for $T(2;Dp8-23)\gamma I337$ ($T\gamma I337/+$) showed that the 2L breakpoint was distal to 21C (Figure 14A). Dp^P2^D chromosomes were not visible in these preparations (possibly due to their small size and their inclusion in the chromocenter) but were expected to be present in the strain since haploinsufficiency associated with deficiencies of this size (Lindsley, et al., 1972) would have resulted in lethality. Analysis of larval brains of $T\gamma I337/T\gamma I337$ individuals confirmed the presence of both halves of the translocation (Figure 14B).

$T\gamma I337$ is *yellow* due to the disruption of the *yellow* gene by the $T\gamma I337$ breakpoint. I used inverse PCR to clone the breakpoint and have shown that it falls within the 1st intron of *yellow* (at +712 base pairs relative to transcription start). Both halves of *yellow* can be independently cloned from $T\gamma I337$, ry^+ animals, confirming that the translocation is reciprocal and no sequence was lost (Figure 15). The two halves of $T\gamma I337$ are denoted $T\gamma I337$, ry^+ (properly $T(2^P;Dp8-23^D)\gamma I337$, ry^+) and $T\gamma I337$, 0 (properly $T(Dp8-23^P;2L^D)\gamma I337$, which is genetically unmarked). The former chromosome can be detected using the $rosy^+$ marker, while the latter

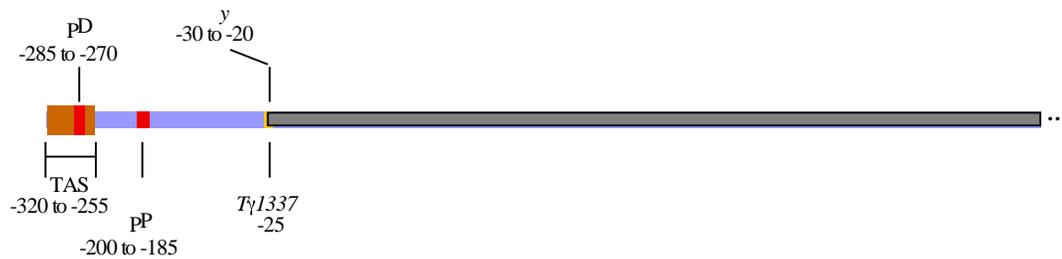


Figure 15: *Tγ1337* Structure and Landmarks

Symbols is the same as Figure 12. The portion of *Tγ1337* containing the test segment is shown. Stippled bar on the right is the chromosome 2 portion.

chromosome may be detected using the presence of one *Tγ1337* breakpoint. The sequence of *Dp8-23* and 2L translocation breakpoints are reported in Appendix B.

Tγ1337/Tγ1337 stocks also display an incompletely-penetrant held-out phenotype. This mutant phenotype is has not been linked to a specific locus.

Use of *Tγ1337* as a substrate for -induced breaks and subsequent recovery of derivative chromosomes required that I assess whether haploinsufficiency or duplication of the tip of 2L would affect viability. Lindsley, Sandler, and colleagues, have shown that small regions of the genome may be viable as haploids or triploid (Lindsley, et al., 1972). To determine if flies carrying one, two, or three copies of

distal 2L are equally viable, I scored the progeny of $T\gamma 1337/CyO, y^+$ males and wild-type females. As shown in Figure 16, the translocation heterozygotes produce four types of gametes with respect to constitution of chromosome 2. Provided all classes were equally viable, these four gamete types should be equally represented in progeny. All four gamete types are expected to produce individuals which could be distinguished by visible markers and molecular criteria (Figure 16). The progeny of this cross are reported in Table 1.

Data from Table 1 show that the region distal to the 2L $T\gamma 1337$ breakpoint is haplolethal and triplosubviable. Only 0.6 of the expected number of $y;+/CyO, y^+; ry; T\gamma 1337,0$ offspring are observed. The fact that the aneuploid individuals have reduced viability has bearing on the screen that I will present in Chapter 3. First, the subviability of the triploid class and inviability of the haploid class must be accounted for in determining the expected recovery of progeny bearing unbroken $T\gamma 1337$ chromosomes. Second, this difference will affect the proportion of neocentromere-containing test fragments to reciprocal breaks.

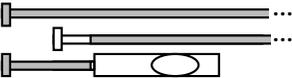
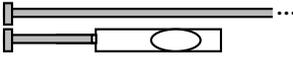
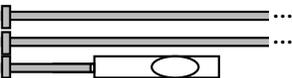
	sperm chromosome 2	zygotic chromosome 2 complement	<i>rosy</i> phenotype	<i>Tγ1337,0</i>	doses of 2L tip
	2L  ... <i>Tγ1337, ry⁺</i>  ...  <i>Tγ1337,0</i>				
1			+	+	2
2			+	-	1
3			-	+	3
4			-	-	2

Figure 16: Segregation Products of *Tγ1337*

Upon outcross, a heterozygous *Tγ1337/+* individual will produce four types of offspring. The genetic contribution by the *Tγ1337/+* gamete will be of four types (sperm chromosome 2), while the contribution by the female fly will be constant (a normal chromosome 2), generating indicated zygotic karyotypes. In row 1, both halves of the *Tγ1337* move together (alternate segregation). In row 2, only the *Tγ1337, ry⁺* chromosome is inherited (adjacent-1 segregation). In row 3, a wild-type 2L homologue is inherited with the *Tγ1337, 0* chromosome (adjacent-1 segregation). In row 4, the wild-type 2L homologue is inherited (alternate segregation). Note that all four classes can be distinguished using *rosy⁺* and PCR phenotypes.

Table 1: Viability of *Tγ1337* Segregation Products

Flies were scored for *rosy*⁺ phenotype, then subjected to PCR-based assay for the presence of *Tγ1337,0*. Offspring were not selected randomly with respect to *rosy*⁺, so the difference in *rosy*⁺ versus *rosy*⁻ is not meaningful. All offspring that contained the *Tγ1337,ry*⁺ also contained the *Tγ1337,0*, showing that the region distal to the *Tγ1337* breakpoint is haplolethal (H_0 states that the small product of the translocation would segregate randomly relative to the larger product of the translocation, and all products would be equally viable, $X^2 = 13.3$, $df = 1$, $p = 3.0 \times 10^{-4}$). For the *yellow*⁺ *rosy*⁻ offspring, the fraction with the small product of the translocation is 0.6. A X-square fails to reject the null hypothesis (that the transmission is 50%) at $\alpha = 0.05$ ($X^2 = 1.87$, $df = 1$, $p = 1.7 \times 10^{-1}$), but corroborating evidence from the irradiation of *Tγ1337* will be presented in Chapter 3.

	number assayed	<i>Tγ1337,0</i> present	<i>Tγ1337,0</i> absent	viability
<i>yellow</i> ⁻ <i>rosy</i> ⁺	20	20	0	0
<i>yellow</i> ⁺ <i>rosy</i> ⁻	51	19	32	0.6

Summary

From detailed molecular, genetic, and cytological analyses, I can confirm that the 290 kilobase test segment is identical in the *Dp8-23*, *Dpγ238*, and *Tγ1337* chromosomes, but that the chromatin context differs in each case. These results justify the use of these three rearrangements as substrates to test whether context affects the ability of a linked segment to manifest neocentromere activity. I will describe in the next chapter how the behaviors of these chromosomes, and their derivatives, serve as an assay for the behaviors of resident centromeres.

Work presented in this chapter has been prepared for submission to peer review and publication.

CHAPTER THREE

The γ -Screen

The γ -Screen

Previously, Murphy and Karpen have reported that *Drosophila* chromosome fragments are stably inherited despite lacking all centromeric heterochromatin (Murphy and Karpen, 1995b; Williams, et al., 1998). This separation of heritable DNA from centromere sequence necessitates the conclusion that a neocentromere has activated or newly specified on the recovered test fragment. Here, I present data demonstrating that the activation of the neocentromere can occur on naïve DNA at a high frequency if and only if that DNA is juxtaposed to an active centromere. This supports the proposition that centromeric activity can spread to neighboring chromatin, and that neocentromere identity is then propagated faithfully through replication and division. These results support an epigenetic model for the determinants of neocentromere and normal centromere activity.

Logic of the Screen

The logic of the screen is dependant on sequence-sufficiency for the generation or manifestation of neocentromere activity. Simply, if neocentromere formation is strictly sequence-dependant, then I expect that test segment sequence placed anywhere in the genome would show neocentromere activity. Observation of this activity is

contingent on the separation of the neocentromere from the linked endogenous centromere. This prediction has no bearing on mechanism, merely that the same segment of a chromosome (whatever its genomic location) would behave the same. Conversely, I test the hypothesis that the origin of the fragment has some bearing on its potential activity. In this case, the location of the segment would affect its ability to manifest neocentromeric activity.

Mechanics of the Screen

To distinguish between the strictly sequence-based and epigenetic models of centromere regulation, I irradiated *Dpγ238*, *Dp8–23*, and *T(2;Dp)γ1337* in parallel. The irradiation was done in females homozygous for the mutagen-sensitive mutation, *mu2^c*. This mutation confers sensitivity to low doses of ionizing radiation, and allows the recovery of terminal deficiencies which lack telomeric repeats. This approach has been shown to be fruitful in previous studies of the centromere (Murphy and Karpen, 1995b). Mason, and Murphy and Karpen have shown that the preponderance of 500 R-induced breaks result in one-break terminal deficiencies (Mason, et al., 1984; Murphy and Karpen, 1995b).

Females of genotype $y^2; mu2^c ry^{506} e$, and bearing *Dpγ238* or *Dp8–23* as monosomes, were exposed to 500 R of ionizing radiation from a ⁶⁰Co source and

immediately outcrossed to $X^{\wedge}Y, y/0; ry^{506}$ males (Figure 17, P generation). Male offspring of this cross were sterile ($X/0$) (Bridges, 1916) and not scored. All female offspring were scored for the loss of one *Dp*-linked marker ($rosy^+ yellow^-$ or $rosy^- yellow^+$), indicating a breakage event that removed part of the *Dp*. The presence of *Y*-chromatin in the female offspring was included to suppress variegated position effects and permit the recovery of strongly variegating $rosy^+$ or $yellow^+$ derivatives that might otherwise be scored as $rosy^-$ or $yellow^-$ (Ashburner and Novitski, 1976). Derivative

-irradiation (500 Rad)

P	$X, y^{\dagger}; mu\mathcal{Z} ry^{506} e; Dp(Dp8-23 \text{ or } Dp\gamma 238)$	♀	x	$X^{\wedge}Y, y^{\dagger}/0; ry^{506}$	♂
F1	$X^{\wedge}Y/0$	x	$X/X^{\wedge}Y; Der(y^- ry^+ \text{ or } y^+ ry^-)$	x	X/Y
F2	$X^{\wedge}Y/X^{\wedge}Y$	x	$X^{\wedge}Y/0; Der$	X/X	x $X/Y \text{ or } X^{\wedge}Y/Y; Der$
F3	$X^{\wedge}Y; Derstock$	X/X	x	$X/Y; Der$	
F4		$X; Derstock$			

Figure 17: Screen to Recover γ -Induced Derivatives of *Dp* γ 238 and *Dp*8-23

Virgin females bearing single copies of *Dp* γ 238 or *Dp*8-23 ($yellow^+ rosy^+$) were irradiated with 500 R of radiation, then outcrossed to $X^{\wedge}Y/0$ males. Progeny bearing new derivatives were scored as $yellow^+ rosy^-$ or $yellow^- rosy^+$. Subsequent generations isogenized the genetic background in order to determine the structural and genetic characteristics of all derivatives.

chromosomes were introduced into strains carrying a regular-*X* or attached-*XY* ($Y^S X.Y^L$, $In(1)EN$, y^1 ; $X^A Y$), as shown in Figure 17 F1-F4, before determining their structures (Chapter 4) and transmission rates (Chapter 5). No attempt was made to remove *mu2* or *e* from the genetic background as tests showed that neither locus has an effect on transmission (data not shown).

For the γ -irradiation mutagenesis of *T γ 1337*, females of genotype y^2 ; $T(2^P;Dp^D)\gamma 1337$, ry^+/CyO , $P\{SuPorP, w^+ y^+\}$; $mu2^c ry^{506} e$; $T(Dp^P;2^D)\gamma 1337$, $0/0$ and were irradiated and

A	-irradiation	
P	$X, y^1; T\gamma 1337, ry^+/CyO, y^+; ; mu2^c ry^{506} e$ ♀	x $X^A Y, y^1/0; 2/2; ry^{506}$ ♂
F1	$XX^A Y; T\gamma 1337, ry/2$	x $X^A Y/Y; +Sp$
F2	$2/Sp$ or $T\gamma 1337, ry/Sp$	
B	-irradiation	
P	$X, y^1; T\gamma 1337, ry^+/CyO, y^+; ; mu2^c ry^{506} e$ ♀	x $X^A Y, y^1/0; 2/2; ry^{506}$ ♂
F1	$XX^A Y; CyO, y^+/2; Det(ry^+)$	

Figure 18: Screen to Recover γ -Induced Derivatives of *T γ 1337*

Virgin females heterozygous for *T γ 1337* ($yellow^+ rosy^+$) were irradiated with 500 R of radiation, then outcrossed to $X^A Y/0$ males. Progeny bearing new derivatives were scored as (A) $yellow^- rosy^-$ or (B) $yellow^+ rosy^+$.

immediately outcrossed to $X^AY, y/0; ry^{506}$ males (Figure 18). Offspring of this cross were scored for $yellow^- rosy^-$ or $Curly yellow^+ rosy^+$ phenotypes to detect new derivative chromosomes.

A terminal deficiency of $T\gamma 1337, ry^+$ that removes $rosy^+$ and a wild-type

A

F2 $XX; Sp/CyO$ x XY or $X^AY; 2/Sp, T\gamma 1337, 0$ or 0

F3 $XX; Sp/CyO$ x $XY; 2/CyO, T\gamma 1337, 0$ or 0

F4 XX and $XY; 2/CyO, T\gamma 1337, 0$ or 0

F5 $X; 2/CyO$ stock

B

F2 $XX; Sp/CyO$ x XY or $X^AY; T\gamma 1337, ry/Sp, T\gamma 1337, 0$

F3 $XX; Sp/CyO$ x $XY; T\gamma 1337, ry/CyO; T\gamma 1337, 0$

F4 XX and $XY; T\gamma 1337, ry/CyO; T\gamma 1337, 0$

F5 $X; T\gamma 1337, ry/CyO$ stock

Figure 19: Scheme to Recover and Balance $T\gamma 1337, ry^-$ Derivatives

$yellow^-$ derivatives of $T\gamma 1337, ry^+$ were balanced as groups. Since the paternal chromosome 2 and the maternal $T\gamma 1337, ry^-$ chromosomes are indistinguishable genetically, I balanced 8-15 individuals from the F2 of Figure 18A. After balancing, I was able to determine when I "caught" new *Der* or a chromosome 2 by assaying for the $T\gamma 1337, 0$ chromosome (which is required in conjunction with $T\gamma 1337, ry^-$ derivatives, yet is subviable in conjunction with chromosomes 2). (A) shows the outcome for "balancing" a chromosome 2, and (B) shows the balancing of a *Der* ($T\gamma 1337, ry^-$).

paternal chromosome 2 are indistinguishable on the basis of phenotypes of offspring since neither bears a visible marker. For this reason, a balancing scheme for these *rosy*⁻ chromosomes was introduced into the screen using 8-15 F2 individuals as outlined in Figure 19. Subsequent genetic and PCR-based assays distinguish between *Tγ1337* derivatives and unbroken chromosome 2. The presence of *Tγ1337,0* is required for viability of derivatives of *Tγ1337*, *ry*⁺, and is a detriment to flies of otherwise wild-type karyotype. After the four generations required for balancing, the presence or absence of *Tγ1337,0*-specific PCR product was a reliable diagnostic for the chromosome 2 constitution of the original F2 individual.

Results of the Screen

Progeny bearing new derivatives of *Dpγ238*, *Dp8-23*, or *Tγ1337* were isolated and stocks made of each new chromosome derivative. Each derivative was placed into a phenotypically distinct class, as detailed below. The frequency of each class of derivative from each chromosome was compared to determine if each substrate chromosome gave rise to derivatives at the same rate.

Progeny from the irradiations of *Dpγ238* and *Dp8-23* belong to four phenotypically distinct classes. Classes 1 and 2 represent F1 progeny receiving

(*yellow⁺ rosy⁺*) or not receiving (*yellow⁻ rosy⁻*) the free and unbroken substrate

Dp(yellow⁺ rosy⁺), respectively (Figure 20). Members of these classes were recovered

frequently and in equal proportion. At a low frequency (Table 2), I also recovered

new derivatives. Class 3 (*yellow⁺ rosy⁻*) or Class 4 (*yellow⁻ rosy⁺*) females represent

Class	<i>Dpγ238</i>	<i>Dp8-23</i>	<i>Tγ1337</i>
1			
2			
3	A ----- B	A ----- B	A ----- B
4	A ----- B ----- C	A ----- B ----- C	A ----- B ----- C

Figure 20: Classes of Chromosomes Recovered from γ -Screens

Chromosomes recovered after the irradiation of *Dpγ238*, *Dp8-23*, and *Tγ1337* are divided into classes, based on phenotype they confer and location of the breakpoint. Chromosome pictures follow the symbolism of Figures 11 and 14. Descriptions of classes can be found in text.

derivatives of the *Dp* that arose through one-break events that remove part of the minichromosome containing one or the other genetic marker. Class 3 derivatives of *Dp* γ 238 and *Dp*8–23 were expected to be breaks within the euchromatin, and retained the entire block of heterochromatin including the *yellow*⁺ gene while losing a portion of the test segment, including the *rosy*⁺ genes. Class 4 derivatives of *Dp* γ 238 were expected to fall within three subclasses. Subclass 4A have breaks between the *rosy*⁺ genes and the centromere in the test segment, generating structurally acentric *Ders* (*SA Ders*) upon which neocentromere activity is unequivocal. Subclass 4B have breaks within the centromeric region. Subclass 4C chromosomes have a break within the heterochromatin, between the centromere and *yellow*⁺. Because of the placement of the *rosy*⁺ and *yellow*⁺ genes in *Dp*8–23, and the phenotypic criteria of the screen, this chromosome was expected to only give *Ders* of Classes 3 and 4A.

Chromosomes recovered from the γ -ray irradiation of *T* γ 1337 fell into four classes, as well. Classes 1 and 2 represented unbroken *T* γ 1337 (*yellow*⁻ *rosy*⁺) and *CyO* (*Curly yellow*⁺ *rosy*⁻) chromosomes (Figure 20). These classes, like Classes 1 and 2 from *Dp* γ 238 or *Dp*8–23, were frequent and discarded. New derivatives were isolated in two new classes. Class 3 females lost the P{*ry*⁺} test fragment from *T* γ 1337, and were scored as *yellow*⁻ *rosy*⁻. Class 4 represents the complementary

chromosomes to Class 3, and were scored as *yellow⁺ rosy⁺*, indicating a cosegregation of the *CyO*, *y⁺* and *Tγ1337*, *ry⁺* chromosomes. Since triploidy for chromosome 2 is lethal (Lindsley, et al., 1972), this class would most likely represent a break proximal to the *rosy⁺* P-elements in *Tγ1337*, followed by cosegregation with the *CyO*, *y⁺* chromosome. It is this class that would contain potentially neocentromeric test fragments. Both Class 3 and Class 4 from *Tγ1337* may break in chromatin distal to (subclasses 3A and 4A) or proximal to (subclasses 3B and 4B) the *Tγ1337* breakpoint. Statistically, only subclasses 3A and 4A are comparable to the *Ders* from *Dpγ238* or *Dp8-23*.

Frequencies of Offspring

Table 2 shows the results of the -screens of *Dpγ238*, *Dp8-23*, and *Tγ1337*. A molecular analysis of breakpoints, and a revised table showing the frequencies of specific subclasses of breaks, will also follow in Chapter 4 (Table 4).

Work presented in this chapter has been prepared for submission to peer review and publication.

Table 2: Chromosome Derivatives: Numbers

These are the uncorrected data from the irradiation of *Dp γ 238*, *Dp8-23*, and *T γ 1337*. Class designations refer to Figure 20. None of these derivatives are subdivided based on structure. Since the target size of chromatin differs between classes, these numbers may not be used to determine break frequency without correction. Those data will be refined and presented in Chapter 4 (Table 4).

	<i>Dpγ238</i>	<i>Dp8-23</i>	<i>Tγ1337</i>
total chromosomes (Class 1)	26,365	94,968	
<i>yellow⁺ rosy⁻</i> (Class 3)	28	21	
<i>yellow⁻ rosy⁺</i> (Class 4)	49	1	
<i>yellow⁻ rosy⁺</i> (Class 1)			25,910
<i>yellow⁺ rosy⁻</i> (Class 2)			32,930
<i>Curly yellow⁻ rosy⁻</i> (Class 3)			28
<i>yellow⁺ rosy⁺</i> (Class 4)			2

CHAPTER FOUR

Structural Analyses of Minichromosome Derivatives

Structural Analyses of Minichromosome Derivatives

Few studies in *Drosophila* allow a high-resolution analysis of the distribution of chromosome breakpoints from multiple sources. Previous studies have been limited to cytological (Mason, et al., 1984; Green, et al., 1987) or genetic (Mason, et al., 1984; Mason, et al., 1986; Green, et al., 1987; Mason, et al., 1997) resolution, or report breakpoint densities without an understanding of chromatin context (Laurenti, et al., 1995; Murphy and Karpen, 1995b). The distributions of breakpoints of chromosomes generated during this study may help us understand what structural requirements exist for chromosome stability and metabolism in *Drosophila*.

It is generally assumed that breakpoints by gamma radiation are randomly distributed throughout the genome (Ashburner, 1989). An analysis of the recovery of new breakpoints in the *Dp γ 238*, *Dp8-23*, and *T γ 1337* chromosomes tests that assumption. Additionally, the frequencies of breaks within defined target sequence can be used to investigate context effects as well as understand the frequency of neocentromere activation on the test fragment.

After a description of the methods used in size and breakpoint determination, I will proceed to apply statistical comparisons of breakpoint distributions on the three substrate chromosomes. My resolution is limited by the resolution of separating

whole chromosomes or large restriction fragments on pulsed-field agarose gels. Based on agreement between multiple preparations of identical derivatives, I estimate my resolution to be of about 5 kilobases (Le, et al., 1995; Murphy and Karpen, 1995b; Williams, et al., 1998). I calculate the size of gaps between successive chromosome breaks, and test whether these conform to a normal distribution. In an attempt to correct for the limits of resolution, statistical analyses done on these continuous data were reproduced with nonparametric tests run on 5 kilobase and 10 kilobase bin sizes. In no case did the results differ appreciably. Populations were compared using Student's *t* and Kolmogorov-Smirnov tests, and relative frequencies were assessed using a X^2 or V^2 (if two or more data were less than 10).

Methodology for Determining Derivative Structure

Techniques for determining the structures of minichromosomes have been developed in the Karpen lab, and were used in this study (Karpen and Spradling, 1992; Le, et al., 1995). Pulse conditions, and restriction protocols can also be found in Appendix A. I determined sizes by comparing unrestricted, NotI-restricted, and SfiI-restricted samples (see Appendix B for restriction maps of these chromosomes). Data from individual samples were weighted to reflect higher resolution of smaller fragments, as described in Appendix A.

Structures of the *Dp* γ 238 *yellow*⁺ Derivatives

Twenty-eight derivatives were generated by irradiation of *Dp* γ 238 and scored as *yellow*⁺ *rosy*⁻, and hence belong to Class 3A. The frequency of recovered derivatives of this class is 1.06×10^{-3} derivatives/chromosome ($f = 1.1\text{E-}3$

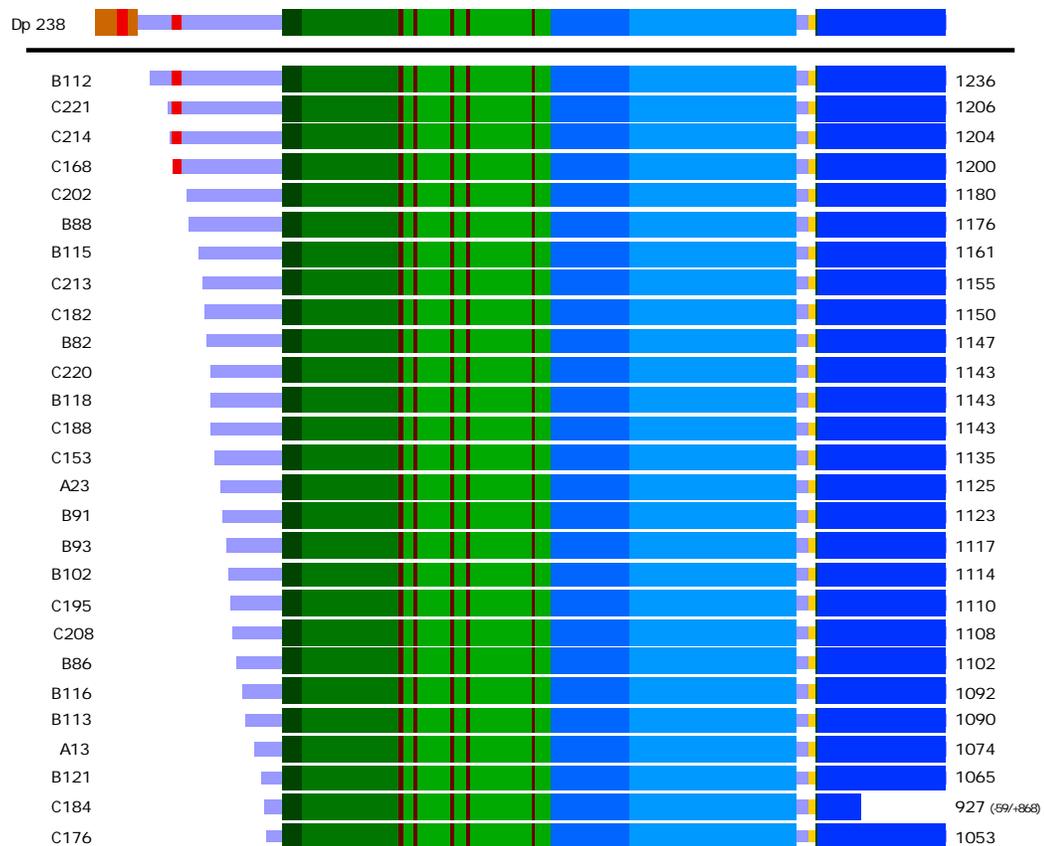


Figure 21: Distribution of Class 3A (*yellow*⁺) Breaks of *Dp* γ 238

Unbroken *Dp* γ 238 is shown above the line. Numbers to the left of chromosomes are their screen names, numbers to the right are sizes in kilobases. Symbolism is that of Figure 12. Derivative C184 is a 2-break event, removing some of the right arm.

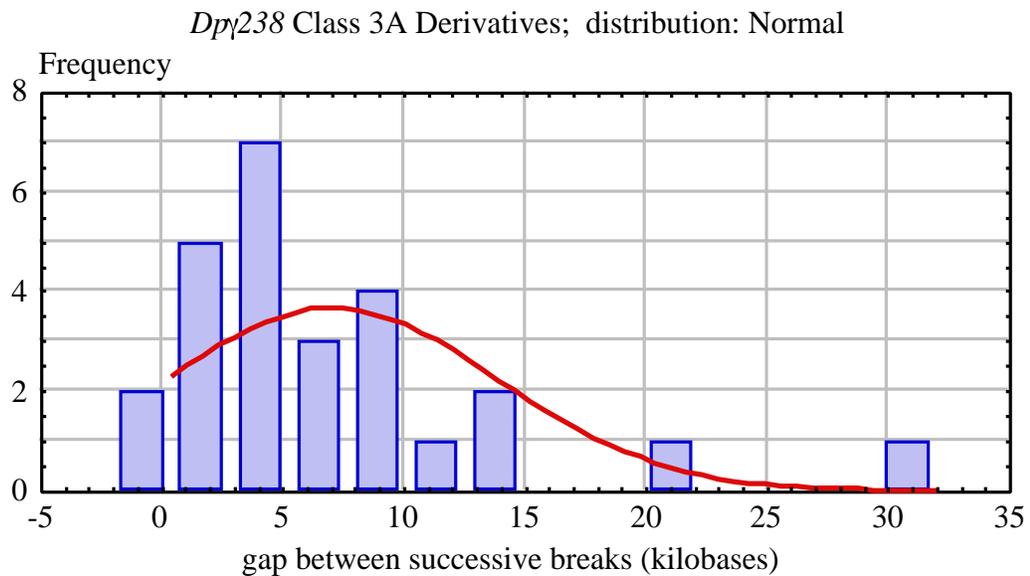


Figure 22: Normal Distribution of Breakpoints of *Dpγ238* Class 3A Derivatives

Comparison of *Dpγ238* Class 3A derivatives with a normal distribution. A measure of distance (in kilobases) between successive chromosomes when sorted by size is shown on the X-axis. χ^2 and Kolmogorov-Smirnov tests show $p > 0.05$.

derivatives/chromosome).

Data showing the distribution of breakpoints are shown in Figure 21. Within the euchromatin, the breakpoints of *Dpγ238* derivatives are randomly distributed (H_0 : derivatives are drawn from a normally distributed population, $p = 0.149$, Figure 22).

Structures of the *Dp8-23 yellow⁺* Derivatives

Twenty-one breaks were recovered in the euchromatic arm of *Dp8-23*, all

recovered as *yellow⁺ rosy⁻* derivatives ($f = 2.2E-4$ derivatives/chromosome). As with the Class 3A derivatives of *Dp γ 238*, these Class 3A derivatives have breakpoints that are distributed randomly. These data are shown in Figure 23 and analyzed for normality in Figure 24 (H_0 : derivatives are drawn from a normal population, $p = 0.20$).

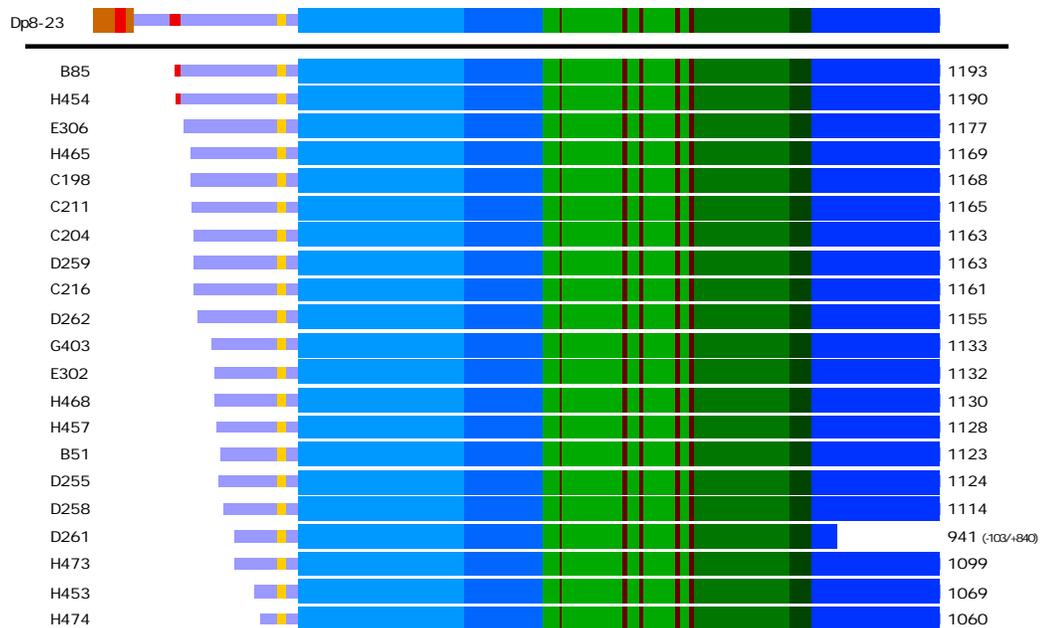


Figure 23: Distribution of Class 3A (*yellow⁺*) Breaks of *Dp8-23*

Unbroken *Dp8-23* is shown above the line. All chromosomes are shown. Symbolism is that of Figure 12. Derivative D261 is a 2-break event, removing most of the right arm.

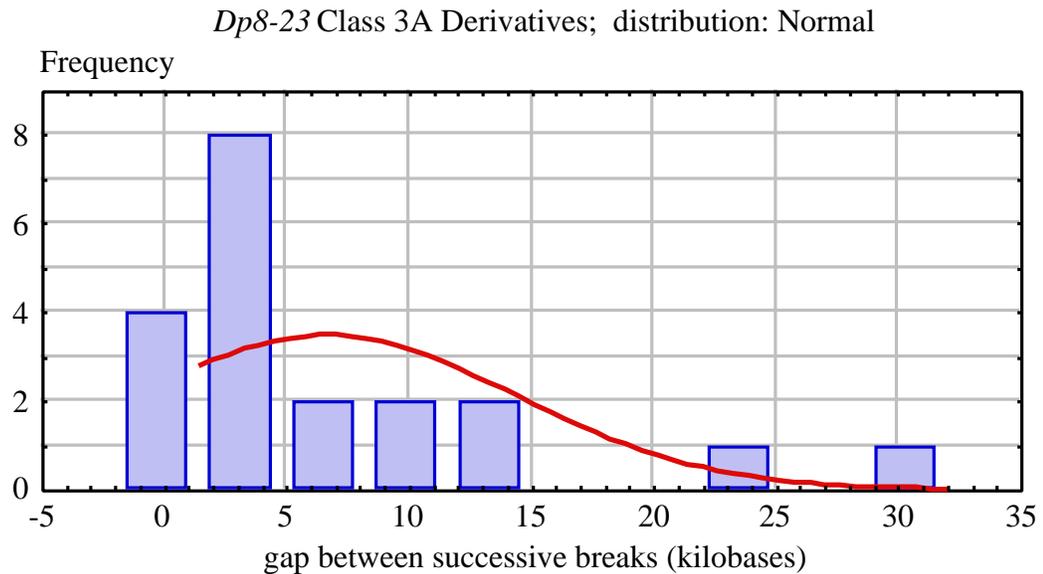


Figure 24: Normal Distribution of Breakpoints of *Dp8-23* Class 3A Derivatives
 Comparison of *Dp8-23* Class 3A derivatives with a normal distribution. See Figure 22 for a description of the methodology. χ^2 and Kolmogorov-Smirnov tests show $p > 0.05$.

Comparison of Class 3A Euchromatic Breakpoints of *Dp γ 238* and *Dp8-23*

I compare the breakpoints of *Dp γ 238* and *Dp8-23* Class 3A derivatives to address two questions. First, I analyze if the substrate chromosomes generate overlapping or identical distributions of breakpoints. Second, I assess whether the substrate chromosomes are equally mutable. Differences in either of these characteristics may be diagnostic of differences in context effect.

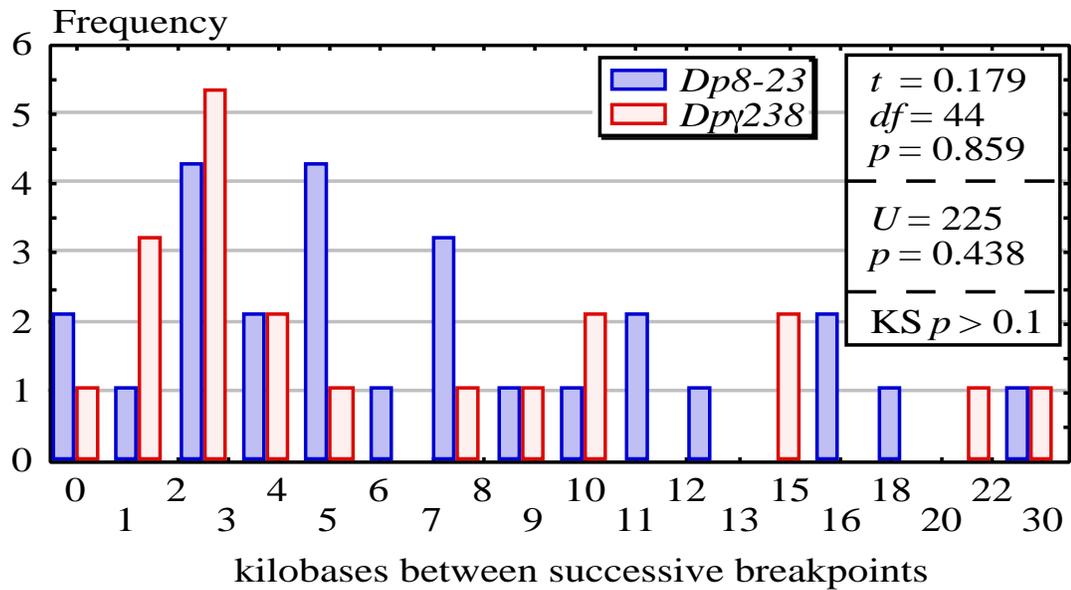


Figure 25: Comparison of $Dp\gamma238$ and $Dp8-23$ Class 3A Derivative Breakpoints

Comparison of $Dp\gamma238$ and $Dp8-23$ Class 3A derivatives and statistics. The distributions are statistically indistinguishable.

The breakpoint distributions of Class 3A derivatives from $Dp\gamma238$ and $Dp8-23$ were compared using Student's t -test. Since the error in breakpoint locations may confuse a model of normal distribution, I also calculated Kolmogorov–Smirnov and Mann–Whitney U statistics on breakpoints rounded to within 5 kilobases. All three tests gave approximately the same result: that the null hypothesis (H_0 ; the distributions

are identical) cannot be rejected ($\alpha = 0.05$). The data are plotted in Figure 25, along with the test statistics. I conclude the test segments of *Dp γ 238* and *Dp8-23* are uniformly mutable. The characteristics that allow breaks to be recovered within the euchromatic region do not differ from *Dp γ 238* to *Dp8-23*. Hence, potential telomeric regions, origins of replication, or other aspects that may limit potential broken ends of chromosomes are sufficiently frequent within the test segment to allow any sequence in these regions to behave as a chromosome end (provided the *mu2^c* background). Moreover, there is no context effect for chromosome ends that distinguishes *Dp γ 238* from *Dp8-23*.

I recovered 21 Class 3A derivatives in the *Dp γ 238* screen, and 21 Class 3A derivatives from the *Dp8-23* screen. I expected to recover Class 3A derivatives from *Dp γ 238* and *Dp8-23* at equal frequencies since there are no selective advantages or disadvantages on viability conferred by Class 3A derivatives from either of these chromosomes (Le, et al., 1995; Murphy and Karpen, 1995b). However, the observed difference in frequencies of breaks seems non-random (H_0 states that the frequency of derivative recovery is equal between *Dp8-23* and *Dp γ 238*, $X^2 = 35.5$, $df = 1$, $p = 2.55E-9$). The reason for the difference in mutability of these two chromosomes, despite the identity of the target sequence, is not clear. It is possible that these

substrates are packaged differently, perhaps due to their chromatin structure. Since no other study using *mu2c* gives resolution of breakpoints as high as this, it is impossible to say whether the irradiation and recovery of derivatives from *Dp8-23* is suppressed, or whether irradiation and recovery from *Dpγ238* is enhanced. Since *Tγ1337* places the test fragment in a euchromatic context, it is possible to compare frequencies of Class 3A derivatives from *Tγ1337* to the frequencies of *Dpγ238* and *Dp8-23* Class 3 derivatives to resolve this issue.

Structures of *Tγ1337 rosy* Derivatives

Twenty-eight derivatives were recovered after irradiation of *Tγ1337*. Of these, three were comparable to the Class 3A derivatives of *Dpγ238* and *Dp8-23*. Class 3A derivatives of *Tγ1337* were distinguished from Class 3B derivatives using a PCR-based assay to detect the *Tγ1337* breakpoint linked to chromosome 2 (see Appendix A). The frequency of breaks within the test fragment ($f = 1.2E-4$ derivatives/chromosome) is comparable to the break frequency of *Dp8-23* (H_0 states that *Dp8-23* generates Class 3A derivatives at the same rate as *Tγ1337*, $X^2 = 1.138$, $df = 1$, $p = 0.28$), suggesting that *Dpγ238* is enhanced for breaks within the test fragment (Table 3). The factors that contribute to breakability and recovery are not known.

Dpγ238 may be enhanced for breaks due to chromatin decompaction and higher accessibility to ionized intermediates (Ljungman, 1991; Warters and Lyons, 1992). This may be due to the characteristics of the chromatin, which in turn may be a function of centromere activity. It may also be enhanced for recovery due to decreased lethality/cell cycle arrest associated with terminal deficiencies, or it may be suppressed for repair. Although I cannot say why *Dpγ238* is a more mutable substrate than *Dp8-23* or *Tγ1337*, I suggest that its location or structure within the nucleus

Table 3: Frequencies of euchromatic breaks from *Dpγ238*, *Dp8-23*, and *Tγ1337*

Frequencies of comparable classes of euchromatic breaks (Class 3A, *yellow*⁺) recovered from three different substrate chromosomes. Frequency is expressed as derivatives per chromosome, and relative to the frequency of *Tγ1337*.

Chromosome	frequency (abs)	frequency (rel)
<i>Dpγ238</i>	1.06×10^{-3}	9.2
<i>Dp8-23</i>	2.21×10^{-4}	1.9
<i>Tγ1337</i>	1.16×10^{-4}	1

differs from these other two substrates, and that difference may be a function of its neocentromeric character.

Since the centromere is removed from the test segment in *Tγ1337*, many of the derivatives recovered as *rosy*⁻ were expected to have breaks in 2L material, proximal

to the *Tγ1337* translocation breakpoint. Although these Class 3B derivatives are not comparable to derivatives generated from *Dpγ238* or *Dp8-23*, I can use their frequency to estimate the target size on 2L material of *Tγ1337*. Classification of these derivatives was confirmed by the presence of *Tγ1337,0* (Figure 19) using a PCR-based assay for the *Dp8-23*-linked *Tγ1337* breakpoint (Appendix B). The target in *Tγ1337,ry⁺* is bound by the *Tγ1337* breakpoint distally, and limited proximally by a haploinsufficient region. If I assume that the frequency of breaks is constant along the *Tγ1337* 2L arm, and on either side of the *Tγ1337* breakpoint, then the frequency of breaks proximal to the *Tγ1337* breakpoint should also be 5.9E-7 derivatives/kilobase. Solving for target size, I expect to have had breaks in about 1.3 megabases of target. I expect to find Class 3B derivative breakpoints distributed throughout this region, limited by segmental aneuploid regions (Lindsley, et al., 1972), but the determination of these breakpoints is beyond the scope of this thesis work.

The Class 3B derivatives of *Tγ1337* are important for interpretation of the irradiation of *Tγ1337*. The reciprocal products of the Class 3B derivatives, Classes 4A and 4B chromosomes, must be neocentromeric, given that the centromere of *Tγ1337* is over 20 megabases from the *Tγ1337, ry⁺* test fragment. Since I have at most 1.3 megabases of target chromatin on *Tγ1337*, I might expect to recover neocentromere-

containing derivatives that are up to 1.3 megabases in size. This possibility will be discussed during the analysis of structurally acentric neocentromere derivatives.

Data from Table 2 can be used to assess the viability of segmental aneuploids for the region distal to the *Tγ1337* breakpoint. As predicted from Table 1, the data from Table 2 show that more Class 2 derivatives were seen relative to Class 1 derivatives. Since *SA Ders* come from Class 4A, which is subject to segmental haploinsufficiency, and Class 3A are subject only to segmental triplosubviability, for every three derivatives in Class 3A, I expect five derivatives in Class 4A. Although the complementary classes (3A and 4A) from *Dpγ238* and *Dp8-23* are directly comparable, Class 3A should be 1.6 times more frequent in *Tγ1337* derivatives than Class 4A.

Although larger *rosy*⁺ derivatives of *Dpγ238* and *Dp8-23* may show neocentromere activation, only those that remove all centromeric chromatin may be scored unequivocally as neocentromere-containing. This criterion limits the potential size of these Class 3A derivatives to less than 290 kilobases. Irradiation of *Tγ1337* removes this maximal size limitation by placing centromere-related sequence very far away from potential breakpoints. Any *rosy*⁺ derivatives originating from irradiation of *Tγ1337* should show neocentromere activation. Given the calculated target size for

breaks resulting in derivatives of *Tγ1337*, I expected that any inherent centromeric activity of the test segment would result in recovery of very many more neocentromere-containing derivatives from *Tγ1337* than either *Dpγ238* or *Dp8-23*.

Structures of *Dp8-23* and *Tγ1337 rosy*⁺ Derivatives

Two derivatives of *Dp8-23* were *rosy*⁺ and represented potential Class 4A neocentromere-containing derivatives. The existence of these derivatives would support Model 1, that an innate centromere was located on the test segment. Both *Dp8-23* derivatives came from the same set of parents, and are likely derived from the same breakage event in premeiotic S-phase of a germ cell. When these derivatives were isolated and run on PFGE, it was clear that they were larger than the 290 kilobase test fragment, indicating that they were internal deficiencies (2-break events) and were not Class 4A. Their breakpoints not relevant to this study, so they were not analyzed further.

Two derivatives of *Tγ1337* were recovered as *rosy*⁺, and may also belong to Class 4A or 4B derivatives. The *Curly yellow*⁺ *rosy*⁺ females (Figure 18B, F1) were outcrossed to *y*¹; *ry*⁵⁰⁶ males in order to determine the genetic stability of the potentially free *rosy*⁺ fragment chromosome. All offspring of this cross were of two

phenotypic classes: *yellow⁻ rosy⁻* or *Curly yellow⁺ rosy⁺*. This observation showed that these derivatives were both translocations with the *CyO, y⁺* chromosome. Since these derivatives were not comparable to the 1-break derivatives in this study, they were not analyzed further.

Structures of *Dpγ238 rosy⁺* Derivatives

Irradiation of *Dpγ238* is known to break that chromosome between the *rosy⁺* and *yellow⁺* genes (Murphy and Karpen, 1995b). Breaks in the *Dpγ238* chromosome can be observed and recovered as phenotypically *yellow⁺* or *rosy⁺* chromosomes. It is from this latter class of chromosomes that neocentromere activation events are recovered.

rosy⁺ derivatives of *Dpγ238* fall into 3 classes, based on the breakpoint relative to the centromere (Figure 20). Class 4A chromosomes have breaks that lie on the left arm, between *rosy⁺* and the centromere. It is this class of chromosome that has neocentromere activity. Class 4B chromosomes have breaks within the centromere, which has been shown to compromise transmission (Murphy and Karpen, 1995b). Lastly, Class 4C chromosomes have breaks within the heterochromatin between the centromere and the *yellow⁺* gene. All derivatives were analyzed for structure using

PFGE. Class 4A chromosomes were rigorously defined by the absence of the *Dp γ 238* breakpoint, detected through Southern- and PCR-based analyses. Although chromosomes that retained some centromeric chromatin were not obviously different in transmission from completely structurally acentric chromosomes, Class 4A chromosomes were limited to those that were truly devoid of all centromeric chromatin.

The structures of Class 4 derivatives of *Dp γ 238* are shown in Figure 26. Those in (A) are of Class 4A, those in (B) are of Class 4B, and those in (C) are of Class 4C. There were six derivatives that lacked all centromeric chromatin (Class 4A), making a break frequency in the euchromatic arm of $f = 2.28\text{E-}4$ derivatives/chromosome. Since these breaks cluster within the 70 kilobases near the euchromatin/centromere border, a frequency corrected for size of target chromatin is $f = 3.25\text{E-}6$ derivatives/kilobase. Breaks within the centromere (Class 4B) had a frequency of $f = 1.17\text{E-}6$ derivatives/kilobase, and a distribution that is randomly distributed (Figure 27). Breaks within the heterochromatin (Class 4C) had a frequency of break of $f = 2.99\text{E-}6$ derivatives/kilobase, and a distribution that is random. However, optimizing correspondence to a normal distribution indicates that there are two subgroups within Class 4C: those that lie proximal to the *Dp1187* breakpoint ($f = 1.75\text{E-}6$ derivatives/kilobase), and those that lie distal to this breakpoint ($f = 3.64\text{E-}6$

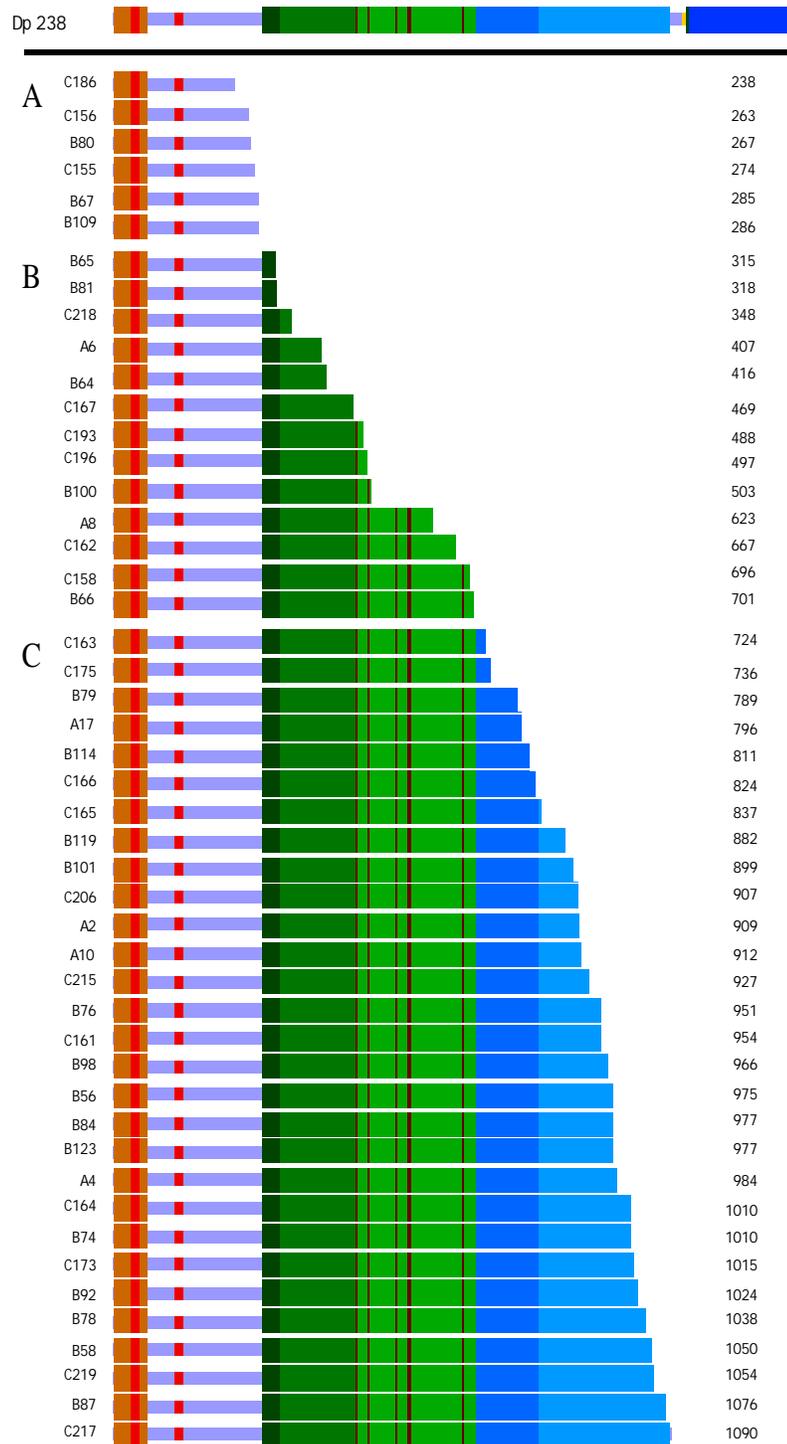


Figure 26: Distribution of Class 4 (*rosy*⁺) Breaks of *Dp*γ238

Unbroken *Dp*γ238 is shown above the line. Chromosomes that are *rosy*⁺ derivatives of *Dp*γ238. (A) Class 4A, broken within the euchromatin, (B) Class 4B, broken within the centromere, and (C) Class 4C, broken within the non-centromeric heterochromatin. Symbolism is that of Figure 21.

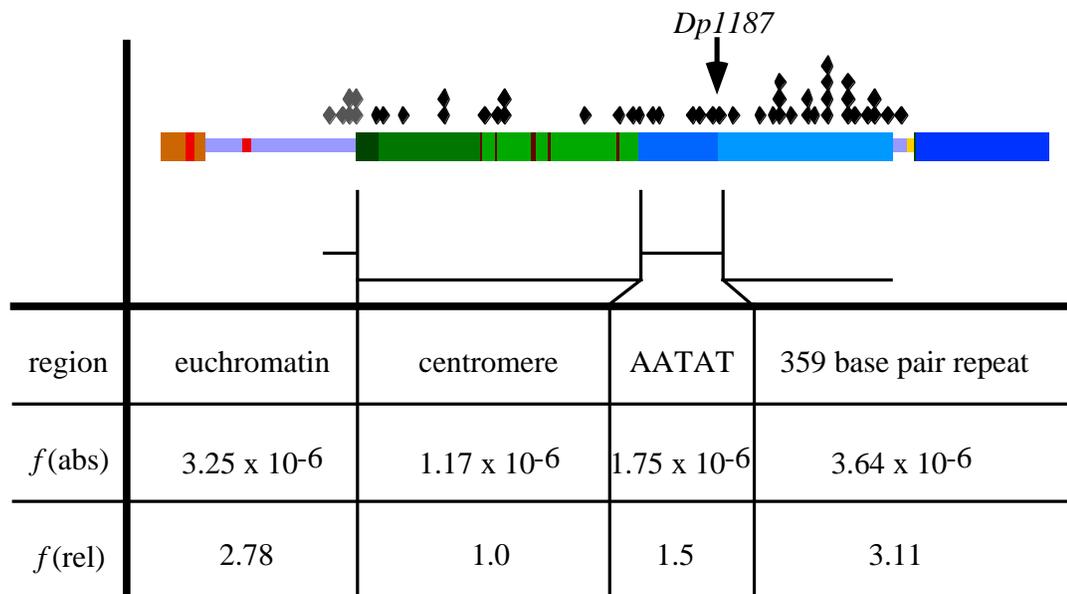


Figure 27: Mutability of Chromatin Types

Analysis of breakpoints along *Dp*γ238. *Dp*γ238 can be divided into four regions (vertical lines) by divisions between euchromatin, centromeric chromatin, and two types of heterochromatin (distinguished by the *Dp*1187 breakpoint). Each has distinct properties, as well as mutabilities. Symbolism is as in Figure 13. Diamonds represent breakpoints. *f* is reported in breaks/kilobase•chromosome and relative to breaks within the centromere.

derivatives/kilobase). These frequencies are reported graphically in Figure 27. Each of these populations alone more closely resembles normally distributed populations.

Analysis of Structurally Acentric Chromosomes

The generation of test fragments with neocentromere activity can be analyzed to assess whether the test segment is capable of manifesting neocentromere activity, despite context effect.

To compare the frequency of neocentromere generation, I analyzed the ability of *Dp8-23* to liberate neocentromere-containing Class 4A derivatives. As comparison, I used the known ability of *Dpγ238* to produce Class 4A derivatives. The frequencies were corrected for target size, and compared using a V-square test. The difference in frequencies of neocentromeric-fragment recovery between *Dpγ238* (6 per 28 complementary chromosomes) and *Dp8-23* (none per 21 complementary chromosomes) is not likely to have occurred by chance alone, arguing that the structure of the parent chromosome affects the ability of a liberated fragment to manifest neocentromere activity (H_0 states that *Dpγ238* and *Dp8-23* will generate neocentromeres at equal frequency; $V^2 = 4.084$, $df = 1$, $p = 0.0435$).

This observation is crucial to the test of my hypothesis. The history of a chromosome fragment is important in its ability to express centromeric activity. This

history cannot be explained by differences in nucleotide sequence, ruling out the possibility that a repressed centromere exists at the tip of the test fragment. History cannot be explained by differences in nucleotide sequence, ruling out the possibility that a repressed centromere exists on test segment. This history is coded onto the chromosome by some other stable mechanism, which is capable of existing in at least two states ("on" and "off" for centromere activity), and is regulated differently when an active centromere is juxtaposed or not. My data and observations support the hypothesis that neocentromere activity is regulated epigenetically – it is stably on or off, this change has no correlation to sequence changes, and suggest that it spreads from active centromeric chromatin to neighboring chromatin.

Supportive data were obtained from the irradiation of *Tγ1337*. Due to the structure and genetics of *Tγ1337*, a direct comparison of frequencies of *Tγ1337* to *Dpγ238* or *Dp8–23* is over-simplified. The number of potentially neocentromere-containing fragments from *Tγ1337* is affected by two factors: the target size and segmental aneuploidy associated with some segregation products of *Tγ1337*. As I will show in Chapter 5, a chromosome must be at least 200 kilobases in size to be recovered, with larger chromosomes increasing in stability. For the 290 kilobase test fragment, maximal fidelity (f_{mt}) is 0.9, and maximal transmission is 40% from males.

Given this, I conservatively estimate that about half of the neocentromeric derivatives generated from *Dp γ 238* and *Dp8-23* would be lost. This limitation does not constrain derivatives of *T γ 1337*, which has megabases of non-centromere euchromatin that could be linked to the test segment. Although Class 4A derivatives of *T γ 1337* should be no more frequent than Class 4A derivatives from *Dp γ 238* or *Dp8-23*, the Class 4B derivatives from *T γ 1337* should be generated at a much higher frequency. Because their sizes could be larger than 290 kilobases, their stability would be expected to be higher than 0.9. Hence, Class 4B derivatives of *T γ 1337* should be more frequent than Class 4A derivatives of *Dp γ 238* or *Dp8-23* not only because of target size, but also because of stability of resultant chromosomes. The fact that, without the size/stability constraint, no Class 4B derivatives were recovered from *T γ 1337* strongly argues that the test fragment in *T γ 1337* possesses no inherent centromeric activity.

The Class 3B derivatives of *T γ 1337* are qualitatively important in this screen. If neocentromeres could be activated and recovered from the *T γ 1337*-linked test segment, either as Class 4A or Class 4B derivatives, then they would be more stable than Class 4A derivatives from *Dp γ 238* or *Dp8-23*, and they would be 1.6 times more frequent than their complementary Class 3B (based on preferential viability of some *T γ 1337* segregation products, see Figure 16 and Table 1). Despite these two points,

none were recovered (H_0 states that the expected 5/25910 is identical to the observed 0/25910, $V^2 = 5.0$, $df = 1$, $p = 0.026$).

Analysis of Breaks within the *Dp γ 238* Test Fragment

I investigated whether the frequency of recovery of complementary breaks (Classes 3A and 4A from *Dp γ 238*) could be used to understand the probability with which a freed test fragment would manifest neocentromere activity. This bears on the understanding of the activation of neocentromeres, or whether that activity exists prior to γ -irradiation.

Since centromeres are required at exactly one copy per chromosome, I must conclude that *Dp γ 238* has one functional centromere prior to irradiation. This may mean that only one centromere (either the heterochromatin or the euchromatic test segment) is active prior to irradiation (Page, et al., 1995), or both regions are active, but cooperate (Page and Shaffer, 1998; Sullivan and Willard, 1998). Since this scheme of irradiation and recovery allows a class of chromosomes complementary to the structurally acentric fragments to be recovered, I can address this issue by examining the frequency of breaks in the test fragment that generate neocentromeres to breaks that do not.

This study produced six structurally acentric Class 4A fragments. Breaks in the identical sequence would be expected to give rise to chromosomes of Class 3A, the *yellow*⁺ derivatives. If the neocentromere on *Dpγ238* is active prior to the breakage of *Dpγ238*, then I would expect to see the frequency of Class 4A and complementary Class 3A *Dpγ238 yellow*⁺ derivatives be equal. If, however, the activation of the neocentromere occurs on a subset of test fragments, then I expect to see the number of *Dpγ238* Class 3A derivatives to exceed the number of recovered Class 4A derivatives.

The frequency of breaks is subject to the constraints of the behaviors of the chromosomes. Given that breaks in *Dpγ238*, giving rise to Class 3A derivatives, occurred with normal distribution out to the proximal P-element, then *Dpγ238* has 155 kilobases of effective target (−185 to −30).

For *Dpγ238* to generate neocentromere-containing Class 4A derivatives, there are different considerations concerning the stability of recovered derivatives. In this case, I never recovered derivatives smaller than 230 kilobases, similar to the finding of Murphy, who never recovered chromosomes smaller than 220 kilobases (Murphy and Karpen, 1995b). An explanation of this observation will be discussed in Chapter 5. For now, I will proceed with the correction that *Dpγ238* has 70 kilobases of effective target, from −100 to −30, for Class 4A derivatives.

Twenty-eight *yellow*⁺ derivatives were recovered from *Dpγ238*. If the target were limited to the 70 kilobases of target of Class 4A derivatives, I expect 12 derivatives at this frequency. Additionally, smaller Class 4A derivatives are mitotically and meiotically unstable. At most, they transit through each cell division only 90% of the time. Given that irradiation damage occurs prior to meiosis, and generated fragments must transit through two cell divisions before being passed to offspring, I can correct this expectation further. The correction for 10% loss through three divisions makes the expectation drop from 12 chromosomes to 8. This is in good agreement with the observed frequency (H_0 states that the corrected number of Class 3A derivatives, 12/28, is the same as the corrected number of Class 4A derivatives, 8/28, $V^2 = 0.583$, $df = 1$, $p = 0.44$), and I conclude that if neocentromere activity does not exist on the test fragment prior to liberation, it is relatively efficient at manifestation after liberation.

Although I cannot state that the neocentromere activity exists on the test fragment prior to liberation, I can say that the manifestation of neocentromere activity is rapid and efficient subsequent to liberation. The simplest model is that it exists before the chromosome is broken, and separation allows it to behave independently from the endogenous *Dpγ238* centromere.

Franz Schrader observed that a centromere split into two halves retained

inheritance activity on both products (p. v) (Schrader, 1939). The activity on these fragments manifested itself immediately, similar to what I have observed for the breakage products of *Dpγ238*. Centromeric activity seems to exist on the test segment of *Dpγ238* prior to irradiation and liberation. This possibility supports a hypothesis that some characteristics of *Dpγ238*, and not of the resultant test fragment, is important in eliciting neocentromeric activation.

Summary

The data presented in Table 2 can now be revisited, placing each derivative within its appropriate class. These data are presented in Table 4.

Table 4: Chromosome Derivatives: Numbers and Classes

	<i>Dp</i> γ 238	<i>Dp</i> 8-23	<i>T</i> γ 1337
total chromosomes (Class 1)	26,365	94,968	
<i>yellow</i> ⁺ <i>rosy</i> ⁻ (Class 3A)	28	21	
<i>yellow</i> ⁺ <i>rosy</i> ⁻ (Class 4A)	6	0	
<i>yellow</i> ⁺ <i>rosy</i> ⁻ (Class 4B)	13		
<i>yellow</i> ⁻ <i>rosy</i> ⁺ (Class 4C)	30		
<i>yellow</i> ⁻ <i>rosy</i> ⁺ (Class 1)			25,910
<i>yellow</i> ⁺ <i>rosy</i> ⁻ (Class 2)			32,930
<i>yellow</i> ⁻ <i>rosy</i> ⁻ (Class 3A)			3
<i>yellow</i> ⁻ <i>rosy</i> ⁻ (Class 3B)			25
<i>yellow</i> ⁺ <i>rosy</i> ⁺ (Class 4A)			0
<i>yellow</i> ⁺ <i>rosy</i> ⁺ (Class 4B)			0

CHAPTER FIVE

Genetic Analyses of Minichromosome Derivatives

Genetic Analyses of Minichromosome Derivatives

Transmission Tests: methods and results

The transmission of each derivative was determined in order to give an approximation of the fidelity of chromosome inheritance, and thus centromere activity (Murphy and Karpen, 1995b). In a transmission test, a chromosome that is transmitted to progeny must pass through 19 divisions, of six types. Specifically, a chromosome must pass from a parent to an offspring through nine syncytial divisions, no fewer than three pole cell divisions, a germline stem-cell division, four premeiotic mitoses, meiosis I, and meiosis II, before being included in a gamete (Ashburner, 1989). Following fertilization, the chromosome is transmitted through twelve syncytial and an number of postblastoderm divisions if recovered in adult somatic tissue (Figure 28). These categories of divisions are almost certainly a simplification; differences may exist between the first and subsequent syncytial divisions (Demerec, 1994), between syncytial and postblastodermal divisions (Sullivan, et al., 1993), between pole cell and somatic postblastodermal divisions (Counce, 1963), and between meiosis-I and meiosis-II (Koehler, et al., 1996).

In practice, transmission tests are performed by outcrossing a fly bearing a monosomic derivative to flies that do not. The offspring are scored for presence or

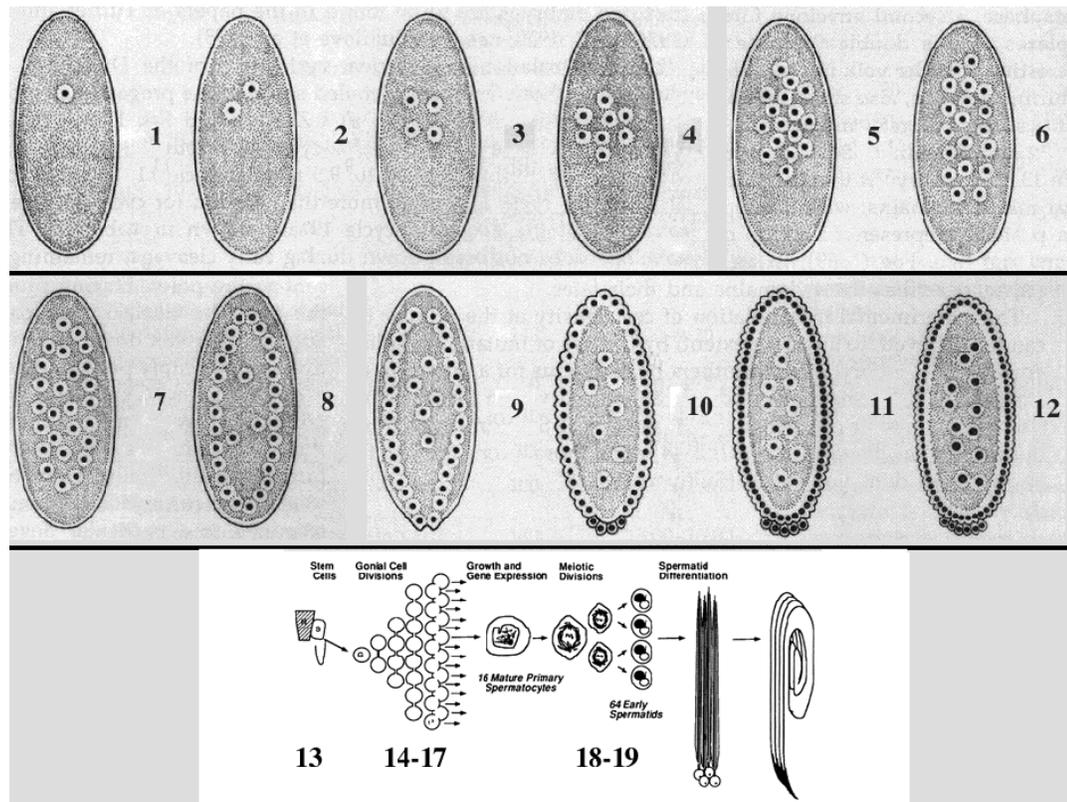


Figure 28: Types of Divisions

There are approximately 19 divisions from sperm to sperm, and a chromosome must transit through all of them in order to be inherited properly. Figure adapted from Ashburner (Ashburner, 1989).

absence of the derivative, usually phenotypically. A monosomic chromosome should be transmitted to half of the progeny. In my assay, derivative-bearing individuals are outcrossed to $y^1; ry^{506}$ mates, and the number of derivative-bearing offspring expressed

as a percentage of the total number of progeny produced. Normal transmission which was exhibited by the parental chromosomes is 50%, while departures from this (always lower) reflect instability of the derivative.

It is useful to examine the relationship between fidelity and transmission.

Fidelity is used to describe the faithful inheritance of a chromosome from one mother cell to two daughter cells. This is a single-cell division phenomenon, and can be categorized as either mitotic fidelity (f_{mt}) or meiotic fidelity (f_{me}). Fidelity differs from transmission, which considers an entire organismal generation. In protozoa, the terms are equivalent, but in metazoa (including *Drosophila*) they are not. Numerically, transmission is always lesser than or equal to fidelity, and they do not equate linearly. These terms are numerical and their use makes no predictions or assumptions about how or when a chromosome is lost or maintained, only the frequency with which a chromosome successfully transits from mother to daughter. High f_{mt} , f_{me} , or transmission requires that chromosomes be replicated, condensed, packaged, separated, and moved appropriately; dysfunction in any of these aspects may have an indistinguishable end-result: chromosome loss.

In the case of an organism with 19 cell divisions, and if the mitotic and meiotic fidelities are constant throughout the organism's lifecycle, then transmission and fidelity generally equate as:

$$\text{Eqn (1)} \quad \mathbf{T} = \mathbf{f}^g$$

or, more precisely,

$$\text{Eqn (2)} \quad \mathbf{T} = \mathbf{f}_{\text{mt}}^{17} \cdot \mathbf{f}_{\text{me}}^2$$

where \mathbf{T} is transmission, g is the number of cell generations in an organism's life (taken as 19 in Eqn 2, with 17 being mitotic and 2 being meiotic), and \mathbf{f}_{mt} and \mathbf{f}_{me} are mitotic and meiotic fidelities, respectively (\mathbf{f} equals $\mathbf{f}_{\text{mt}} \cdot \mathbf{f}_{\text{me}}$). \mathbf{T} is given through a transmission test as the fraction of offspring with the derivative, and \mathbf{f} can readily be solved in Eqn (1). Data from Williams, *et al.* (Williams, et al., 1998), show that a difference in transmission between males and females can readily be explained by differences in \mathbf{f}_{me} , specifically, $\mathbf{f}_{\text{me}}^{\text{♀}}$ is much lower than $\mathbf{f}_{\text{me}}^{\text{♂}}$. Data from neocentromeric derivatives in that study (26C and 22A) show that $\mathbf{f}_{\text{me}}^{\text{♂}}$ may approximate \mathbf{f}_{mt} . For this reason, analyses of \mathbf{f} and \mathbf{T} are done in males, where instability in female meiosis is circumvented. Data collected for the derivatives generated in this study are shown in Appendix B.

Brooding Assays: methods and results

The assumption that f_{mt} is constant throughout all types of cell divisions is not based on experimental data and is no doubt an oversimplification. However, without direct observation, it is difficult to assay chromosome loss in any one division type in an organism with multiple cell division types. Since Eqn 1 is exponential, the problem is compounded when f differs greatly from unity (f is small), or when an organism has many divisions (g is large). There is one aspect of spermatogenesis that can be considered – the difference between sperm pools of young males and old males. These sperm pools differ in the number of stem cell divisions that their progenitor cells have undergone. If germline stem cell divisions occur every ten hours (Demerec, 1994), then sperm derived from males that are ten hours different in age will correspond to sperm pools derived from n and $n+1$ divisions. Very young males produce sperm approximately one day after eclosion, and chromosomes in these sperm have seen 19 divisions before final sperm differentiation (Demerec, 1994). Using this logic, we are able to test at least one division for a specific value of f_{mt} (here f_{mt}^{gsc} for f_{mt} of germline stem cell).

Mechanistically, the assay is done by collecting very young males and placing them with aged, receptive females. At each subsequent time point, the females are removed and allowed to lay their eggs, while the males are mated to a new set of

receptive females. In practice, a male may mate with twelve females in one day (data not shown). The resolution of the assay is maximized if the male is placed with an excess of females, but the assay has been done with three females for three days (Williams, et al., 1998), or with three females for one day. The results of brood assays done on two chromosomes generated from *Dpγ238* (22A and 26C) showed that frequency of brooding produces results that are indistinguishable, indicating that sperm pools are not diluted by older sperm pools, even at this low resolution.

In interpretation, the initial transmission (from the youngest males) will be maximal, and loss of the derivative in germ line stem cell divisions will be observed as a decay in transmission rates for each clade of offspring from successively older males. The decay in transmission may be plotted and subjected to logarithmic regression to determine decrease in transmission per ten hour period. The rate of loss is $1 - f_{mt}^{gsc}$.

This assay was done by Wahlstrom for a suite of minichromosomes, and the data are presented in Williams, *et al* (Williams, et al., 1998). I have added to these studies by confirming the values of f_{mt}^{gsc} of 26C and 22A (290 kilobase and 220 kilobase derivatives, respectively) at a higher (48-hour) resolution. My results for 26C are shown in Figure 29, and are in good agreement with those of Wahlstrom and Murphy (Williams, et al., 1998).

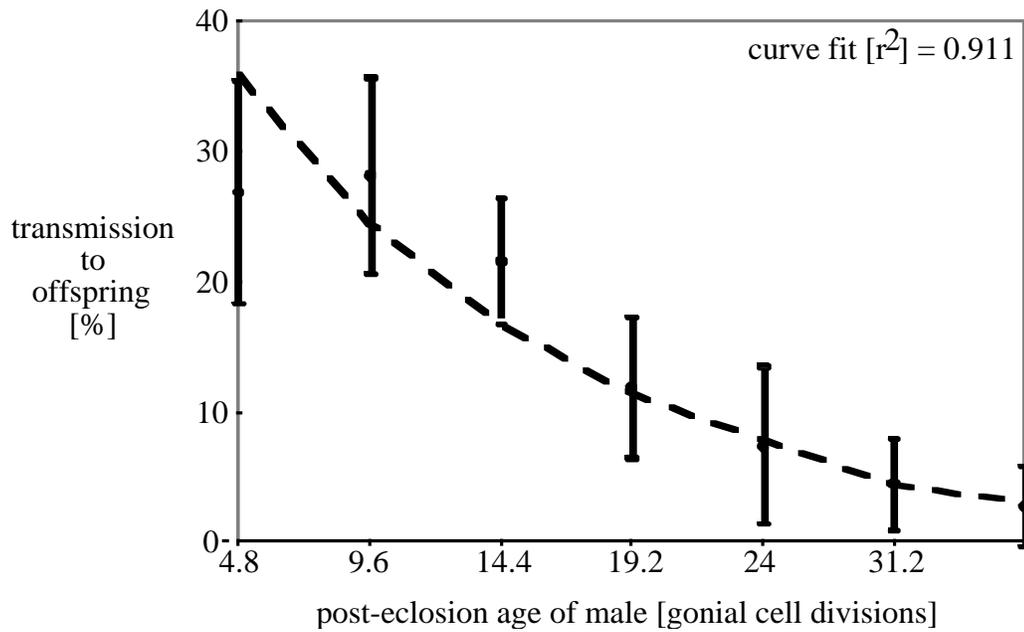


Figure 29: Brooding Analysis of Derivative 26C Loss

Loss of derivative 26C from each brood set. X-axis is in increments measuring from eclosion, and showing number of germ line stem cell divisions. Each datum is the mean of 9 individual tests (± 1 SD).

Acentrism and Size-Dependency: a treatment of chromosome loss

As described (Williams, et al., 1998), Murphy noted that f_{mt}^{gsc} is less than the solution for f in Eqn 1. Since $f^{26C} > f_{mt}^{gsc,26C}$, there must be divisions in the life cycle of the germ cells where f_{mt}^{26C} is higher than either $f_{mt}^{gsc,26C}$ or f^{26C} (as solved from Eqn 1). Solving for g in Eqn 1 with the value of f as f_{mt}^{gsc} and T as the observed transmission,

26C's instability may be explained as approximately nine divisions of 100% fidelity, followed by ten divisions at f_{mt}^{gsc} . It is compelling to draw parallels between the hypothetical nine stable divisions and about nine syncitial divisions in early embryogenesis, but this correlation has not been tested.

The fidelity f_{mt}^{gsc} is not the same for every chromosome tested. In fact, the value of f_{mt}^{gsc} value was size dependent, as the largest Class 4A derivative (26C) was lost at only 8% per cell division while the smallest (22A) was lost at 13% (Williams, et al., 1998). Here, mitotic fidelity (f_{mt}^{gsc}) follows the formula

$$Eqn (3) \quad \mathbf{f} = \mathbf{h} \cdot \ln \mathbf{L}$$

where \mathbf{L} is length in kilobases, and \mathbf{h} is a constant. \mathbf{h} can be solved using brooding assay data for the Class 4A derivatives, and has a numerical value of 0.16 ± 0.01 . \mathbf{h} is a good indicator, even for extrapolation, as \mathbf{h} applied to Class 4B derivatives of *Dp γ 238* are predictive of \mathbf{f} . By solving for \mathbf{L} with the requirement that $\mathbf{f} = 1.0$, it follows that a structurally acentric fragment of greater than ~500 kilobases would be expected to be fully stable. Since the largest Class 4A derivative from *Dp γ 238* can only be 290 kilobases in size without including centromeric heterochromatin, I cannot determine directly if this expectation is true. However, the smallest fully stable *rosy*⁺ derivative generated from *Dp γ 238* was ~700 kilobases, supportive of this

extrapolation. The contributions of size to chromosome stability will be discussed in Chapter 6.

Class 4B derivatives may differ from Class 4A derivatives in neocentromere activation, as well as size. Since Class 4B derivatives retain part of the *Dpy238* centromere, their transmission may be directed by centromeric DNA rather than neocentromeric DNA known to be required for Class 4A derivatives. However, Class 4B derivatives follow the trend described in Equation 3, suggesting that addition of *bona fide* centromeric sequence does not increase the stability of a chromosome over addition of euchromatin (Appendix B, Williams, et al, 1996.). As I will discuss in Chapter 6, the presence of a centromere is not different from the presence of a neocentromere. Beyond the activity of a centromeric/neocentromeric region, there is a size-dependency for stability (Murphy, 1998).

Selection for Stability

A centromere may be generally resistant to changes in its activity, but subject to stochastic or environmental alteration. Such alterations in centromeric activity would be selectable and heritable, if influenced by an epigenetic underpinning. This premise is based on observations in *S. pombe*, which have demonstrated that stable and unstable chromosomes may be selected from a common population.

Chromosomes with stable centromere activity, once isolated, remain stable; centromeres with unstable centromere activity give rise to populations containing unstable centromere activity.

I devised two experiments to test for selectable heritable changes in *Drosophila* neocentromere-containing derivatives. The first assay relied on selection of slightly more- or less-stable chromosomes, based on transmission test data. Chromosomes that exhibited higher-than-mean or lower-than-mean transmissions were used for subsequent generations of transmission testing. I chose conservative changes in transmission, excluding individuals that had transmission much lower or higher than average. Extremely unstable chromosomes would be lost, limiting the ability to select for very low stability, and would not represent a good model for lowered centromere activity. Alternatively, nondisjunction could generate disomic individuals that would exhibit high transmission, even though this high transmission was not caused by high fidelity.

Derivative 26C was transmission tested for four consecutive generations, and representative data are shown in Figure 30. At each generation, chromosomes selected as either "high" or "low" transmitting-chromosomes recapitulated the entire distribution. This suggests that there are not subpopulations within clades of neocentromeric derivatives that have higher-than-average or lower-than-average

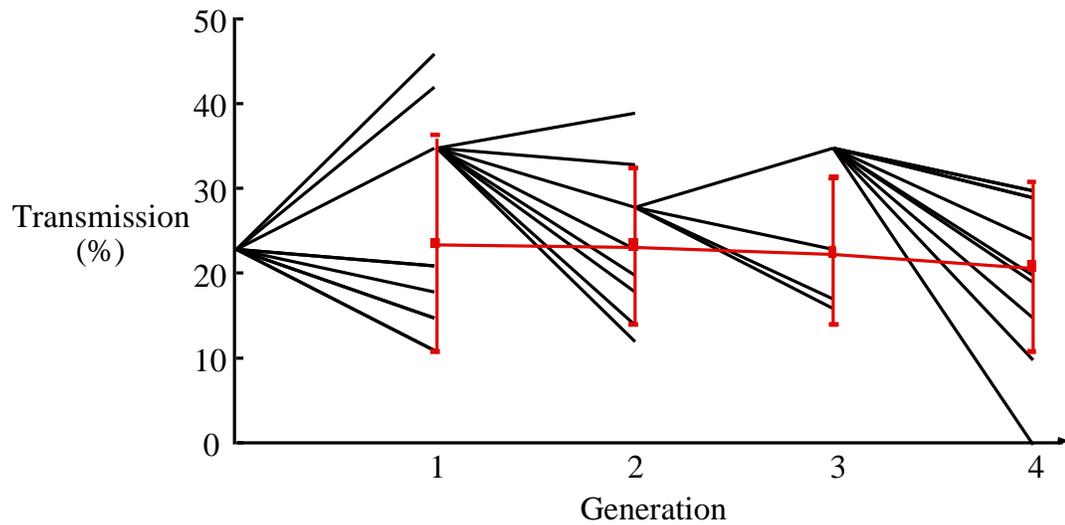


Figure 30: Test for *in vialo* Evolution of 26C Stability

Each generation of males produces a population of transmission values to its offspring. This is represented by the fanning of the lines, from a single male to many females with varied transmissions. At each generation, an individual with higher-than-average transmission was selected to start the subsequent transmission test. The red points are the mean of the population, and the bars represent one standard deviation. The mean remains constant each generation.

heritable stabilities. Minichromosomes *J21A* and *22A* were also subjected to this test, and revealed results consistent with those of *26C*.

The second experiment compared the coefficient of variation (C.V.) of transmission tests from an individual versus the C.V. of a population of individuals. The logic behind such an analysis is that the stability in fidelity would be slow to

change, and so a single chromosome would give rise to chromosomes with stabilities more similar to each other than would a disparate population of chromosomes. If a chromosome's transmission is dictated by its fidelity, and the fidelity changes slowly one generation to the next, then the C.V. of an individual transmission tested male should be different from the C.V. of a population of transmission tested males. Specifically, a derivative transmitted through a male to many females should have a lower C.V. than derivatives transmitted from a heterogeneous population of males to many females.

Coefficients of variance were tested for homoscedasticity using the Bartlett X^2 .

Data from this analysis are shown in Table 5. It is evident that the C.V. of transmission tests from a single male is not lower than the C.V. of transmission tests from a group of ten males. The difference in C.V. does not appear to be sufficiently improbable to preclude chance occurrence, and I cannot reject the null hypothesis of homoscedasticity.

Table 5: Comparison of Coefficients of Variance from Transmission Tests

Data from individual versus groups transmission tests. In the case of the 'individual' test, 10 males bearing 26C were outcrossed to 10 virgins, and the transmission to each was calculated. In the 'grouped' test, a single male bearing 26C was outcrossed to 10 virgins, and the transmissions to each was calculated. std dev is standard deviation, and C.V. is the coefficient of variance. Statistic reported is Bartlett X^2 .

	mean (%)	std dev (%)	C.V.	
individual	38.4	10.9	0.284	$X^2 = 0.0775$
grouped	31.8	8.5	0.267	$p = 0.78$

Summary

The neocentromeres encoded on the liberated test fragments bears the hallmarks of a fully-active centromere, yet the chromosomes are observed to be unstable. Their loss rates are higher than the loss rates for endogenous chromosomes, likely due to their size. The centromeric DNA, then, encompasses a locus as small as 220 kilobases. Additional chromatin is required to effect faithful inheritance, and I suggest that size is required either to align the chromosome at the metaphase plate during prometaphase congression or to stabilize its location on the metaphase plate prior to anaphase.

CHAPTER SIX

Synthesis and Discussion

Synthesis and Discussion

In this study, I have demonstrated that the activation of neocentromeric activity on structurally acentric chromosomes requires the proximity of an active centromere. I have shown that I can induce breaks in three minichromosome derivatives of the *X* chromosome tip, and that these breaks led to the liberation of identical fragments from three different chromosomal contexts. These fragments, although identical in sequence, behave differently in their ability to express centromeric activity.

In organisms with regional centromeres, large blocks of heterochromatin separate the centromere from the euchromatic arms. This situation is retained in *Dp8-23*. In this derivative, a test segment which consists of 290 kilobases of euchromatin and subtelomeric heterochromatin is separated from the centromere by 380 kilobases of centric heterochromatin. The situation is similar in the translocation derivative, *T(2;Dp)γ1337*, where the identical test segment is separated from the endogenous chromosome 2 centromere by normal chromosome 2 arm (~35 megabases of euchromatin and heterochromatin) (Adams, et al., 2000). Irradiation of either of these substrates fails to release the 290 kilobase fragment with neocentromeric activity, indicating that this fragment does not possess inherent centromeric activity.

In contrast the 290 kilobase test segment is directly juxtaposed to a functional

centromere in *Dpγ238*. Once irradiated, this chromosome gives rise to stably transmitted neocentromere-containing fragments. Despite the fact that the liberated chromosome tips of all of these substrates are identical in sequence, I can reject the null hypothesis that states that identical substrates will manifest neocentromeric activity with equal frequency ($X^2 = 10.5$, $df = 2$, $p = 5.29E-3$). I propose that something must mark chromatin that is capable of manifesting neocentromere activity, distinguishing it from chromatin that cannot. This mark differs between *Dpγ238* and its close relatives *Dp8-23* and *Tγ1337*. I propose that the difference is a function of proximity to an endogenous, functional centromere.

Two salient features define centromere activity as I have investigated it. First, a fully-functional centromere must be capable of nucleating a kinetochore. As such, centromeres have a "sufficiency of mark," and once marked they are self-propagating. Second, a chromosome must have a "sufficiency of size." Whether size affects kinetochore activity or others aspects of chromosome stability cannot be distinguished by my experiments. These two features, as I will now discuss, are sufficient for a chromosome to possess a centromere.

Centromeres and Neocentromeres: Sufficiency of Mark

Centromeres and neocentromeres differ in the sequence that they overlie

(Barry, et al., 1999), and different centromeres may also overlies different sequences (Sun, et al., 1997; Talbert and Henikoff, 2000). Clearly, then, sequence cannot be used to distinguish between centromeres and neocentromeres. If there is a difference between a centromere and a neocentromere, then some difference may be obvious beyond their location within the genome.

My results show that once some mark is acquired by a region of chromatin, it becomes sufficient to direct chromosome inheritance for thousands of cell divisions. This epigenetic marking must be both stable and plastic. Neocentromeres are rare, and limited in their occurrence. Centromeres are also not seen to arise or disappear at a high rate, nor are they seen to migrate during the life of either an organism or of a species (White, 1954). Once active, both centromeres and neocentromeres are fully capable of nucleating kinetochores and propagating their activities to each daughter chromatid.

The nature of this mark is not known. Recently, work in *Drosophila* has identified a homologue to a known centromere protein in humans (Henikoff, et al., 2000). The gene product of the *Cid* locus is homologous to CENP-A, which is strongly correlated with centromere activity in mammals (Palmer, et al., 1987; Palmer, et al., 1989; Palmer, et al., 1991; Vafa and Sullivan, 1997; Warburton, et al., 1997; Vafa, et al., 1999). CENP-A is a histone-H3 variant, and whether it may actually

identify chromatin as centromerically active or simply be targeted to it is not known. Additionally, it may remain associated with DNA through replication (Kass and Wolffe, 1996), and it may preferentially recruit new CENP-A molecules to DNA through protein-protein interactions (Shelby, et al., 1997). Investigation of these last two points is critical, as proof of these two activities is sufficient for explaining epigenetic inheritance. CENP-A found on both centromeres and neocentromeres in humans, and in fact 19 mammalian centromere-specific antigens are found at both centromeres and neocentromeres (Saffery, et al., 2000b). In *Drosophila*, the gene product of the *Cid* locus is a CENP-A-like molecule, and associates with centromeres and neocentromeres alike (Blower and Karpen, in preparation). Other factors known to be assembled into kinetochores are also present on *Drosophila* neocentromeres, including *l(1)zw10* (Williams, et al., 1998), *polo*, and others (Blower and Karpen, in preparation). These observations suggest that neocentromeres and centromeres utilize the same kinetochore machinery for attachment to spindles to direct their movement during cell divisions. Additionally, they may utilize the same factors for identifying DNA that is capable of nucleating a kinetochore.

Based on associated factors, it is difficult to distinguish a neocentromere from a centromere. However, my experiments show that they do differ in at least one respect, stability.

Centromeres and Neocentromeres: Sufficiency of Size

On the surface, the genetic results that I have presented suggest that the neocentromere is a poorer cousin to an endogenous centromere. This is reasonable considering the arguments for chromosome mass contributing to the strength of a centromere (Novitski, 1951, 1955; Ptashne, 1960). The neocentromere, devoid of centric heterochromatin, may be expected to be exceedingly weak or inactive. To the point, very small chromosomes with neocentromeric activity exhibit transmission and fidelity lower than those of endogenous chromosomes, or even than those of their parent chromosome *Dpγ238*, despite the fact that they have an active neocentromere. This instability could be a function of the neocentromeric character of their inheritance, revealing some difference between neocentromeres and centromeres. Perhaps neocentromeres are less faithful than *bona fide* centromeres. Alternatively, some other aspect of chromosome structure in neocentromere-containing chromosomes may be responsible for their compromised stability. Neocentromeres in humans do not show meiotic or mitotic instability (Tyler-Smith, et al., 1999), yet brooding analyses of neocentromeres in *Drosophila* suggest that most (or all) instability of these chromosomes is attributable to mitotic loss.

Work done by Wahlstrom and Murphy (Williams, et al., 1998), and Lopez (Lopez, et al., submitted), offer an explanation for chromosome instability in

neocentromere-bearing derivatives. In the first study, neocentromeric derivatives were analyzed for mitotic loss using brooding assays, as I have reported again here (Chapter 5). The results of those analyses, as well as mine, show that loss of these derivatives can be explained by a steady loss through about half of the divisions in an adult fly. Moreover, f_{me}^{σ} seems to be no worse than f_{mt} , and f_{me} can be treated as f_{mt} for neocentromere-containing chromosomes in males. This observation is bolstered by a genetic response to *nod* (Murphy and Karpen, 1995a), a chromokinesin responsible for generating antipoleward forces on non-exchange chromosomes (Zhang and Hawley, 1990), including the derivatives in this study. In females, this gene product is necessary for spindle-chromosome stability. *nod* interacts with chromosomes along their entire length, and not just at the centromere (Afshar, et al., 1995; Murphy and Karpen, 1995a). Hence, the instability seen in female divisions can be attributed to dysfunction in antipoleward forces. Antipoleward forces do not emanate from the centromere (Ault, et al., 1991; Wise, et al., 1991; Liang, et al., 1994; Rieder and Salmon, 1994), and so are not a function of centromere identity or kinetochore activity, but rather the length of the chromosome.

The exaggerated loss of neocentromere-containing derivatives in females can be linked to meiosis. Lopez and co-workers have observed that these derivatives bind the gene product of the *mei-S322* locus in primary oocytes at a frequency about equal

to their transmission from females (Lopez, et al., submitted). This can be interpreted as the loss of these chromosomes occurring because lack of *mei-S322* binding. *mei-S322* mutants show no defect in mitotic chromosome loss, and neocentromeric derivatives are observed in female cells prior to meiosis at a frequency higher than transmission after meiosis, indicating that loss of neocentromeric fragments in females occurs mostly in meiosis (Lopez, et al., submitted). The gene product of *mei-S322* is involved in chromatid cohesion in meiosis, and not kinetochore nucleation or function (Moore, et al., 1998). The instability of neocentromeres in females, then, is almost entirely attributable to steady low rates of mitotic loss plus extremely high rates of meiotic loss. Both of these loss types are responsive to the length of the chromosome.

In both males and females, the instability observed in the behavior of neocentromeres can be attributable to a defect in chromosome size, and is therefore not likely to be due to defects in kinetochore activity.

If the function equating size and transmission is extrapolated to include predictions about larger chromosomes, I expect a chromosome of approximately one-half megabase to be "fully" functional and show 100% fidelity. Although the absence of neocentromere-containing fragments from the irradiation of *Ty1337* show unequivocally that size is not the only factor in neocentromere activation or recovery, I cannot state that the full function of a centromere is not conferred on any

neocentromere-containing chromosome of sufficient size. It would be interesting to induce a neocentromere on a fragment of size hypothesized to give full activity.

Transient juxtaposition of a segment of DNA to the *Dpy238* centromere would presumably give neocentromere activation, and any liberated fragment of greater than about 500 kilobases would be expected to be fully stable.

Two factors are important in centromere identity: neocentromeric activation and sufficient chromosome size. I suggest that there is no mechanistic difference between a neocentromere on a sufficiently large stretch of chromatin, and a *bona fide* centromere. I propose that both are maintained using the same epigenetic system, and both assemble identical kinetochores, and ultimately both perform identical functions.

Commonalities Among Centromeres

The point centromeres of *S. cerevisiae* and the regional centromeres of many other eukaryotes may represent entirely different mechanisms for identity and maintenance. A strong sequence-dependency is clearly evident in conferring *S. cerevisiae* centromere identity (Hahnenberger, et al., 1991), while the lack of a sequence sufficiency/necessity for conferring regional centromere identity strongly implicates an epigenetic mechanism (Karpen and Allshire, 1997). While the lowest tier of identity (sequence of DNA) is not conserved in these types of centromeres, the

factors which play active roles in centromere function are clearly homologous (*e.g.* – *Cse4* and *CENP-A*) (Sullivan, et al., 1994; Stoler, et al., 1995; Meluh, et al., 1998).

Ultimately, centromere function will depend on the factors brought to the chromatin, and not the mechanism with which they are delivered (Dobie, et al., 1999). Epigenetic and genetic demarcation of centromere identity differ only in the method of recruitment, and not in the mechanism of centromere activity.

Similarly, the differences between holocentric behavior and discrete centromere loci may also represent grossly different modes of centromere regulation. But these differences may be rather minor. My data show that centromere activity may spread, and indeed is probably bound by some factor that prevents the centromere from spreading in a normal context in a monocentric organism. One would predict that the removal of a boundary is sufficient to cause a centromere to "leak" onto the euchromatin that neighbors it, as in *Dpy238*. If all boundaries were removed, one would predict that centromere activity would "leak" over the entire chromosome, stopped only at the ends of the chromosome. This case would be indistinguishable from a holocentric chromosome, and perhaps a holocentric chromosome merely demonstrates the deregulation of centromere boundaries, and not a novel form of centromere identity.

Activating and Deactivating Centromeres

My work has demonstrated that centromere function can spread *in cis* to juxtaposed DNA. It is clear that proximity to an active centromere greatly increases the frequency of neocentromere formation on substrate euchromatin, although the mechanism for this is unknown. Spreading of an epigenetic state may also occur *in trans*, as established for paramutation (Chandler, et al., 1996). Recently, spreading has also been demonstrated in *Drosophila* dosage compensation (Kelley, et al., 1999). Although interactions between chromosome regions and spreading *in trans* or *in cis* are possible, it is also possible that centromere activity can be acquired spontaneously. The rare appearance of neocentromeres precludes determination of the probability of centromere ON-to-OFF and OFF-to-ON rates without some contextual contribution (such as the juxtaposition to an active centromere in *Dpy238*). I will postulate a mechanism to explain centromere spreading in the next section. To do so, I will first introduce a model for explaining the sporadic activation and deactivation of centromeres.

I would like to revisit the energetics of epigenetics with an emphasis on centromere state identity. Figure 31 shows how the paradigm introduced in Figure 8 applies to the centromere. In this case, conditions 1, 2, and 3 may represent a normal state, a pharmacological treatment in *S. pombe*, or a null mutation for a centromere-

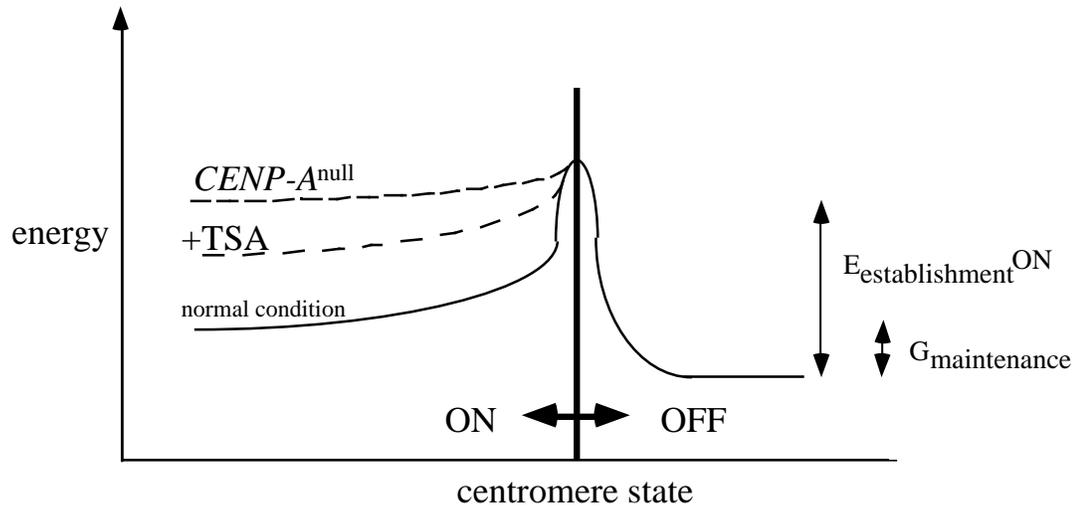


Figure 31: Energetics of Centromere Identity

See text for description.

identity factor in mammals. In each case, there is an equilibrium defined by the difference in energy levels between centromere ON and centromere OFF. In conditions 2 and 3, the centromere ON state is destabilized by dramatically increasing the amount of energy required to stay in that state. In those conditions, centromere OFF is favored, and most centromeres will be inactivated or fail to be assembled.

In all three conditions, an equilibrium exists between centromere ON and centromere OFF. As such, there is a flux of any locus between centromere ON and centromere OFF, and *vice versa*, even without environmental or genetic influence. In

cases where a centromere OFF improbably becomes centromere ON, a neocentromere is formed. In the improbable case where a centromere ON becomes centromere OFF, an extant centromere is deactivated. It is important to note that the kinetics of transition from centromere OFF to centromere ON, or ON to OFF, is defined by the difference in energies of the states (ΔG), as well as by the activation energy (E). For this reason, changes in state will occur with a defined frequency. In most cases, the centromere is either energetically unfavorable, or the alteration of a state is so energetically costly that neocentromeres arise or centromeres are lost only rarely. Additionally, since most of these instances of centromere activation or deactivation result in cellular lethality, they are maintained and witnessed even more rarely.

Chromosomes must suppress illicit centromere activation or deactivation. Even infrequent (energetically disfavored) events can be rendered effectively non-existent by requiring multiple rare events to coincide. Centromeres that are activated and stabilized may be a subset of OFF-to-ON events if centromere maturation requires additional steps. For instance, if one molecule of centromere identity factor were sufficient to cause a neocentromere to form, neocentromeres may be quite common. However, if multiple centromere identity factors are required to exceed a threshold, or more than one factor were required at the same chromosomal locus, then neocentromere activation would be dramatically reduced in frequency.

Additionally, some sequences are more capable of forming centromere-permissive structures than are others. In *S. pombe*, the *cen* loci are linked to centromere activity, and only DNA reintroduced into cells shows centromere activity. But there is a clear sequence-independent circumstance that moderates centromere activity once the DNA is introduced. Hence, I view the epigenetic regulation of the centromere as a continuum. At one extreme is the situation exemplified by *S. cerevisiae*, where sequence alone is necessary and sufficient to confer full mitotic and meiotic centromere activity. At the other extreme would be a strictly-epigenetic phenomenon. Whether such a condition exists in nature is not known, but what is certainly known is the contribution by epigenetic factors on relatively loose sequence requirements in *H. sapiens* and *D. melanogaster*. In these latter cases, the energy required to form a centromere on some sequences may be lower than the energy required for centromere activity to associate with other sequences, but centromere activity may associate with a broad range of sequences that share no homology.

An understanding of centromere energetics can be extended to explain spreading. In my experiments, the centromere is capable of moving to modify neighboring DNA. This spreading occurs only in *Dp γ 238*, where the centromere juxtaposes euchromatin. It does not occur in *Dp8-23*, where the centromere is 380 kilobases from euchromatin. Some activity of spreading allows a centromere to move

at least 70 kilobases (to generate the smallest neocentromere-containing chromosome 22A), but not 380 kilobases. An explanation is that a boundary exists between the centromere and test fragment in *Dp8-23*, but not in *Dpγ238*. Perhaps, then, it is the presence of heterochromatin in general that prevents spreading of centromeric activity from the centromere to outlying DNA. Heterochromatin may be refractory to centromere spreading in a way that DNA packaged as euchromatin is not.

This is supported by the lack of derivatives that do *not* contain the test fragment. Figure 27 shows that the heterochromatin between the centromere and the *yellow*⁺ gene is highly breakable. Despite this, no *yellow*⁺ derivatives were recovered that broke in that region (H_0 states that fragments with active neocentromeres on the heterochromatin to the right of the centromere can be recovered at the same frequency as fragments with neocentromeres active on the euchromatin to the left of the centromere, $X^2 = 11.42$, $df = 1$, $p = 7E-4$). The centromere activity spreads over euchromatin at least 70 kilobases, but not at all through the heterochromatin.

Although the presence of neocentromeres in euchromatin rules out a necessity for heterochromatin in centromere identity, centromeres are generally bound by large blocks of heterochromatin. Perhaps it is this placement that limits a centromere's ability to inappropriately modify neighboring DNA and confer centromeric activity. If centromere spreading were not limited, centromeres would be expected to grow or

move – two conditions not observed in nature.

Alternatively, the centromere may be limited in spreading by the presence of discrete centromere boundary elements, the structure of the simple satellite itself, or limiting concentrations of essential centromere proteins . In any case, the inversion *Dpγ238* defeated at least one side of this regulation, allowing centromere activity to spread onto the *Dpγ238* euchromatic arm.

Spreading can be explained energetically, as can limits to spreading. Figure 32 redraws Figure 31 in a pictorial way. Here, (A) shows how centromeres epigenetically maintain themselves. Each factor associates with similar factors, allowing centromere identity factors to recruit more centromere identity factors preferentially to the centromere. This occurs along an entire strand after replication. I have shown data that demonstrate this templating ability for centromere factors (*i.e.*, only the constant of association *in trans*, $k_{\text{CEN,trans}}$ need be non-zero). The ability of heterochromatic factors to possess this templating ability may be inferred from clonality of position effects (Wakimoto, 1998). The ability for euchromatic factors to template themselves after replication has not been demonstrated. In truth, these factors may be directed to DNA by sequence-specificity, or they may associate with chromatin that has not associated with other factors.

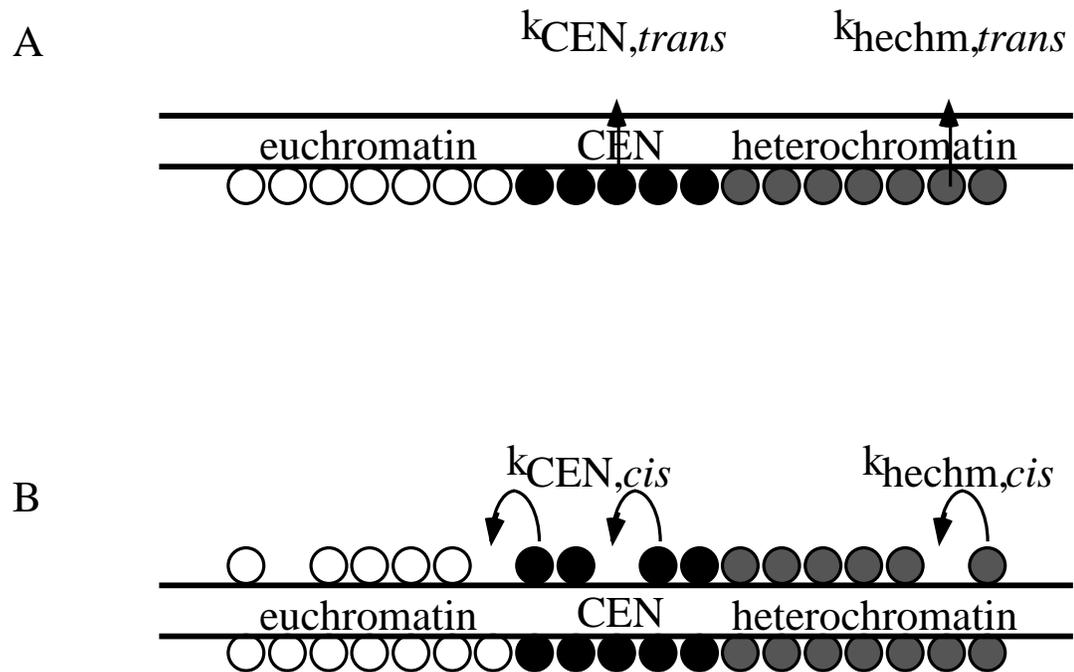


Figure 32: Energetic Explanation of Centromere Spreading

See text for description. Black circles represent centromere activity, as dictated by centromere-specific factors overlying centromeric DNA (CEN); heterochromatin-specific factors are shown as gray circles; open circles represent euchromatin-specific factors.

The model in (B) shows how this applies to spreading and boundaries. If each factor can associate like factors to nascent DNA strands, then each factor may also have a k_{cis} , or constant of association with factors *in cis*. For spreading to occur from a centromere to flanking euchromatin, then $k_{CEN,cis} > k_{euchm,cis}$. This predicts that at an

interface between a centromere and euchromatin, the centromere identity factors are more avidly recruited, and so tend to spread into euchromatin, perhaps even displacing euchromatin-associated factors. Since spreading does not occur through heterochromatin, then $k_{\text{hechm,trans}} > k_{\text{CEN,cis}}$. Here, heterochromatin factors are more avidly recruited by heterochromatic factors from the native strand, preventing the centromere from spreading *in cis*. This allows centromere activity to spread, but not through regions where heterochromatic structure is associated with either strand. In general, spreading *in trans* is more likely, favorable, or processive than spreading *in cis*, assuring that a chromatin state is conserved in location and extent. Only in instances where a chromatin state is removed or altered do neighboring chromatin states spread.

Although spreading *in cis* and spreading *in trans* may seem mechanistically distinct, they may be features of a common mechanism. For any epigenetic system to work, a state must be copied from a template strand to a nascent strand. Hence, spreading *in trans* must occur in epigenetic systems. The observation of spreading *in cis* in this study may represent a new mechanism, distinct from spreading *in trans*, or it may be a combination of spreading from template strand to nascent strand with some 'offset' in location of the templated mark. If this spreading occurs back, from the nascent strand back to a template strand, with a similar offset, spreading *in cis* will

have occurred along the template strand.

This model of chromatin structure determinants spreading *in cis* and *in trans* is supported by data from *S. pombe*, that shows that some centromere proteins may associate with euchromatic sequence when placed in a centromere (Partridge, et al., 2000). Other factors are not capable of impinging on this euchromatin, suggesting that each centromere protein within the centromere has a different **k**, which is a property not only of its activity, but of the DNA that it may bind to. In this way, some sequences may lend themselves to centromeric activity, while others may be incapable of forming centromeres, and yet centromere activity does not rely on sequence specificity.

These models predict a characteristic of some centromere-specific and heterochromatin-specific factors. Some of these factors will be preferentially recruited to chromatin by protein-protein interactions. Once in place, their activities will be stable and propagated during multiple rounds of DNA synthesis. This model also predicts that differences in the strengths of association of centromere-specific or heterochromatin-specific factors will affect the ability of centromere activity to spread.

Conclusion

The regional centromere is a largely elusive structure, due in part to sequence-

independent aspects of its activity and location within centric heterochromatin. I have presented data that show that at least some components are epigenetic in regulation, and amenable to investigation. Although all the factors responsible for coding centromere identity are not yet known, work from this thesis and work from many other labs defines some characteristics of centromere identity and activation. My data show that the neocentromere is a good model for understanding the factors that contribute to the function of a centromere. I have presented and supported models for the ability of centromeres to be epigenetically regulated, as well as spread from active to inactive DNA. These observations have led to a model for centromere maintenance and spreading that relies on energetic characteristics of centromere identity factors. Analysis of this model, and predictions that it makes, will help clarify how centromeres are determined and maintained on a chromosome, in an organism, and during evolution.

APPENDICES

Ancillary Information

Appendix A: Materials and Methodology

Transmission tests

Transmission rates of derivatives was determined by outcrossing single flies bearing derivatives to three flies of the opposite sex. Derivative-bearing flies were allowed to mate for three days at 25°C, were transferred to new vials and allowed to mate for three more days, then were discarded. Progeny were scored on days 14 and 18 after the vials were set. Transmission was recorded as percentage of offspring that received the derivative. Individual tests were excluded from analyses if the total number of progeny were fewer than 35 or the transmission was greater than 65% (as such a value probably indicated a nondisjunction event giving rise to a disomic individual).

Brooding Assays

Brooding assays were performed on freshly eclosed males. Individual males were given three to five aged and receptive females and allowed to mate at 25°C. Every twenty-four or forty-eight hours, the females were discarded and the males transferred to new vials with new females. Progeny were scored on days 14 and 18 after each vial was set.

Chromosome I Background

Stocks were of either rod-*X* or attached-*XY* ($Y^{\delta}X.Y^L$, *In(I)EN*, *y*; $X^{\wedge}Y$) background.

The *X* constitution was swapped as follows. $X^{\wedge}Y$ to *X*: virgins of genotype $X^{\wedge}Y/X^{\wedge}Y$;

Der were collected and outcrossed to males of genotype *X/Y*. Male progeny bearing

the *Der* were collected and outcrossed to *X/X* virgins for two generations to generate

X/X/Y; Der stocks. *X* to $X^{\wedge}Y$: virgins of genotype *X/X; Der* were collected and

outcrossed to males of genotype $X^{\wedge}Y/0$. Female progeny bearing the *Der* were

collected and outcrossed to males of genotype $X^{\wedge}Y/0$. Male offspring of that cross

containing the *Der* were collected and crossed to $X^{\wedge}Y/X^{\wedge}Y$ virgins to establish a stock

of genotype $X^{\wedge}Y/X^{\wedge}Y/0$; *Der*.

Plasmid DNA Preparation

Plasmid DNA of quality sufficient for cloning, restriction, sequencing or injection was prepared using the Maxed-Out-Miniprep (MOM). Fifty milliliter overnight cultures (in Luria Broth with antibiotic) were spun down at low speed (3000 g) and resuspended in 5 mL Solution I (1% glucose, 10 mM EDTA, 25 mM Tris pH 8.0). 10 mL Solution II (1 N NaOH, 1 % SDS, made fresh) were added at room temperature, inverted, and mixed with 7.5 mL solution III (3 M KOAc). The solution

was spun at low speed (3000 g) for 5 minutes, and filtered through a Kimwipe. Equal volumes 3-propanol was added and spun at low speed for 20 minutes. After resuspending the pellet in 800 μ L water, 4 μ L RNase A (20 mg/mL) was added and allowed to incubate at 37°C for 30 minutes. Next, 4 μ L Proteinase K (20 mg/mL), 4 μ L SDS (20%) and 40 μ L NaOAc (3 M) were added and allowed to incubate at 65°C for 30 minutes. Two rounds of phenol-chloroform extraction was followed by ethanol precipitation. Finally, the DNA was PEG precipitated by adding 0.6 volumes 20% PEG 3350/2 M NaCl, freezing, and centrifugation for 15 minutes. Pellets were resuspended in TE or water to a final concentration of 1 mg/mL.

Adult Fly DNA Preparation

DNA was extracted from adult flies or larvae by grinding flies in Extraction Buffer (100 mM NaCl, 200 mM sucrose, 100 mM Tris pH 9.1, 50 mM EDTA, 0.5% SDS), and incubating at 65°C for 30 minutes. DNA was precipitated by adding 60 μ L KOAc (5 M) and incubating on ice for 30 minutes, followed by two rounds of centrifugation, ethanol addition, and centrifugation to precipitation DNA as a pellet. Pellets were washed in 70% ethanol and resuspended in 20 μ L TE.

Genomic DNA Inserts

Flies bearing derivatives were allowed to lay eggs on apple juice plates, and embryo collections were made every 24 hours. Pooled samples were kept at -20°C until a few hundred microliters of embryos were collected. Genomic DNA was made without dechoriation by grinding embryos in $40\ \mu\text{L}$ Buffer A (100 mM NaCl, 30 mM Tris pH 8.0, 50 mM EDTA, with 0.5% Triton X-100 and 7.7 mM β -mercaptoethanol added fresh) per 25 μL embryos. Embryos were ground and spun lightly in a centrifuge. Finally, equal volumes 1.5% FMC InCert Agarose in 125 mM EDTA was added and inserts poured into 25 μL wells until hardened. Inserts were incubated overnight in 500 mM EDTA, 10 mM Tris pH 9.5, 1% Sarkosyl, 1 mg/mL Proteinase K, then transferred to TE (2 washes in TE with 0.1 mM PMSF, 6 washes in TE), and stored at 4°C . Digestions were done overnight in appropriate digestion buffer with 100 $\mu\text{g}/\text{mL}$ BSA, 10 $\mu\text{g}/\text{mL}$ DNase-free RNase A, and 0.1% Triton X-100.

Pulse Field Gel Electrophoresis and Molecular Weight Determination

PFGE followed one of three programs, to optimize separation of whole chromosomes or restriction fragments. All programs were run at 180V in 1/2 X TBE.

Program A (for separation of 215 kilobases to 945 kilobases): 20 seconds to 80 seconds with a 2 second ramp, run for 18 hours. Program B (for separation of 215 kilobases to 1200 kilobases): 60 seconds to 120 seconds with a 2 second ramp, run for 28 hours. Program C (for separation of 40 kilobases to 450 kilobases): 1 second to 30 seconds with a 1 second ramp, run for 24 hours.

Molecular mass was determined by weighted average of the masses derived from three or more gels. Smaller fragments (i.e. - the products of restriction digest) were weighed more than uncut chromosomes to account for poor resolution of large chromosomes in Pulse Field Gel Electrophoresis. Fragments were weighted as $((\text{chromosome size}_1/\text{fragment size}_1)+\dots+(\text{chromosome size}_n/\text{fragment size}_n))/(\text{fragment size}_1+\dots+\text{fragment size}_n)$, where chromosome size is the size of the entire chromosome based on the restriction fragment size.

Southern Hybridization

Transfers from PFGE or conventional agarose gels were done by acid depurination (15 minutes in 250 mN HCl), equilibrating in ABB (1.5 M NaCl, 400 mN NaOH) and wick-transferring overnight. Blot membranes were Hybond-N+ and were cross-linked using 120 mJ UV radiation. Southern hybridizations were done at 65°C in a rotating oven, using Quick-Hyb as prehybridization and hybridization

solutions. Prehybridization times were 30-90 minutes, and hybridization times were 120-300 minutes. 20 ng DNA was labeled for 20-45 minutes using random hexamer incorporation. Blots were washed twice in C+G1 (40 mM NaPO₄ pH 7.2, 5% SDS, 1 mM EDTA, 0.5% BSA) and twice in C+G2 (40 mM NaPO₄ pH 7.2, 1% SDS, 1 mM EDTA); all washes were done at 65°C for 20 minutes. Blots were exposed to film or PhosphorImager (Barium-Fluoride-Bromo-Europium, BaFBrEu) plates overnight.

Statistical Analyses

Statistical analyses were performed using custom spreadsheets designed for Excel from the Microsoft Office 98 suite (commonly utilizing AVERAGE, STDEV, CHIDIST, CHITEST, CHIINV, and TREND functions), or using Statistica 4.1 for the Macintosh from StatSoft, Inc.

Polymerase Chain Reaction

Polymerase Chain Reaction was done on either DNA inserts, MOM-prepped DNA, Quick Fly DNA, or ligated Inverse-PCR DNA. All reactions were carried out in a Perkin Elmer 9600 thermocycler, using Taq 2000 Polymerase in buffer containing 250 nM primers and 25 µM dNTPs. For inverse PCR: 94°C for 2 minutes; (94°C for 45 seconds, 60°C for 45 seconds, 72°C for 4 minutes) repeated 35 cycles; 72°C for 6

minutes. For all others: 94°C for 2 minutes; (94°C for 45 seconds, 55-63°C for 45 seconds, 72°C for 45 seconds) repeated 30 cycles; 72°C for 6 minutes.

Appendix B: Data

List of Clones

The following is a list of subclones that have been cloned from the *yellow* region. A (*) indicates that the site is artificial and introduced during cloning; a (†) indicates polymorphism between reported contig sequence (125DH110) and chromosomal (*Dp8-23*, *Dpγ238*, and *Tγ1337*) sequence; all coordinates are relative to transcription start. KS1.4 contains the *Dpγ238* inversion breakpoint, and BK0.6 contains the *Tγ1337* translocation breakpoint.

clone name	di stal		proxi mal	
	enzyme	coordi nate	enzyme	coordi nate
RS4. 0	*EcoRI	- 7694	Sal I	- 3819
RK2. 6	*EcoRI	- 7694	KpnI	- 5195
KS1. 4	KpnI	- 5195	Sal I	- 3819
HH1. 2	HhaI	- 7024	†HhaI	- 5836
HH0. 9	HhaI	- 4336	HhaI	- 3394
BK0. 5	KpnI	- 340	BamHI	+190
BK0. 6	BamHI	+190	KpnI	+795

Primers and Oligonucleotides

KM1	GCGCGAAAAGCTGCGGC
KM2	GCACTTCTCGCCATTCGCGC
KM4	GAGTGGTAGCTATTCGCC
KM5	ATTCGCGAGCCCGTCTCACAC
KM6	TAATGGCAAAGTAGGACTT
KM7	GATACCGCCTCATCCAGTAT
KM8	CCGAAGGCTAGAGAAGAACC
KM9	CTGGCTCCAACTATATCGCTCCTG
KM10	GGCTGGAGTGGGATCTTCC

KM1 1 CTCATCCGCCACATATCC
KM1 2 CTGCGCTGATTTGTGCCG
KM1 3 GCAATGTTCCAGGACAAAGGG
KM1 4 TAATCCTCTTCTGTGGACCG
KM1 5 CGCATTTTGAATCGATACAGCCG
KM1 6 CGCATGGAGAGGTATTGAGTATCGC
KM1 7 CCCAACTGTGTTTGTGTGTGTGCCG
KM1 8 GGGGAAAGAGATGCCAACTAAGCG
KM1 9 GCGGACCTCTCATAAATCCGG
KM2 0 CGCTAAAACCTTTACATGAGGG
KM2 1 GTATTAGATAAGCGATTCCG
KM2 2 CCGCCTCGCAAGTGTAACACG
KM2 3 CGTGTTACACTTGCGAGGCGG
KM2 4 CGTACTATATATCAAATCTAGC
KM2 5 GCTAGATTTGATATATAGTACG
KM2 6 CGATAATTGTCGGTTTATACC
KM2 7 GGTATAAACCGACAATTATCG
KM2 8 GCCATGTTACCATGTTGCTGC
KM2 9 GAAGCTCTTGACTGGCGTC
KM3 0 GTTACGTAATATATCAAATCTAGCAAGC
KM3 1 GCTTGCTAGATTTGATATATAGTACGTAAC
KM3 2 CGATAATTGTCGGTTTATACCATTCC
KM3 3 GGAATGGTATAAACCGACAATTATCG
KM3 4 AAAAGAATTCAACGAATTGGATCGGTTATAAGTAAGC
KM3 5 AAAAGGATCCAACCAATTCGATAATTATTAATCGATGG
KM3 6 AATTGAATGCAATTCAAAAATTAAGATATTGGC
KM3 7 GAACGGTATGCGAATAACGGAGC
KM3 8 CATAAGGTGGTCCCGTCGATAGCC
KM3 9 GGCTCTATCACTTTACTAGGTACGGC
KM4 0 GCGCTTATTTAATAGGTAGTCTAAG
KM4 1 GCCCACAACCTGTACGCTGCCAGCTG
KM4 2 CGTCTTCGGAACCTAGCAG
KM4 3 GCGTGAAATACTTCATTGTTTGTCCC
KM4 4 TTAATGTACGCATGCATAATTGAGC
KM4 5 GTGAACGGTATGCGAATAACGG
KM4 6 CCCCTTTCGCCAGCTGGCG

Maps of Chromosomes

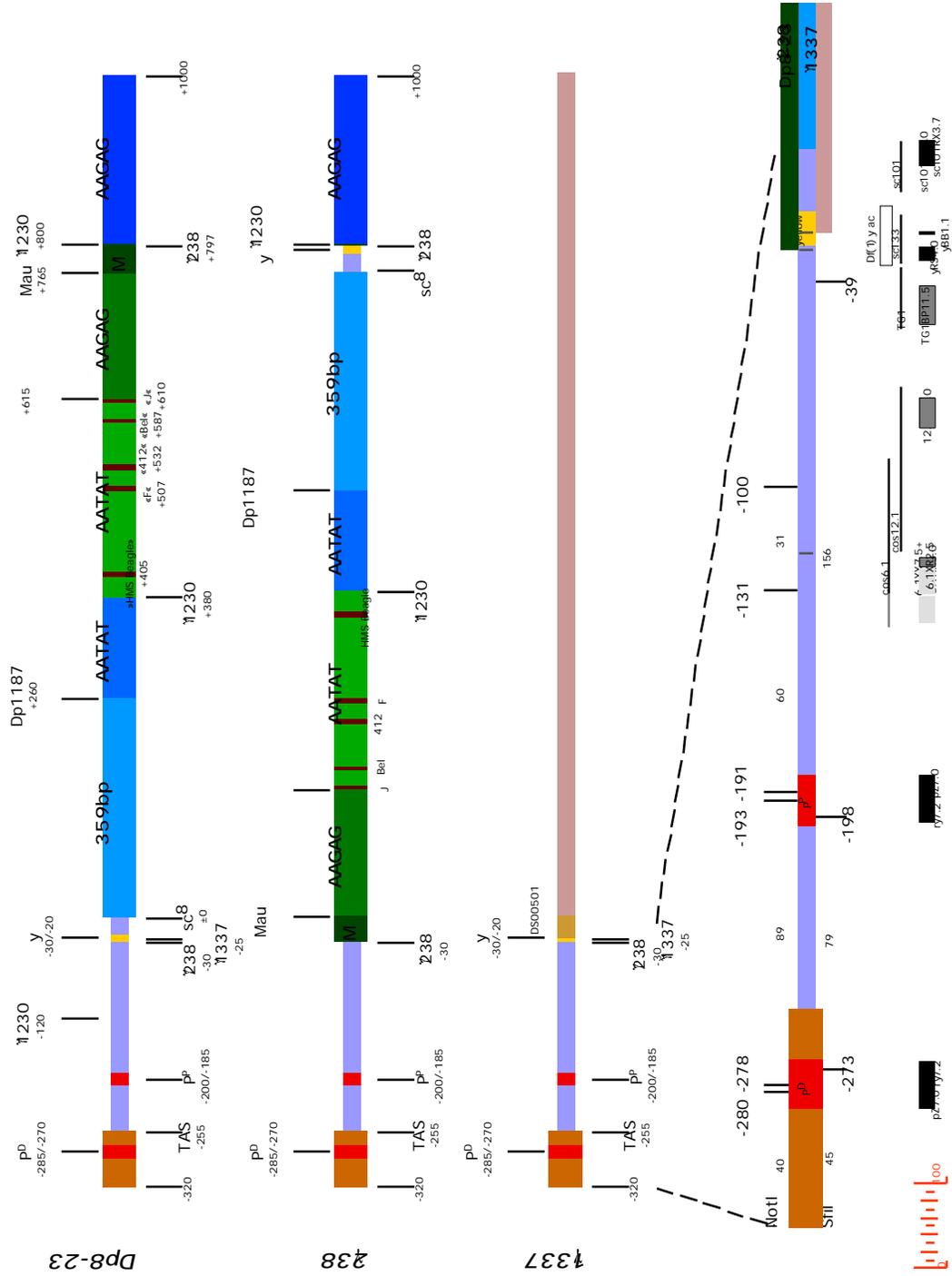


Figure 34: Structural Maps of *Dp8-23*, *Dp γ 238*, and *T γ 1337*

Symbolism is the same as Figures 12, 13, and 15. Additionally, subclones used for Southern analyses are shown below the test segment (black, well-mapped location; gray, fairly accurate location; white, approximate location).

Map of yellow

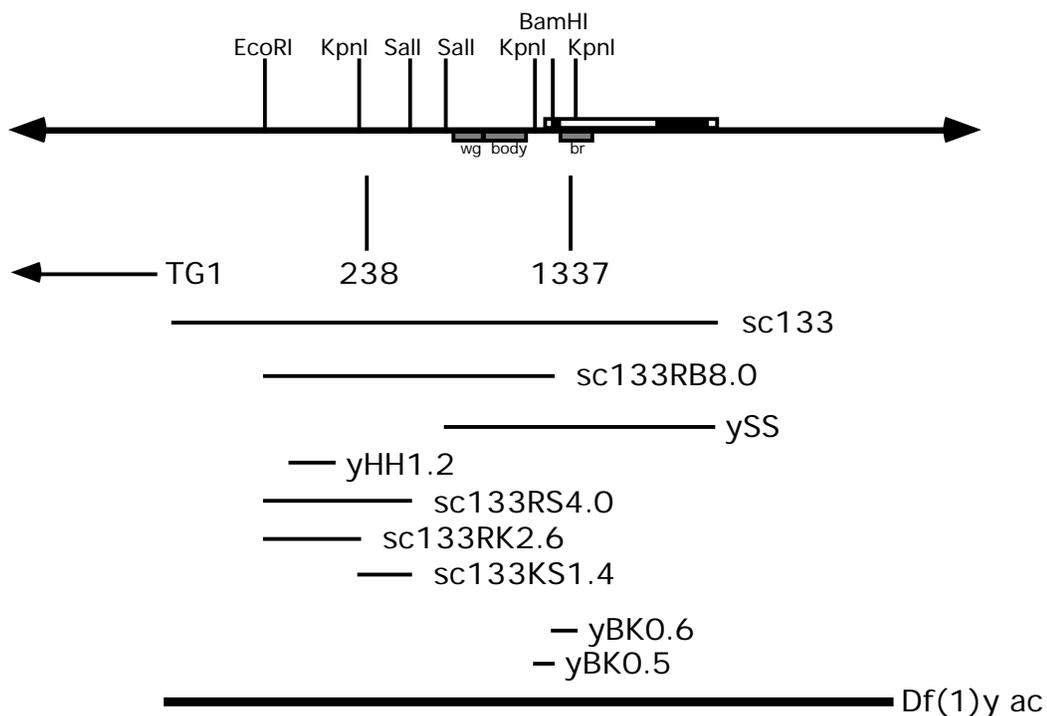


Figure 35: Map of *yellow*

Df(1)y ac shows the approximate extent of *Df(1) y ac*. All other lines below the chromosome line are subclones. Diagnostic restriction sites are shown relative to the *yellow* transcript (rectangle); EcoRI is not naturally occurring, and merely denotes the approximate location based on the RS4.0 subclone. 238 and 1337 show the breakpoints for *Dp γ 238* and *T γ 1337*, respectively.

Transmission Data

Line	sex	mean	stdev	flies	vials	Line	sex	mean	stdev	flies	vials
A2	♀	0.4548	0.0967	1954	14	A22	♂	0.4708	0.048	2055	11
A2	♂	0.4338	0.0423	2093	11	A22	♀	0.5037	0.0519	708	6
A2	♀	0.4732	0.0544	592	5	A22	♂	0.4483	0.0214	643	5
A2	♂	0.4718	0.0822	875	7	A23	♀	0.4974	0.0623	1128	15
A4	♀	0.5038	0.06	1777	9	A23	♂	0.4483	0.0593	1110	11
A4	♂	0.4341	0.0395	1105	7	A23	♀	0.4864	0.055	287	4
A4	♀	0.4737	0.0829	558	5	A23	♂	0.4745	0.0655	482	7
A4	♂	0.4477	0.0931	776	6	B51	♀	0.4727	0.1192	275	4
A6	♀	0.1313	0.0938	2551	15	B51	♂	0.4131	0.0361	547	7
A6	♂	0.3554	0.1164	2662	14	B56	♀	0.4964	0.0376	1036	7
A6	♀	0.1054	0.0648	1763	16	B56	♂	0.447	0.0524	856	7
A6	♂	0.2479	0.0888	3284	17	B58	♀	0.4522	0.2329	553	6
A6	♀	0.0861	0.0629	564	7	B58	♂	0.4443	0.0561	1087	7
A6	♂	0.2439	0.1154	4689	30	B64	♀	0.1033	0.1057	155	2
A6	♀	0.1357	0.0974	1453	14	B64	♂	0.2352	0.1348	1401	10
A6	♂	0.3327	0.0867	1453	11	B65	♀	0.039	0.0298	583	6
A8	♀	0.296	0.1379	1727	13	B65	♂	0.2416	0.1855	866	7
A8	♂	0.4336	0.0815	3067	16	B66	♀	0.3857	0.1035	767	6
A8	♀	0.2891	0.1502	2770	18	B66	♂	0.3756	0.0605	1489	8
A8	♂	0.4614	0.0937	3389	20	B66	♀	0.3849	0.0995	997	11
A8	♀	0.257	0.0915	1878	16	B66	♂	0.4442	0.0859	1738	9
A8	♂	0.4164	0.1009	2185	15	B67	♀	0.0334	0.0448	922	7
A10	♀	0.4761	0.0805	2785	15	B67	♂	0.1878	0.1391	905	7
A10	♂	0.4825	0.0605	2224	11	B74	♀	0.5465	0.0245	331	3
A10	♀	0.4479	0.0601	909	7	B74	♂	0.4961	0.0756	307	3
A10	♂	0.4318	0.1196	612	5	B76	♀	0.5446	0.0251	342	3
A13	♀	0.5105	0.0445	1653	17	B76	♂	0.4377	0.03	393	3
A13	♂	0.4593	0.067	1564	16	B78	♀	0.6228	0.0002	167	2
A13	♀	0.5027	0.0691	252	4	B78	♂	0.4416	0.0405	373	3
A13	♂	0.4336	0.1064	414	6	B79	♀	0.5721	0.0009	187	2
A17	♀	0.5308	0.0541	1905	11	B79	♂	0.5705	0.0715	312	3
A17	♂	0.4639	0.0453	2372	12	B80	♀	0.0546	0.0361	1196	9
A17	♀	0.414	0.0267	536	4	B80	♂	0.1845	0.1167	1690	12
A17	♂	0.4118	0.0319	664	5	B81	♂	0.194	0.1635	1581	8
A21	♀	0.4907	0.0294	734	5	B81	♀	0.092	0.0557	1455	6
A21	♂	0.4366	0.0343	504	3	B82	♀	0.4742	0.0409	160	3
A21	♀	0.5298	0.0513	2342	15	B82	♂	0.4674	0.1197	141	3
A21	♂	0.4897	0.0421	3617	21	B84	♀	0.5201	0.052	279	3
A22	♀	0.4977	0.0401	2392	13	B84	♂	0.408	0.0491	338	3

Line	sex	mean	stdev	flies	vials	Line	sex	mean	stdev	flies	vials
B85	♀	0.5304	0.0198	179	3	C153	♀	0.4478	?	67	1
B85	♂	0.4782	0.0317	162	3	C155	♀	0.0444	?	135	1
B86	♀	0.4937	0.0921	131	2	C155	♂	0.2361	0.1037	1267	6
B86	♂	0.4956	0.058	107	2	C156	♀	0.0508	?	118	1
B87	♀	0.4765	0.0412	1437	12	C156	♂	0.1011	0.0338	1383	6
B87	♂	0.433	0.0388	1692	11	C158	♀	0.3284	?	67	1
B88	♀	0.4537	0.0921	160	3	C158	♀	0.374	0.0925	1652	11
B88	♂	0.4522	0.0321	172	3	C161	♀	0.3361	?	119	1
B91	♀	0.4952	0.0328	152	3	C161	♀	0.4555	0.0938	653	9
B91	♂	0.4354	0.1015	113	2	C162	♀	0.2872	?	94	1
B92	♀	0.5489	0.0744	370	3	C162	♀	0.3373	0.079	1120	8
B92	♂	0.4117	0.0497	342	3	C163	♀	0.4235	?	85	1
B93	♀	0.4454	0.0099	115	2	C163	♀	0.2678	0.1666	1520	10
B93	♂	0.4389	0.0529	159	3	C164	♀	0.4604	?	139	1
B98	♀	0.5671	0.0281	99	2	C164	♀	0.4939	0.0708	735	10
B98	♂	0.5147	0.0334	343	3	C165	♀	0.4697	?	132	1
B100	♀	0.0779	0.0456	602	8	C165	♀	0.4216	0.0862	517	9
B100	♂	0.4431	0.1358	1319	8	C166	♀	0.4078	?	103	1
B101	♀	0.4723	0.0716	838	8	C166	♀	0.5351	0.0619	718	12
B101	♂	0.402	0.0377	1418	10	C167	♀	0.0676	?	133	1
B102	♀	0.4214	0.1245	92	2	C167	♂	0.3436	0.0868	1751	8
B102	♂	0.2879	0.2503	188	3	C168	♀	0.46	?	50	1
B109	♀	0.0804	0.0529	1087	10	C173	♀	0.542	?	131	1
B109	♂	0.2187	0.0805	1575	10	C173	♀	0.4024	0.119	2136	12
B112	♀	0.3864	?	44	1	C175	♀	0.446	?	139	1
B112	♂	0.3704	0.0262	108	2	C175	♀	0.46	0.0513	726	12
B113	♀	0.451	?	51	1	C176	♀	0.3704	?	81	1
B113	♂	0.3696	0.0382	182	3	C178	♀	0.45	?	140	1
B114	♀	0.4318	0.0425	356	3	C182	♀	0.4364	?	55	1
B114	♂	0.432	0.0627	292	3	C184	♀	0.4603	?	63	1
B115	♀	0.4861	0.041	136	2	C184	♀	0.4971	0.048	214	3
B115	♂	0.4615	0.027	181	3	C184	♂	0.5085	?	177	1
B118	♀	0.4716	0.2094	105	2	C186	♀	0.0822	?	85	1
B118	♂	0.463	0.0585	196	3	C186	♂	0.1297	0.0854	876	4
B119	♀	0.4367	0.0624	239	2	C188	♀	0.4783	?	69	1
B119	♂	0.4857	0.1046	273	3	C193	♀	0.0608	?	115	1
B120	♀	0.4609	0.0861	81	2	C193	♂	0.3508	0.1259	1904	8
B120	♂	0.4889	0.0187	187	3	C195	♀	0.3385	?	65	1
B121	♂	0.4519	0.1113	160	3	C196	♀	0.0438	?	137	1
B123	♀	0.524	0.1073	294	3	C196	♂	0.4024	0.1327	1739	8
B123	♂	0.4198	0.0371	355	3	C198	♀	0.5806	?	62	1
C151	♀	0.5	?	102	1	C202	♀	0.425	?	40	1

Line	sex	mean	stdev	flies	vials	Line	sex	mean	stdev	flies	vials
C204	♀	0.4	?	50	1	D259	♀	0.5176	?	85	1
C206	♀	0.4873	?	158	1	D261	♀	0.4604	?	139	1
C206	♀	0.5464	0.0451	596	11	D262	♀	0.5145	?	138	1
C208	♀	0.5079	?	63	1	E302	♀	0.413	?	92	1
C211	♀	0.3529	?	51	1	E306	♀	0.4561	?	114	1
C213	♀	0.3455	?	55	1	E316	♂	0.4898	0.0169	663	3
C214	♀	0.4595	?	37	1	H453	♀	0.6122	?	49	1
C215	♀	0.4497	?	149	1	H454	♀	0.6479	?	71	1
C215	♀	0.4656	0.034	795	13	H456	♀	0.5926	?	27	1
C216	♀	0.5556	?	90	1	H457	♀	0.5	?	52	1
C217	♀	0.3711	?	97	1	H465	♀	0.3867	?	75	1
C217	♀	0.3811	0.0797	1705	11	H468	♀	0.551	?	49	1
C218	♀	0.036	?	139	1	H470	♀	0.6154	?	52	1
C219	♀	0.522	?	159	1	H473	♀	0.4932	?	73	1
C219	♀	0.481	0.0531	860	10	H474	♀	0.5102	?	49	1
C220	♀	0.4154	?	65	1						
C221	♀	0.6102	?	59	1						
D255	♀	0.5352	?	142	1						

Brooding data

	26C	A6	B65	B81	C155	C156	C167	C196	C218
days 0-2	0.3036	0.3029	0.2087	0.3388	0.2679	0.3556	0.2967	0.3478	0.1939
days 2-4	0.3019	0.3981	0.1930	0.3772	0.3000	0.3388	0.4490	0.4079	0.2317
days 4-6	0.2702	0.2805	0.2010	0.3257	0.2535	0.2293	0.3330	0.3415	0.1727
days 6-8	0.2130	0.2650	0.1559	0.3247	0.1246	0.1997	0.2618	0.3304	0.0913
days 8-10	0.2121	0.2342	0.0904	0.2898	0.0921	0.1334	0.2315	0.2882	0.0636
days 10-12	0.0500	0.3050	0.1050	0.2291	0.0690	0.0694	0.2830	0.2740	0.0929
days 12-14	0.0968	0.3302	0.0397	0.2404	0.0763	0.0202	0.2152	0.2538	0.1066
days 14-18	0.2000	0.4098	0.0561	0.2706	0.0320	0.0385	0.3367	0.2835	0.0174

Selection Data

22A	gen 1	gen 2	gen 3	gen 4	gen 5	22A	gen 2	gen 3	gen 4	gen 5
up 1	0.3770	0.3657		0.1930	0.3093	down 1	0.2535	0.1702	0.2143	0.3548
up 2	0.4058	0.3125	0.4805		0.3789	down 2	0.1795		0.2308	0.2500
up 3	0.4359	0.2963	0.2685	0.2899	0.2527	down 3			0.3433	0.2051
up 4	0.0000	0.2920	0.4494	0.3143	0.3452	down 4		0.1827	0.3265	0.3800
up 5	0.4589	0.3143	0.3107	0.3000	0.3093	down 5		0.3200	0.3387	0.2899
up 6	0.4800	0.3167	0.6939		0.2778	down 6			0.3889	0.2165
up 7	0.3462	0.2637			0.2174	down 7				0.3537
up 8	0.0340	0.1415	0.2300			down 8				0.5673
up 9	0.2993	0.2037	0.4952			down 9				
up 10	0.1899	0.1857	0.0000			down 10				
average	0.3027	0.2692	0.3660	0.2743	0.2987	average	0.2165	0.2243	0.3071	0.3272
stdev	0.1726	0.0702	0.2102	0.0551	0.0548	stdev	0.0523	0.0831	0.0690	0.1174

26C	gen 1	gen 2	gen 3	gen 4	gen 5	26C	gen 2	gen 3	gen 4	gen 5
up 1	0.0000	0.2149		0.1591	0.1545	down 1	0.1593	0.2021	0.3864	0.3765
up 2	0.2342	0.1452	0.1750		0.2410	down 2	0.1608	0.0000	0.2000	0.2174
up 3	0.0609	0.4158	0.2000	0.3472	0.2048	down 3		0.2941	0.2029	0.0833
up 4	0.0000	0.1121	0.3882	0.1692	0.0000	down 4		0.2059	0.2813	0.1368
up 5	0.3094	0.1504			0.3043	down 5		0.4220		0.2817
up 6	0.0000	0.1803	0.2336		0.3021	down 6		0.0594		0.1579
up 7	0.0000	0.4560	0.1429	0.2299	0.1935	down 7		0.3190		0.1757
up 8	0.1484	0.3507	0.3298		0.1000	down 8		0.1062		0.5362
up 9	0.0000	0.1102	0.2768		0.2875	down 9		0.3250		
up 10	0.0000	0.2149	0.1182		0.2879	down 10				
average	0.0753	0.2351	0.2331	0.2264	0.2076	average	0.1601	0.2149	0.2676	0.2457
stdev	0.1153	0.1268	0.0935	0.0864	0.1000	stdev	0.0011	0.1390	0.0876	0.1485

J21A	gen 1	gen 2	gen 3	gen 4	gen 5	J21A	gen 2	gen 3	gen 4	gen 5
up 1		0.4404	0.4366	0.3814	0.3506	down 1	0.2339		0.3623	0.3500
up 2		0.3178	0.3548	0.5000	0.3605	down 2	0.2308	0.3905	0.3684	0.4632
up 3	0.7444	0.4224	0.5000	0.3958	0.4468	down 3	0.3723		0.5147	0.4095
up 4	0.7007	0.6667	0.2979		0.4028	down 4	0.3333	0.4286		0.3793
up 5	0.3627	0.3524	0.3014	0.3519	0.2817	down 5	0.4141		0.4762	0.4000
up 6	0.1813	0.3661		0.4286	0.3750	down 6	0.3398	0.4946	0.4000	0.4124
up 7	0.4392	0.3333	0.3448		0.3299	down 7	0.3468	0.3514	0.3700	0.3258
up 8	0.3496	0.3217			0.4563	down 8		0.3371	0.6211	0.4783
up 9	0.6581	0.2764	0.3333		0.4111	down 9	0.4524	0.4270		0.4505
up 10	0.4231	0.3025	0.4778		0.3711	down 10	0.3212			0.3951
average	0.4824	0.3800	0.3808	0.4115	0.3786	average	0.3383	0.4048	0.4447	0.4064
stdev	0.1983	0.1129	0.0794	0.0566	0.0530	stdev	0.0731	0.0579	0.0976	0.0483

Dp γ 238 Breakpoint

Maupiti (Wahlstrom map)

⁷⁶⁶⁹⁰ ACTGCACTTAATTAATTCT TATTACATT CGATCAATTTGAAACTCAAACACAAAGA ⁷⁶⁷⁴⁴
 TGACGTGAATTAATTAAGA ATAATGTAA GCTAGTTAAACTTTGAGTTTGTGTTTCT

deleted

yellow (relative to transcription start)

⁻⁵¹⁰⁸ TAATTATTAATCGATGGGTGGATA CAGGGG TGACAGTATCATAAAGGATATTGTAG ⁻⁵⁰⁵³
 ATTAATAATTAGCTACCCACCTAT GTCCCC ACTGTCATAGTATTTCCCTATAACATC

deleted

In(Dp8-23) γ 238 left (ry⁺) arm

TAATTATTAATCGATGGGTGGATA AGAATTAATTAAGTGCAGT
 ATTAATAATTAGCTACCCACCTATTGTTAATTAATTCACGTCA

<--TEL KS1.4 Maupiti^L CEN-->

In(Dp8-23) γ 238 right (y⁺) arm

CTACAATATCCTTTATGATACTGTCATCGATCACATTCGATCGATCGATCAATTTGAAACTC
 GATGTTATAGGAAATACTATGACAGTAGCTAGTGTAAGCTAGCTAGCTAGTTAAACTTTGAG

<--CEN <-*yellow* novel Maupiti^S TEL-->

T γ 1337 Breakpoints

2L (P1 DS00501)

⁷⁶⁶⁸⁰ GATTCAGACCTATAAACTTGGCCGTTGCTGTAGCTGATTATGTTGGTATAGGTG ⁷⁶⁷³⁵
 CTAAGTCTGGATATTTGAACCTTCAACGACATCGACTAATACAACCATATCCAC

yellow (relative to transcription start)

⁺⁶⁸⁶ TTATTTGTATGCCTTTTCACCATTTTGGCAGAGGAATAAATATGACAATATAT ⁺⁷³⁸
 AATAAACATACGGAAAAGTGGTAAAACCGTCTCCTTATTTATACTGTTATATA

T(2^P;Dp8-23^D) γ 1337, ry⁺

TTATTTGTATGCCTTTTCACCATTTTGGTGA^{AAATCA}GCTGTAGCTGATTATGTTGGTATAGGTG
 AATAAACATACGGAAAAGTGGTAAAACCACTTTTGTAGTCGACATCGACTAATACAACCATATCCAC

<--TEL^{Dp} -> *yellow* novel 2L CEN^{2L}-->

T(Dp8-23^P;2^D) γ 1337, 0

GATTCAGACCTATAAACTTGGCCGTTGCTATATTCCTTGGCAGAGGAATAAATATGACAATATAT
 CTAAGTCTGGATATTTGAACCTTCAACGATATAAGGAACCGTCTCCTTATTTATACTGTTATATA

<--TEL^{2L} 2L novel -> *yellow* CEN^{Dp}-->

Bibliography

- M. D. Adams, S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, H. O. Smith, R. A. Gibbs, E. W. Myers, G. M. Rubin, J. C. Venter and et al. *The genome sequence of *Drosophila melanogaster**. **Science** 287:2185-95, 2000.
- K. Afshar, N. R. Barton, R. S. Hawley and L. S. B. Goldstein. *DNA binding and meiotic chromosomal localization of the *Drosophila* NOD kinesin-like protein*. **Cell** 81:129-138, 1995.
- K. Ahmad and K. G. Golic. *Telomere loss in somatic cells of *Drosophila* causes cell cycle arrest and apoptosis*. **Genetics** 151:1041-51, 1999.
- B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts and J. D. Watson. Molecular Biology of the Cell. Garland Publishing, Inc., New York, 1989.
- M. Ashburner. *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989.
- M. Ashburner and E. Novitski. *The Genetics and Biology of *Drosophila**. Academic Press, London ; New York, 1976.
- J. G. Ault, A. J. Demarco, E. D. Salmon and C. L. Rieder. *Studies on the Ejection Properties of Asters - Astral Microtubule Turnover Influences the Oscillatory Behavior and Positioning of Mono-Oriented Chromosomes*. **Journal of Cell**

Science 99:701-710, 1991.

- A. E. Barry, E. V. Howman, M. R. Cancilla, R. Saffery and K. H. Choo. *Sequence analysis of an 80 kb human neocentromere*. **Human Molecular Genetics** 8:217-227, 1999.
- A. E. Barry, M. Bateman, E. V. Howman, M. R. Cancilla, K. M. Tainton, D. V. Irvine, R. Saffery and K. H. A. Choo. *The 10q25 neocentromere and its inactive progenitor have identical primary nucleotide sequence: Further evidence for epigenetic modification*. **Genome Research** 10:832-838, 2000.
- T. Bechert, S. Heck, U. Fleig, S. Diekmann and J. H. Hegemann. *All 16 centromere DNAs from *Saccharomyces cerevisiae* show DNA curvature*. **Nucleic Acids Research** 27:1444-1449, 1999.
- M. Blower and G. H. Karpen. in preparation.
- S. Bongiorno, O. Cintio and G. Prantera. *The relationship between DNA methylation and chromosome imprinting in the coccid *planococcus citri**. **Genetics** 151:1471-8, 1999.
- C. B. Bridges. *Non-disjunction as proof of the chromosome theory of heredity*. **Genetics** 1:1-52, 1916.
- V. L. Chandler, K. M. Kubo and J. B. Hollick. *b and pl Paramutation in Maize: Heritable Transcriptional States Programmed During Development, in Cold Spring Harbor monograph series 32: Epigenetic mechanisms of gene regulation*. Cold spring Harbor Laboratory Press, Plainview, New York, 1996.
- K. H. Choo. *Centromere DNA dynamics: latent centromeres and neocentromere formation*. **American Journal of Human Genetics** 61:1225-33, 1997.
- K. H. A. Choo. *Centromerization*. **Trends in Cell Biology** 10:182-188, 2000.
- K. R. Cook, T. D. Murphy, T. C. Nguyen and G. H. Karpen. *Identification of trans-acting genes necessary for centromere function in *Drosophila melanogaster* using centromere-defective minichromosomes*. **Genetics** 145:737-747, 1997.

- S. J. Counce. *Developmental morphology of polar granules in Drosophila. Including observations on pole cell behavior and distribution during embryogenesis.* **Journal of Morphology** 112:129-146, 1963.
- S. Cumberledge and J. Carbon. *Mutational analysis of meiotic and mitotic centromere function in Saccharomyces cerevisiae.* **Genetics** 117:203-12, 1987.
- C. D. Darlington. *Misdivision and the Genetics of the Centromere.* **Journal of Genetics** 37:342-363, 1939.
- A. de la Puente, E. Velasco, L. A. Perez Jurado, C. Hernandez-Chico, F. M. van de Rijke, S. W. Scherer, A. K. Raap and J. Cruces. *Analysis of the monomeric alphoid sequences in the pericentromeric region of human chromosome 7.* **Cytogenetics and Cellular Genetics** 83:176-81, 1998.
- M. Demerec. *Biology of Drosophila.* Cold Spring Harbor Laboratory Press, Plainview, N.Y., 1994.
- K. W. Dobie, K. L. Hari, K. A. Maggert and G. H. Karpen. *Centromere proteins and chromosome inheritance: a complex affair.* **Current Opinion in Genetics & Development** 9:206-217, 1999.
- K. M. Donaldson and G. H. Karpen. *Trans-suppression of terminal deficiency-associated position effect variegation in Drosophila.* **Genetics** 145:325-337, 1997.
- K. Ekwall, T. Olsson, B. M. Turner, G. Cranston and R. C. Allshire. *Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres.* **Cell** 91:1021-32, 1997.
- W. Flemming. *Beitrag zur Kenntnis der Zelle und ihrer Lebenserscheinungen, Teil II.* **Archiv Mikrosk Anat** 18:151-259, 1880.
- C. Goday and S. Pimpinelli. *Centromere organization in meiotic chromosomes of Parascaris univalens.* **Chromosoma** 98:160-6, 1989.

- C. Goday, J. M. Gonzalez-Garcia, M. R. Esteban, G. Giovinazzo and S. Pimpinelli. *Kinetochores and chromatin diminution in early embryos of *Parascaris univalens**. **Journal of Cell Biology** 118:23-32, 1992.
- L. S. Goldstein. *Kinetochore structure and its role in chromosome orientation during the first meiotic division in male *Drosophila melanogaster**. **Cell** 25:591-602, 1981.
- K. G. Golic, M. M. Golic and S. Pimpinelli. *Imprinted control of gene activity in *Drosophila**. **Current Biology** 8:1273-1276, 1998.
- M. M. Green, M.-T. Yamamoto and G. L. G. Miklos. *Genetic instability in *Drosophila melanogaster*: Cytogenetic analysis of MR-induced X-chromosome deficiencies*. **Proceedings of the National Academy of Science USA** 84:4533-4537, 1987.
- S. I. S. Grewal and A. J. S. Klar. *Chromosomal Inheritance of Epigenetic States in Fission Yeast During Mitosis and Meiosis*. **Cell** 86:95-101, 1996.
- K. M. Hahnenberger, J. Carbon and L. Clarke. *Identification of DNA regions required for mitotic and meiotic functions within the centromere of *Schizosaccharomyces pombe* chromosome I*. **Molecular and Cellular Biology** 11:2206-15, 1991.
- P. Hahnfeldt, L. R. Hlatky, D. J. Brenner and R. K. Sachs. *Chromosome Aberrations Produced by Radiation: The Relationships between Excess Acentric Fragments and Dicentrics*. **Radiation Research** 141:1995.
- J. J. Harrington, G. Van Bokkelen, R. W. Mays, K. Gustashaw and H. F. Willard. *Formation of de novo centromeres and construction of first-generation human artificial microchromosomes [see comments]*. **Nature Genetics** 15:345-55, 1997.
- J. H. Hayden, S. S. Bowser and C. L. Rieder. *Kinetochores capture astral microtubules during chromosome attachment to the mitotic spindle: direct visualization in live newt lung cells*. **Journal of Cell Biology** 111:1039-1045, 1990.

- J. A. Heddle and A. V. Carrano. *The DNA Content of Micronuclei Induced in Mouse Bone Marrow by gamma-Irradiation: Evidence That Micronuclei Arise From Acentric Chromosomal Fragments*. **Mutation Research** 44:63-69, 1977.
- S. Henikoff, K. Ahmad, J. S. Platero and B. van Steensel. *Heterochromatic deposition of centromeric histone H3-like proteins*. **Proceedings of the National Academy of Sciences of the United States of America** 97:716-721, 2000.
- A. A. Hyman and P. K. Sorger. *Structure and function of kinetochores in budding yeast*. **Annual Review of Cell and Developmental Biology** 11:471-95, 1995.
- S. E. Jacobsen. *Gene silencing: Maintaining methylation patterns*. **Current Biology** 9:R617-9, 1999.
- J. Jakowitsch, I. Papp, E. A. Moscone, J. van der Winden, M. Matzke and A. J. Matzke. *Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans*. **Plant Journal** 17:131-40, 1999.
- P. L. Jones and A. P. Wolffe. *Relationships between chromatin organization and DNA methylation in determining gene expression*. **Seminars in Cancer Biology** 9:339-47, 1999.
- A. L. Jorgensen. *Alphoid repetitive DNA in human chromosomes*. **Danish Medical Bulletin** 44:522-34, 1997.
- G. H. Karpen and A. C. Spradling. *Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in Drosophila*. **Cell** 63:97-107, 1990.
- . *Analysis of subtelomeric heterochromatin in the Drosophila minichromosome Dp1187 by single P element insertional mutagenesis*. **Genetics** 132:737-53, 1992.
- G. H. Karpen and R. C. Allshire. *The case for epigenetic effects on centromere identity and function*. **Trends in Genetics** 13:489-496, 1997.

- S. U. Kass and A. P. Wolffe. Histones, Histone Modifications, and the Inheritance of Chromatin Structure, in Cold Spring Harbor monograph series 32: Epigenetic mechanisms of gene regulation. Cold spring Harbor Laboratory Press, Plainview, New York, 1996.
- R. L. Kelley, V. H. Meller, P. R. Gordadze, G. Roman, R. L. Davis and M. I. Kuroda. *Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin*. **Cell** 98:513-22, 1999.
- K. Kitada, E. Yamaguchi, K. Hamada and M. Arisawa. *Structural analysis of a Candida glabrata centromere and its functional homology to the Saccharomyces cerevisiae centromere*. **Current Genetics** 31:122-127, 1997.
- J. Koch. *Neocentromeres and alpha satellite: a proposed structural code for functional human centromere DNA*. **Human Molecular Genetics** 9:149-54, 2000.
- K. E. Koehler, C. L. Boulton, H. E. Collins, R. L. French, K. C. Herman, S. M. Lacefield, L. D. Madden, C. D. Schuetz and R. S. Hawley. *Spontaneous X chromosome MI and MII nondisjunction events in Drosophila melanogaster oocytes have different recombinational histories*. **Nature Genetics** 14:406-14, 1996.
- A. L. Koen. Maternal aging and chromosomal defects, in Aging and Reproductive Physiology. Ann Arbor Science, Ann Arbor, 1976.
- P. Laurenti, Y. Graba, R. Rosset and J. Pradel. *Genetic and molecular analysis of terminal deletions of chromosome 3R of Drosophila melanogaster*. **Gene** 154:177-181, 1995.
- M. H. Le, D. Duricka and G. H. Karpen. *Islands of complex DNA are widespread in Drosophila centric heterochromatin*. **Genetics** 141:283-303, 1995.
- B. Lewin. Genes IV. Oxford University Press, Oxford ; New York, 1990.
- H. Liang, W. H. Wright, C. L. Rieder, E. D. Salmon, G. Profeta, J. Andrews, Y. G.

- Liu, G. J. Sonek and M. W. Berns. *Directed Movement of Chromosome Arms and Fragments in Mitotic Newt Lung Cells Using Optical Scissors and Optical Tweezers*. **Experimental Cell Research** 213:308-312, 1994.
- D. L. Lindsley and G. G. Zimm. *The Genome of Drosophila melanogaster*. Academic Press, Inc., San Diego, 1992.
- D. L. Lindsley, L. Sandler, B. S. Baker, A. T. Carpenter, R. E. Denell, J. C. Hall, P. A. Jacobs, G. L. Miklos, B. K. Davis, R. C. Gethmann, R. W. Hardy, A. H. Steven, M. Miller, H. Nozawa, D. M. Parry, M. Gould-Somero and M. Gould-Somero. *Segmental aneuploidy and the genetic gross structure of the Drosophila genome*. **Genetics** 71:157-84, 1972.
- M. Ljungman. *The Influence of Chromatin Structure on the Frequency of Radiation-Induced DNA Strand Breaks: A Study Using Nuclear and Nucleoid Monolayers*. **Radiation Research** 126:58-64, 1991.
- A. R. Lohe, A. J. Hilliker and P. A. Roberts. *Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster*. **Genetics** 134:1149-74, 1993.
- J. Lopez, G. H. Karpen and T. L. Orr-Weaver. *The Relationship between Sister-Chromatid Cohesion and Kinetochore Formation at the Centromere*. **Current Biology** submitted.
- K. A. Maggert and G. H. Karpen. *Acquisition and metastability of centromere identity and function: Sequence analysis of a human neocentromere*. **Genome Research** 10:725-728, 2000.
- B. Marzais, S. G. Vorsanova, G. Roizes and Y. B. Yurov. *Analysis of alphoid DNA variation and kinetochore size in human chromosome 21: evidence against pathological significance of alphoid satellite DNA diminutions*. **Tsitologiya i Genetika** 33:25-31, 1999.
- J. M. Mason, E. Strobel and M.M. Green. *mu-2: Mutator gene in Drosophila that potentiates the induction of terminal deficiencies*. **Proceedings of the National Academy of Science USA** 81:60900-6094, 1984.

- J. M. Mason, L. E. Champion and G. Hook. *Germ-line effects of a mutator, mu2, in Drosophila melanogaster*. **Genetics** 146:1381-1397, 1997.
- J. M. Mason, R. A. Voelker, D. Rosen, A. R. Campos, K. White and J. K. Lim. *Localization of terminal deficiency breakpoints on the X chromosome*. **Drosophila Information Service** 63:164-165, 1986.
- B. McClintock. The discovery and characterization of transposable elements : the collected papers of Barbara McClintock. Garland Pub., New York, 1987.
- P. B. Meluh, P. Yang, L. Glowczewski, D. Koshland and M. M. Smith. *Cse4p is a component of the core centromere of Saccharomyces cerevisiae*. **Cell** 94:607-13, 1998.
- G. Mendel. *Experiments in Plant Hybridization*. 1865.
- D. P. Moore, A. W. Page, T. T. Tang, A. W. Kerrebrock and T. L. Orr-Weaver. *The cohesion protein MEI-S332 localizes to condensed meiotic and mitotic centromeres until sister chromatids separate*. **Journal of Cell Biology** 140:1003-12, 1998.
- M. R. Murphy, D. M. Fowlkes and M. Fitzgerald-Hayes. *Analysis of Centromere Function in Saccharomyces-Cerevisiae Using Synthetic Centromere Mutants*. **Chromosoma** 101:189-197, 1991.
- T. D. Murphy. Characterization of Cis-acting Chromosomal elements and Trans-acting factors required for chromosome inheritance in drosophila. Dissertation of University of California San Diego, 1998.
- T. D. Murphy and G. H. Karpen. *Interaction between the nod+ kinesin-like gene and extracentromeric sequences are required for transmission of a Drosophila minichromosome*. **Cell** 81:139-148, 1995a.
- . *Localization of centromere function in a Drosophila minichromosome*. **Cell** 82:599-609, 1995b.

- . *Centromeres take flight: alpha satellite and the quest for the human centromere*. **Cell** 93:317-320, 1998.
- H. H. Ng and A. P. Bird. *DNA methylation and chromatin modifications*. **Current Opinion in Genetics and Development** 9:158-163, 1999.
- E. Novitski. *The Genetic Consequences of Anaphase Bridge Formation in Drosophila*. **Genetics** 37:270, 1951.
- . *Genetic measures of centromere activity in Drosophila melanogaster*. **Symposium on Genetics and Recombination 1954** 45:151--169, 1955.
- N. L. Oleincik, U. Balasubramaniam, L. Xue and S. Chiu. *Nuclear structure and the microdistribution of radiation damage in DNA*. **International Journal of Radiation Biology** 66:523-529, 1994.
- S. L. Page and L. G. Shaffer. *Chromosome stability is maintained by short intercentromeric distance in functionally dicentric human Robertsonian translocations*. **Chromosome Research** 6:115-22, 1998.
- S. L. Page, W. C. Earnshaw, K. H. Choo and L. G. Shaffer. *Further evidence that CENP-C is a necessary component of active centromeres: studies of a dic(X; 15) with simultaneous immunofluorescence and FISH*. **Human Molecular Genetics** 4:289-94, 1995.
- S. L. Page, J. C. Shin, J. Y. Han, K. H. A. Choo and L. G. Shaffer. *Breakpoint Diversity Illustrates Distinct Mechanisms for Robertsonian Translocation Formation*. **Human Molecular Genetics** 5:1279-1288, 1996.
- D. K. Palmer, K. O'Day and R. L. Margolis. *Biochemical analysis of CENP-A, a centromeric protein with histone-like properties*. **Progress in Clinical Biology Research** 318:61-72, 1989.
- D. K. Palmer, K. O'Day, M. H. Wener, B. S. Andrews and R. L. Margolis. *A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones*. **Journal of Cell Biology** 104:805-15, 1987.

- D. K. Palmer, K. O'Day, H. L. Trong, H. Charbonneau and R. L. Margolis. *Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone.* **Proceedings of the National Academy of Science USA** 88:3734-8, 1991.
- J. F. Partridge, B. Borgstrom and R. C. Allshire. *Distinct Protein Interaction Domains and Protein Spreading in a Complex Centromere.* **Genes & Development** 14:783-791, 2000.
- J. S. Platero, K. Ahmad and S. Henikoff. *A Distal Heterochromatic Block Displays Centromeric Activity When Detached from a Natural Centromere.* **Molecular Cell** 4:995-1004, 1999.
- A. F. Pluta, A. M. Mackay, A. M. Ainsztein, I. G. Goldberg and W. C. Earnshaw. *The Centromere: Hub of Chromosomal Activities.* **Science** 270:1591-1594, 1995.
- H. Potter. *Review and hypothesis: Alzheimer disease and Down syndrome-chromosome 21 nondisjunction may underlie both disorders.* **American Journal of Human Genetics** 48:1192-1200, 1991.
- M. Ptashne. *The Behavior of Strong and Weak Centromeres at Second Anaphase of Drosophila melanogaster.* **Genetics** 45:499-506, 1960.
- M. M. Rhoades. *Preferential segregation in maize, in Heterosis.* The Iowa State College Press, Ames, Iowa, 1952.
- C. L. Rieder. *The formation, structure, and composition of the mammalian kinetochore and kinetochore fiber.* **International Review of Cytology** 79:1-58, 1982.
- C. L. Rieder and E. D. Salmon. *Motile Kinetochores and Polar Ejection Forces Dictate Chromosome Position on the Vertebrate Mitotic Spindle.* **Journal of Cell Biology** 124:223-233, 1994.
- . *The vertebrate cell kinetochore and its roles during mitosis.* **Trends in Cell Biology** 8:310-8, 1998.

- A. D. Riggs and T. N. Porter. Overview of Epigenetic Mechanisms, in Cold Spring Harbor monograph series 32: Epigenetic mechanisms of gene regulation. Cold spring Harbor Laboratory Press, Plainview, New York, 1996.
- V. E. A. Russo, R. A. Martienssen and A. D. Riggs. Cold Spring Harbor monograph series 32: Epigenetic mechanisms of gene regulation. Cold spring Harbor Laboratory Press, Plainview, New York, 1996.
- R. Saffery, D. V. Irvine, B. Griffiths, P. Kalitsis, L. Wordeman and K. H. Choo. *Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins*. **Human Molecular Genetics** 9:175-185, 2000a.
- R. Saffery, D. V. Irvine, B. Griffiths, P. Kalitsis, L. Wordeman and K. H. A. Choo. *Human centromeres and neocentromeres show identical distribution patterns of > 20 functionally important kinetochore-associated proteins*. **Human Molecular Genetics** 9:175-185, 2000b.
- E. Schmid and M. Bauchinger. *Analysis of Primary Processes in the Formation of Acentric Fragments*. **Radiation and Environmental Biophysics** 17:143-149, 1980.
- F. Schrader. The Structure of The Kinetochore at Meiosis. 1939.
- D. D. Sears, J. H. Hegemann, J. H. Shero and P. Hieter. *Cis-Acting Determinants Affecting Centromere Function, Sister-Chromatid Cohesion and Reciprocal Recombination During Meiosis in Saccharomyces Cerevisiae*. **Genetics** 139:1159-1173, 1995.
- R. D. Shelby, O. Vafa and K. F. Sullivan. *Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites*. **Journal of Cell Biology** 136:501-13, 1997.
- A. C. Spradling, D. Stern, A. Beaton, E. J. Rhem, T. Lavery, N. Mozden, S. Misra and G. M. Rubin. *The Berkeley drosophila genome project gene disruption project. Single P-element insertions mutating 25% of vital drosophila genes [In Process Citation]*. **Genetics** 153:135-77, 1999.

- N. C. Steiner and L. Clarke. *A novel epigenetic effect can alter centromere function in fission yeast.* **Cell** 79:865-74, 1994.
- S. Stoler, K. C. Keith, K. E. Curnick and M. Fitzgerald-Hayes. *A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis.* **Genes & Development** 9:573-86, 1995.
- B. A. Sullivan and H. F. Willard. *Stable dicentric X chromosomes with two functional centromeres.* **Nature Genetics** 20:227-228, 1998.
- K. F. Sullivan, M. Hechenberger and K. Masri. *Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere.* **Journal of Cell Biology** 127:581-92, 1994.
- W. Sullivan, D. R. Daily, P. Fogarty, K. J. Yook and S. Pimpinelli. *Delays in anaphase initiation occur in individual nuclei of the syncytial Drosophila embryo.* **Molecular Biology of the Cell** 4:885-96, 1993.
- X. Sun, J. Wahlstrom and G. Karpen. *Molecular Structure of a Functional Drosophila Centromere.* **Cell** 91:1007--1019, 1997.
- C. E. Sunkel and P. A. Coelho. *The elusive centromere: sequence divergence and functional conservation.* **Current Opinion in Genetics & Development** 5:756-767, 1995.
- C. P. Swanson, T. Merz and W. J. Young. Cytogenetics. Prentice-Hall, Englewood Cliffs, N.J., 1967.
- P. B. Talbert and S. Henikoff. *Mapping the centromere of the X chromosome.* Annual Drosophila Research Conference 41, 2000.
- C. Tyler-Smith, G. Gimelli, S. Giglio, G. Floridia, A. Pandya, G. Terzoli, P. E. Warburton, W. C. Earnshaw and O. Zuffardi. *Transmission of a fully functional human neocentromere through three generations.* **American Journal of Human Genetics** 64:1440-4, 1999.

- O. Vafa and K. F. Sullivan. *Chromatin containing CENP-A and alpha-satellite DNA is a major component of the inner kinetochore plate.* **Current Biology** 7:897-900, 1997.
- O. Vafa, R. D. Shelby and K. F. Sullivan. *CENP-A associated complex satellite DNA in the kinetochore of the Indian muntjac.* **Chromosoma** 108:367-374, 1999.
- B. K. Vig, D. Latour and M. Brown. *Localization of Anti-CENP Antibodies and Alphoid Sequences in Acentric Heterochromatin in a Breast Cancer Cell Line.* **Cancer Genetics and Cytogenetics** 88:118-125, 1996.
- L. E. Voullaire, H. R. Slater, V. Petrovic and K. H. A. Choo. *A Functional Marker Centromere with No Detectable Alpha-Satellite, Satellite-Iii, or Cenp-B Protein - Activation of a Latent Centromere.* **American Journal of Human Genetics** 52:1153-1163, 1993.
- B. T. Wakimoto. *Beyond the nucleosome: epigenetic aspects of position-effect variegation in Drosophila.* **Cell** 93:321-324, 1998.
- P. E. Warburton and e. al. *Molecular cytogenetic Analysis of Eight Inversion Duplications of Human Chromosome 13q That Each Contain a Neocentromere.* **American Journal of Human Genetics** 66:2000.
- P. E. Warburton, C. A. Cooke, S. Bourassa, O. Vafa, B. A. Sullivan, G. Stetten, G. Gimelli, D. Warburton, C. Tyler-Smith, K. F. Sullivan, G. G. Poirier and W. C. Earnshaw. *Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres.* **Current Biology** 7:901-4, 1997.
- R. L. Warters and B. W. Lyons. *Variation in Radiation-Induced Formation of DNA Double-Strand Breaks as a Function of Chromatin Structure.* **Radiation Research** 130:309-318, 1992.
- M. J. White. *Animal Cytology & Evolution.* Cambridge University Press, Cambridge, 1954.

- H. F. Willard. *X Chromosome Inactivation, Xist, and Pursuit of the X-Inactivation Center*. **Cell** 86:5-7, 1996.
- B. C. Williams, T. D. Murphy, M. L. Goldberg and G. H. Karpen. *Neocentromere activity of structurally acentric minichromosomes in Drosophila*. **Nature Genetics** 18:30-37, 1998.
- D. Wise, L. Cassimeris, C. L. Rieder, P. Wadsworth and E. D. Salmon. *Chromosome Fiber Dynamics and Congression Oscillations in Metaphase Ptk2 Cells at 23-Degrees-C*. **Cell Motility and the Cytoskeleton** 18:131-142, 1991.
- D. A. Wise and B. R. Brinkley. *Mitosis in cells with unreplicated genomes (MUGs): spindle assembly and behavior of centromere fragments*. **Cell Motility and the Cytoskeleton** 36:291-302, 1997.
- P. Zhang and R. S. Hawley. *The genetic analysis of distributive segregation in Drosophila melanogaster. II. Further genetic analysis of the nod locus*. **Genetics** 125:115-27, 1990.