Gene Amplification in *Drosophila* Ovarian Follicle Cells as a Developmental Strategy and Model for Metazoan DNA Replication

By

Julie Michelle Claycomb

B.S. Biology
University of Pittsburgh
Pittsburgh, PA 15260

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Biology

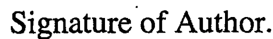
at the

Massachusetts Institute of Technology
[June 2004]
April, 2004

© 2004 Julie M. Claycomb. All rights reserved.

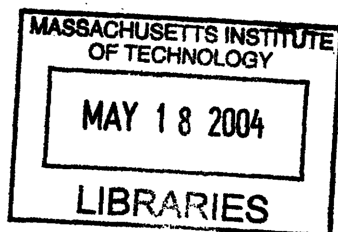
The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author. ....
Department of Biology
April 16, 2004



Certified by..
Terry L. Orr-Weaver
Professor of Biology
Thesis Supervisor

Accepted by..
Stephen P. Bell
Chair, Committee on Graduate Students
Department of Biology


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MAY 18 2004
LIBRARIES

SCIENCE

Gene Amplification in *Drosophila* Ovarian Follicle Cells as a Developmental Strategy and Model for Metazoan DNA Replication

by

Julie Michelle Claycomb

Submitted to the Department of Biology on April 16, 2004 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology

Abstract

The process of gene amplification in *Drosophila* ovaries provides a means of increasing the amount of template for transcription, thus increasing the amount of protein that can be made over a short developmental period. At a specific developmental point (egg chamber stage 10B-13), several clusters of genes encoding the eggshell (chorion) proteins in the follicle cells of each egg chamber are overreplicated 20 or 60 fold (for the *X* chromosome and third chromosome amplicons, respectively). Gene amplification is accomplished using the normal eukaryotic DNA replication machinery and a bidirectional DNA replication mechanism, and as such, is a powerful system for the study of metazoan DNA replication. Furthermore, the nature of the ovaries, with egg chambers of various ages arrayed in the order they were created, coupled with the use of cell biology, allows for the visualization of gene amplification at multiple timepoints in a single sample. We employed confocal and deconvolution microscopy to visualize the replication proteins ORC2, DUP/Cdt1, PCNA, and MCM2-7, as well as the nucleotide analog BrdU, at sites of gene amplification. These studies revealed that the BrdU staining pattern resolves from a focus of incorporation at the third chromosome locus in egg chamber stage 10B, to a coffee-bean structure in stage 11 egg chambers, to a double-bar structure in egg chamber stages 12 and 13. When coupled with quantitative real-time PCR calculations of copy number at the third chorion cluster during egg chamber stages 10B-13, these studies demonstrated that amplicon origin firing ends by stage 11 and that only the existing replication forks move out during stages 12 and 13 to produce the double bar BrdU pattern. The localization patterns of replication initiation and elongation factors also support this model. The initiation protein, ORC2 is only found in foci during stages 10A to 11, while the elongation factors PCNA and MCM2-7 resolved from foci at origins in stage 10B into the double bar staining structure representing replication forks in stages 12 and 13, similar to BrdU. We also observed that the replication initiation factor DUP/Cdt1 colocalized with BrdU throughout amplification, and resolved into double bars, suggesting that DUP/Cdt1 travels with replication forks during elongation. We hypothesize that DUP/Cdt1 may be necessary for the nuclear trafficking and/or the adherence of the MCM2-7 to replicating DNA. In sum, this work has increased our understanding of the process of gene amplification and has provided a powerful tool for the study of replication fork progression and the proteins involved, an aspect of replication that has proven difficult to examine *in vivo* in other systems.

Related BrdU studies revealed that there were two uncharacterized amplified regions in the follicle cells, thus we devised a comparative genomic hybridization microarray approach to systematically identify amplified portions of the genome. This approach identified the two uncharacterized amplicons, at cytological positions 62D5 and 30B10. Using FISH/BrdU co-labeling and real-time PCR, we verified that these regions were amplified over a 75-100kb region. The new amplicon DAFC-62D was shown to have a final origin firing in stage 13, a time when the other amplicons are only elongating. RNA *in situ* hybridization showed that the amplified genes were highly expressed, and that amplification was necessary for high levels of expression. Mutant analysis established that *yellow-g* is essential for proper eggshell formation and female fertility, and we hypothesize that *yellow-g* may be necessary for vitelline membrane crosslinking. This work extends the number of examples of gene amplification and model replicons available for study, and suggests that amplification may be a more widespread phenomenon throughout nature.

Thesis Supervisor: Terry L. Orr-Weaver
Title: Professor of Biology

Dedicated to

Alexander Ensminger

and

Joseph & Sandra Claycomb

Acknowledgements

There are a number of people who have contributed to my progress as a student over the years and have helped me to find this path. First, I would like to thank my advisor Terry-Orr Weaver for her guidance and support. Not only is she an outstanding scientist, but is also a strong role model for anyone wishing to pursue this type of career and have a family. I am always amazed at her ability to do so much on a daily basis, on very little sleep, nonetheless, and her ability to recount volumes of scientific knowledge about *Drosophila*.

I feel that I was very fortunate to be able to join the Orr-Weaver lab. The work here has been fascinating, and the people in the lab have been extremely engaging and supportive. I would like to thank Giovanni Bosco for motivating me to think critically, and to not forget the "big picture," as well as for teaching me to have fun with experiments, and to not be afraid to ask questions. Kim Dej, Irena Ivanovska, Astrid Clarke and Laurie Lee have been very supportive and extremely knowledgeable, and have provided many helpful discussions both in and out of the lab. I thank Janice Lee for her guidance and support when I was thinking of joining the lab and throughout graduate school, especially through the thesis writing process.

My thesis committee members, Steve Bell and Ilaria Rebay, have been wonderful mentors and a lot of fun. Their support and knowledge have been extremely valuable in helping to make graduate school a pleasant and productive experience for me. Along the same lines, perhaps I should now acknowledge Pepperidge Farms for making the delectable treats that kept my graduate committee members happy at each meeting! Thanks also to the David Koch Foundation for funding my work during the past year.

Since our late nights of studying Biology together at Pitt, Christin Cvetic has been a source of support, motivation, love, and fun! From helping me to choose an undergraduate research lab, to traveling the world and moving to Boston together, her friendship and support have never wavered, and for all of this I thank her.

It seems as though my classmate and friend, Tina Tootle, and I have gone through all the milestones of graduate school together. She has been a teacher, study buddy, a source of guidance, support, and motivation, and most of all, a good friend and shopping buddy. For all of this, and for helping me to see the brighter side of things, I thank her.

My fiancé, Alex Ensminger, has been a rock and an inspiration to me throughout the past two years. He encourages and helps to pick me up on those days when I don't feel like anything in life or lab is going right. He pushes me to think critically about my work and science in general, and has offered many helpful suggestions on experiments and my projects. He has helped me learn to not take myself too seriously and reminds me to enjoy life when I get too wrapped up in my work. Alex has been a balancing force in my life and for his friendship, motivation, and love, I thank him.

I would like to express my gratitude to my parents for their love and support throughout my life. They have driven me to pursue my goals and curiosities, and have helped me to traverse the ups and downs of both graduate school, and life. Thanks to my dad for teaching me, by example, the value of hard work and the importance of doing whatever job I set out to do well. Thanks to my mom for teaching me to temper my ambitions with patience and to see the lighter side in life. Thanks to my grandmother, Dale Claycomb and my friend, Donna Everhart for their support and for teaching me that nothing is impossible if you set your mind to it and work for it. Thanks to my entire family for supporting me and loving me, even when they didn't understand why I would want to move so far away or know what exactly I was doing in school. And finally, in response to the question all of you have asked me for years: yes, I am finally done with school...but somehow I think I will never stop being a student.

TABLE OF CONTENTS

Chapter One:
Introduction: Developmental Gene Amplification 7
The Strategy of Developmental Gene Amplification 8
The Mechanisms Used in Developmental Gene Amplification 10
cis-Control Elements Regulating Developmental Gene Amplification 17
Cell Cycle Controls on Gene Amplification 24
trans-Factors Known to Play a Role in Developmental Gene Amplification 28
Replication Fork Movement through Amplified Intervals 35
The Future of Developmental Gene Amplification 39
Summary of Thesis 40
References 43

Chapter Two:
Visualization of Replication Initiation and Elongation in *Drosophila* 50
Abstract 51
Introduction 52
Results 56
Localization of ORC2 and BrdU throughout chorion amplification 56
Quantitative real-time PCR measurement of DNA copy number along the third chromosome chorion amplicon 63
Localization patterns of PCNA and MCM2-7 during chorion amplification 66
The localization pattern of DUP/Cdt1 during chorion amplification 70
DUP/Cdt1 is necessary to localize MCM2-7 during amplification 76
Discussion 79
Materials and Methods 84
Acknowledgements 89
References 90

Chapter Three:
Gene Amplification as a Developmental Strategy: Isolation of Two Developmental Amplicons in *Drosophila* 97
Abstract 98
Introduction 99
Results 103
Identification of amplified genes on microarrays 103
Confirmation that DAFC-30B and DAFC-62D are follicle cell amplicons 104
Predicted amplified gene products 110

Expression patterns of the amplified genes 114
yellow-g is essential for proper eggshell formation 117
Amplification is necessary for gene expression 120
Discussion 124
Materials and Methods 129
Acknowledgements 133
References 134

Conclusions/Afterword 139
Conclusions 140
Rescue of *dup* Mutants 142
The role of Dup in elongation 142
Regulation of DAFC-30B and DAFC-62D 143
Other replication factors involved in gene amplification 144
The role of *yellow-g* and *yellow-g2* 145
Examination of other genes in DAFC-30B and DAFC-62D 147
Summary 148
References 148

Appendix One: 150
Generation and Characterization of Mutants in DAFC-30B

Appendix Two: 158
Expression of a *dup* Transgene Under UAS-Gal4 Control

Appendix Three: 177
Preliminary Studies of the Role of CDC6 in *Drosophila* Gene Amplification

Appendix Four: 188
Analysis of Vitelline Membrane Integrity in *yellow-g* and Replication Factor Mutants Using the Neutral Red Assay

Appendix Five: 196
Real-time PCR Determination of *ACE1* (DAFC-7F) and DAFC-30B Developmental Timing of Amplification

Introduction:

Developmental Gene Amplification

The Strategy of Developmental Gene Amplification

Developmental gene amplification is a DNA replication-based process whereby the genes from a given portion of the genome are replicated above the copy number of surrounding sequences, with the end result being an increase in the number of template molecules available for transcription. This leads to a situation in which large quantities of gene products can be produced over relatively short periods of time, to facilitate various developmental processes in the formation of a complex organism. Besides increasing the copy number of amplified genes, it is possible that the process of amplification promotes transcription by interactions between replication and transcription proteins or by creating an open chromatin conformation.

Although the fundamental concept of developmental gene amplification is similar to that which occurs in tumor cells, where cell cycle promoting genes, multi-drug resistance transporters, and other cancer-promoting genes are represented above their normal diploid copy number, we must emphasize that these two processes are distinct: developmental gene amplification occurs by differential DNA replication, initiated at specific points throughout the genome at strategic developmental times; tumor cell amplification occurs by a less well-defined mechanism that is likely to involve chromosomal recombination, initiated at a variety of genomic loci, and in response to largely unknown cues. This review will only deal with gene amplification in its developmental context, for review of amplification as it occurs in cancer, see (Stark et al., 1989).

Developmental gene amplification is initiated in a variety of organisms at specific stages of their development. The ribosomal RNA genes of the protist *Tetrahymena thermophila* are amplified during the development of the transcriptionally-active macronucleus (Yao et al., 1974). Similarly, ribosomal RNA genes are amplified in amphibian oocytes in order to stockpile the egg with the machinery necessary for rapid embryonic development (Brown and Dawid, 1968; Gall, 1968). Dipteran flies, including *Rhynchosciara americana* (Glover et al., 1982), *Bradysia hygida* (Laicine et al., 1984), and *Sciara coprophila* (Wu et al., 1993) all utilize gene amplification at multiple loci throughout the genome in the larval salivary glands, presumably for the production of large quantities of the structural proteins for the construction of the cocoon.

Another Dipteran fly, *Drosophila melanogaster*, undergoes amplification of four groups of genes in the ovarian follicle cells (Claycomb et al., 2004; Spradling, 1981; Spradling et al., 1980), with two of these gene clusters encoding the major structural proteins of the chorion (eggshell) (Spradling et al., 1980). The other amplified gene clusters in *Drosophila* follicle cells were recently identified by a comparative genomic hybridization (CGH) array approach, and have been found to encode a wide variety of proteins, including transporters, proteases, chitin binding proteins, and two putative enzymes, *yellow-g* and *yellow-g2*, thought to be necessary for crosslinking proteins of the vitelline membrane or eggshell (Claycomb et al., 2004). The observation that genes encoding enzymes, and not just those encoding ribosomal RNA or structural proteins, are amplified in *Drosophila* opens the possibility that developmental gene amplification may be a much more widely spread mechanism for coping with demands on gene expression throughout development than previously appreciated. Furthermore, CGH arrays provide

a systematic means to probe various tissues for developmentally important gene amplification events.

The Mechanisms Used in Developmental Gene Amplification

Although developmental gene amplification occurs in all of the organisms mentioned via a DNA replication-based mechanism, the details of how the amplification process occurs vary from organism to organism, both in terms of the position of the amplified region in relation to the rest of the genome (intra- vs. extrachromosomal), and in the mechanism used to replicate the sequences (rolling circle vs. semi-discontinuous replication). It is likely that each of the tissues has adopted a different means of performing gene amplification that is complementary to the differentiation state or developmental context of the particular cell type. For instance, amphibian oocytes utilize rolling circle DNA replication to amplify the rDNA genes from an extrachromosomal, circular DNA molecule (Gilbert and Dressler, 1968; Hourcade et al., 1973; Rochaix et al., 1974). It is possible that amphibians employ this extrachromosomal gene amplification strategy because carrying a highly-duplicated region of the chromosome in the genome of the oocyte could lead to chromosomal rearrangements and aberrations in the resultant embryo, which would be deleterious for its development. It would be necessary to possess the extrachromosomal rDNA in a circular molecule, because without telomeres, a linear molecule could not be maintained during amplification. This mechanism of gene amplification is unique among the organisms that do gene amplification, thus we will focus the remainder of the discussion on *Tetrahymena* and the Dipteran flies.

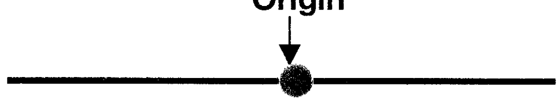
In *Tetrahymena* and the Dipteran flies, gene amplification occurs by semi-discontinuous replication, with repeated firings from initiation zones, containing several origins. 2-D (Delidakis and Kafatos, 1989; Heck and Spradling, 1990; Liang et al., 1993; Osheim et al., 1988; Yokosawa et al., 1999; Zhang et al., 1997) and 3-D (Liang and Gerbi, 1994) gel experiments, as well as electron microscopy of replication intermediates, have cumulatively demonstrated that amplification generates an onionskin structure of nested replication bubbles/forks in these organisms (Figure 1). In the Dipteran flies amplification occurs within the polytene chromosomes of the salivary gland or ovarian follicle cells, while in *Tetrahymena*, gene amplification occurs extrachromosomally.

In the development of the *Tetrahymena* macronucleus from one of the micronuclei after conjugation, the genome is fragmented and rearranged in a very specific, reproducible manner, where intervening non-coding sequences are removed. *de novo* telomere synthesis also occurs to generate stable linear chromosomes. During this process, the 10.3kb rDNA locus is specifically excised, converted to a ~21kb head-to-head palindrome, and telomeres are added. The entire remaining genome is copied to a ploidy of approximately 45C, and then over the course of twelve hours, the rDNA chromosomes are preferentially amplified up to 10,000-fold (reviewed in Prescott, et. al. 1996). It seems there are two phases of rDNA amplification, with the first phase occurring to a low level and perhaps within the native chromosomes (Kapler and Blackburn, 1994; Ward et al., 1997). Excision and palindrome formation are necessary for the second, pronounced amplification phase (Kapler and Blackburn, 1994; Kapler et al., 1994). During this phase, amplification initiates from multiple positions within a central region of the palindrome, the 5' Nontranscribed Spacer region (5'NTS), and it is

Figure 1. Gene amplification occurs by repeated firings from replication origins and bi-directional replication fork movement to form an onionskin structure. The onionskin structure may be intrachromosomal, as in the salivary glands of *Sciara*, *Rhynchosciara*, and *Bradysia* or the follicle cells of *Drosophila*, or it may occur extrachromosomally, as in the macronucleus of *Tetrahymena*. *cis*-acting sequences (see Figure 2) and many *trans*-factors contribute to the regulation of gene amplification. Here, initiation factors that recognize origins are shown in green and the factors traveling with the replication forks are red. In *Drosophila*, after a period of origin firing coupled with replication fork movement, origins stop firing and only the existing replication forks move outward. [Modified from Bosco, G. and Orr-Weaver T.L. (2002). Regulation of the cell cycle during oogenesis and early embryogenesis in *Drosophila*. In Regulation of Gene Expression at the Beginning of Animal Development, M. DePamphilis, ed. (Amsterdam: Elsevier), pp. 107-154.]

Gene Amplification via Origin Re-firing and Bidirectional Replication Fork Movement

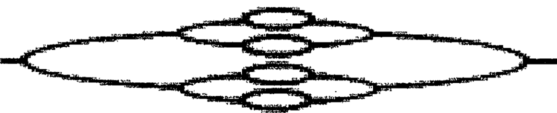
Origin

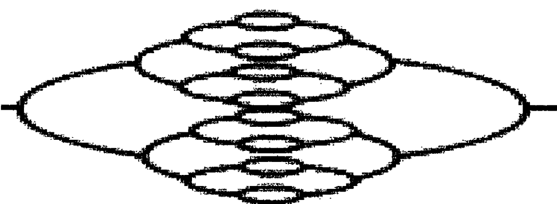


Origins Fire

Replication Forks Move Out

Origins Refire, Forks Progress





(Elongation Only-Drosophila)

● Replication Initiation Protein (ORC)

● Replication Elongation Protein (PCNA)

not clear whether high molecular weight intermediates observed on 2-D gels and thought to represent onionskin structures are stably maintained throughout further development or resolved in some way (Zhang et al., 1997). It appears that at least some portion of the amplified molecules separate from each other, as FISH studies demonstrate the presence of several hundred rDNA loci in nucleoli throughout the macronucleus during amplification stages (Ward et al., 1997). After macronuclear development, the rDNA palindromic chromosomes are each replicated one time per cell cycle, during vegetative growth.

Why might it be useful for *Tetrahymena* to perform gene amplification on a DNA molecule that has been separated from the remainder of the genome? Perhaps it is necessary to remove the rDNA locus from surrounding sequences that serve to inhibit gene amplification by excising the rDNA region from the rest of the chromosome. It has also been suggested that palindrome formation may serve to promote amplification by placing sequences that positively regulate origin firing in closer proximity to the origin. Alternatively, there may be no inherent benefit to the organism to perform gene amplification extrachromosomally, and this may simply be a by-product of the substantial genomic rearrangements occurring throughout the genome during macronuclear differentiation. Dipteran flies, on the other hand, may be able to tolerate the intrachromosomal structures generated by the amplification process due to the terminally differentiated state of the larval salivary gland or ovarian follicle cells. Both of these cell types are lost during further development, and as these cell types are nondividing, genomic aberrations accumulated during developmental gene amplification would not be passed on to daughter cells.

During gene amplification in Dipteran flies, a gradient of copy number is generated within the amplified region, with sequences closest to the origin of DNA replication being amplified, and thus expressed, to the greatest extent, and those proximal and distal to the origin amplified to progressively decreasing levels. Amplification levels vary between amplified regions, organisms, and methods used to determine copy number. In Rhynco*Sciara*, the salivary gland puff C3-22 is amplified over a distance of about 50kb to a maximum of 32-fold, and the C8 puff is amplified over about 60kb to a level of 16-fold (Glover et al., 1982; Penalva et al., 1997). In *Bradysia*, the amplified region spans 18kb at the salivary gland puff C4, which is amplified to 21-fold (Coelho et al., 1993; Monesi et al., 1995). The B10 puff in *Bradysia* is also amplified 10-fold (Fontes et al., 1992). *Sciara* amplifies the salivary gland puff II/9A 18-fold over at least 35kb, and amplifies the puff II/2B 17-fold (Liang et al., 1993; Wu et al., 1993). Rhynco*Sciara*, *Bradysia*, and *Sciara* all require about 15 hours to perform gene amplification. Amplification is completed before transcription begins in *Sciara*, and the same may be true in Rhynco*Sciara* (Gabrusewycz-Garica, 1971; Santelli et al., 1991; Wu et al., 1993).

In the follicle cells of *Drosophila*, amplification occurs over regions similar in size to the other Dipteran flies. Amplified regions each span approximately 75-100kb, and the process requires about 10-12 hours to reach completion, occurring in egg chamber developmental stages 10B, 11, 12, and 13 (Claycomb et al., 2004; Claycomb et al., 2002; Spradling, 1981; Spradling, 1993). Amplification levels have been determined to reach 18-20 fold at DAFC-7F and 64-80 fold at DAFC-66D by quantitative Southern blotting (Delidakis and Kafatos, 1989; Spradling, 1981). This is in contrast to the 12-fold and 30-fold observed for DAFC-7F and 66D, respectively, by quantitative real-time PCR

(Claycomb et al., 2002). This difference in quantitation of fold amplification is likely due to differences in the sensitivity of the two assays, with the fluorescence used in the real-time assay, as well as the size of the regions detected (50 to 70 basepairs by real-time PCR vs. several kilobases by quantitative Southern blots), and the ability to quantitate copy number within the linear range of the assay in real time all contributing to the differences.

The chorion DAFCs are amplified to a greater extent than the newly characterized DAFC-30B and 62D, which amplify to approximately 4 and 6.5-fold, respectively, by the real-time PCR method (Claycomb et al., 2004). Amplification for the most highly amplified region, DAFC-66D, begins during the final endocycle in stage 9 of egg chamber development (Calvi et al., 1998; Royzman et al., 1999). By stage 10B, all of the loci have begun amplification.

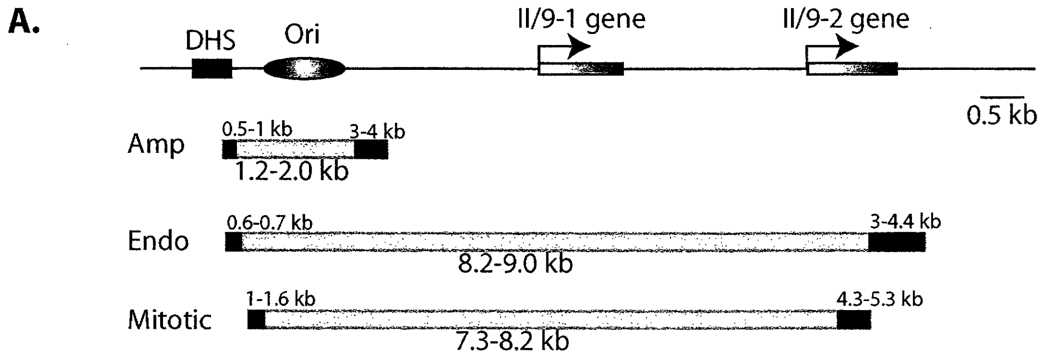
One of the most interesting aspects of amplification at the *Drosophila* loci is that replication initiation and elongation occur during distinct phases of egg chamber development. Real-time PCR measurement of copy number shows that during stages 10B and 11 of egg chamber development, which last 6 hours and 45 minutes respectively, origin firing coupled with replication fork movement occurs. In stages 12 and 13, which each last one to one and one-half hours, only elongation from existing forks occurs at DAFC-7F, 66D, and 30B (Claycomb et al., 2002). DAFC-62D is an exception, in that an additional round of origin firing occurs during stage 13 (Claycomb et al., 2004). It has been proposed that this additional round of origin firing may serve to increase the copy number of the genes closest to the origin, *yellow-g* and *yellow-g2*, to provide a burst of gene expression very late in egg chamber development.

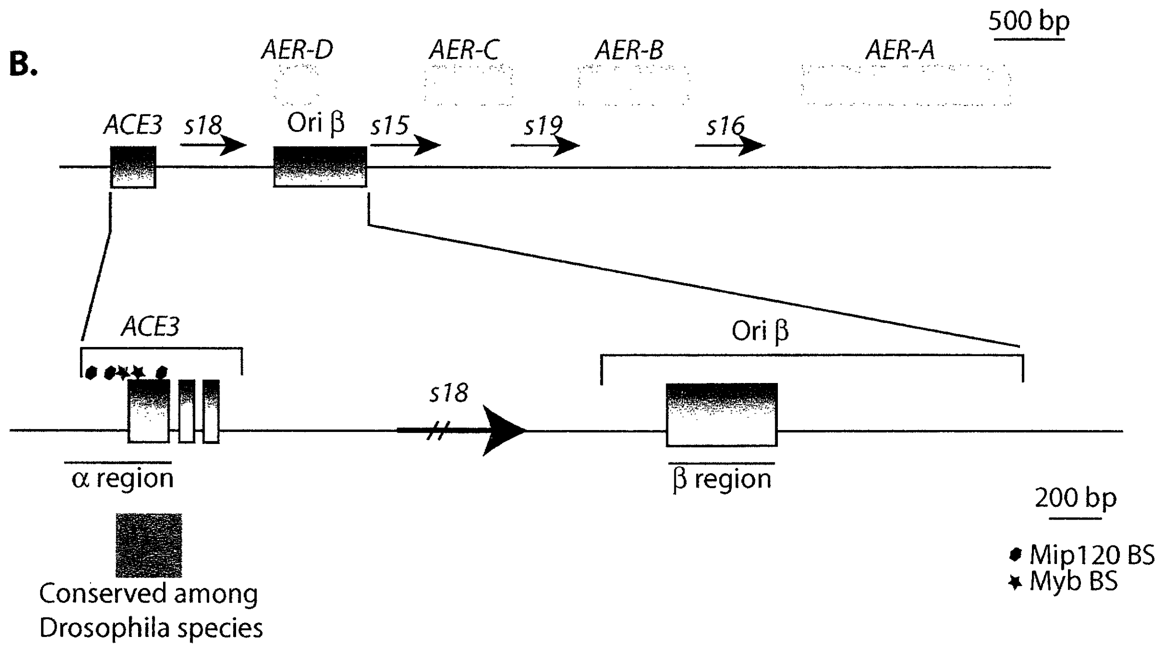
***cis*-Control Elements Regulating Developmental Gene Amplification**

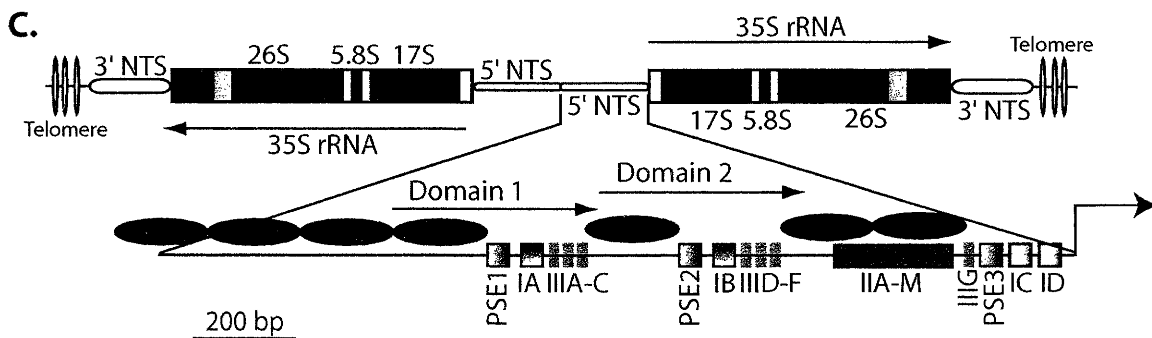
In the organisms for which the replication properties of gene amplification have been studied, a common characteristic has emerged in terms of the cis-regulatory sequences of the amplified regions. There is a zone from which replication initiates, containing usually one to two preferred origins and several auxiliary origins. These origins are regulated by multiple replication determinant sequences that on their own are not origins of DNA replication. Some of the replication determinant sequences serve as stimulatory elements, while others may inhibit replication. In this section we will review, on a per organism basis, what is known about the preferred origins of replication and replication determinants governing gene amplification.

In the *Sciara* salivary puff II/9A, several origins reside within an initiation zone. 2-D and 3-D gel analyses indicate that initiation occurs in a zone extending over 5.5kb, and within this region a preferred 1kb portion accounts for the majority of the origin firings (Liang and Gerbi, 1994; Liang et al., 1993) (Figure 2A). Furthermore, the precise nucleotide within the 1kb region at which DNA synthesis initiates has been isolated by the technique of Replication Initiation Point (RIP) mapping, and both recombinant Origin Recognition Complex (ORC) protein from *Drosophila* and endogenous *Sciara* ORC2 have been shown to bind to an 80bp segment adjacent to this initiation site (Bielinsky et al., 2001). Although much is known about the sites of initiation in amplicon II/9A, little is known about the sequences controlling or stimulating amplification and Amplification Control Elements like *ACE3* in *Drosophila* (see below) have yet to be isolated, although a DNAseI hypersensitive site at the left boundary of the initiation zone has been

Figure 2. Diagrams of the known regulatory sequences within the amplified loci of *Sciara, Drosophila,* and *Tetrahymena.* A. The salivary puff II/9A from *Sciara* initiation zone is bounded by a DNaseI hypersensitive site (DHS) on the left. The right boundary is dynamic throughout development, and the preferred regions of initiation in mitotic, pre-amplification (endocycle), and amplification stages are shown below the diagram of the II-9A region. The preferred origin region during amplification that was previously identified is shown (Ori), as are the two genes within the region, II/9-1 and II/9-2. B. The *Drosophila* third chromosome amplicon, DAFC-66D contains two major and at least four minor regulatory elements. The previously 320bp *ACE3* region has been further dissected and contains three segments necessary for amplification. Within *ACE3* is a region with homology (the α region) to a portion (the β region) of the preferred origin region, Ori β. *ACE3* possesses two Myb and three Mip120 binding sites and a region of high conservation among several Drosophilid species. Ori β was previously narrowed to 884bp, but has recently been further limited to an essential core region. The elements *AER-A* to *D*, like *ACE3*, are stimulatory to replication. The chorion protein genes are shown, as well (arrows). C. The *Tetrahymena* rDNA palindrome initiates replication within the 5' NTS. Within the 5'NTS are two nucleosome-free repeated Domains, (D1 and D2), three directional pause sites (PSEI-III) that regulate replication fork movement, Multiple Type I elements (of which, Type IA and IB are origins of replication), as well as Type II and III elements, some of which play stimulatory roles in replication and/or transcription. The 35S rRNA gene is shown, with the spliced form of the rRNA in black. [Adapted from Lunyak et al., 2002; Lu et al., 2001; J. Tower unpublished results; and Mohammad et al., 2003.]


A.
DHS
Ori
II/9-1 gene
II/9-2 gene
0.5 kb
Amp
0.5-1 kb
3-4 kb
1.2-2.0 kb
Endo
0.6-0.7 kb
3-4.4 kb
8.2-9.0 kb
Mitotic
1-1.6 kb
4.3-5.3 kb
7.3-8.2 kb


B.
500 bp
AER-D
AER-C
AER-B
AER-A
ACE3
s18
Ori β
s15
s19
s16
ACE3
Ori β
s18
α region
β region
200 bp
Conserved among
Drosophila species
Mip120 BS
Myb BS


C.
3' NTS
26S
5.8S
17S
5' NTS
35S rRNA
Telomere
Telomere
5' NTS
17S 5.8S
26S
3' NTS
35S rRNA
Domain 1
Domain 2
PSE1
IA
IIIA-C
PSE2
IB
IIID-F
IIA-M
IIIG
PSE3
IC
ID
200 bp

postulated to play a special regulatory role (Urnov et al., 2002). In the related Sciarid fly, *Rhynchosciara*, 2-D gel analyses demonstrated that replication initiates in the salivary puff C3 from at most 3 sites in a zone of about 6kb, and that this zone resides approximately 2kb upstream of the amplified gene *C3-22* (Yokosawa et al., 1999).

Using nascent strand analysis and quantitative PCR, the zone of initiation in puff II/9A has been determined in mitotic, endocycling, and gene amplification developmental stages (Lunyak et al., 2002) (Figure 2A). These data indicate that the zone of initiation within II/9A spans approximately the same distance in mitotic and endocycling stages. Subsequently, the zone from which initiation can be detected becomes restricted to the previously discussed 1.2-2.0kb region during gene amplification. The initiation zone of each of these stages possesses the same left-hand boundary, while the right-hand boundary is static. This is the first demonstration that an origin used during gene amplification resides within the same region utilized for initiation during a mitotic cell cycle, yet indicates that the boundaries of initiation set up during embryogenesis can change throughout development.

The *Drosophila* developmental amplicon DAFC-66D is the best understood of the four follicle cell amplicons in terms of the sequences regulating replication. 2-D gel analysis has identified three potential replication origins within the peak amplified region, with one of these, the 884bp sequence downstream of the *s18* chorion protein gene known as Ori β, being the preferred site of origin activity (70-80% of origin firings occur in the Ori β region) (Delidakis and Kafatos, 1989; Heck and Spradling, 1990) (Figure 2B). A number of studies have also delineated the evolutionarily well-conserved 320bp

Amplification Control Element, *ACE3*, which is located approximately 1.5kb upstream of Ori β, to the 5' end of the *s18* gene (de Cicco and Spradling, 1984; Delidakis and Kafatos, 1989; Heck and Spradling, 1990; Orr-Weaver et al., 1989; Spradling et al., 1987; Swimmer et al., 1990). The X chromosome amplified region, DAFC-7F also contains an *ACE* element (*ACE1*) that is important for the amplification of this gene cluster (Spradling et al., 1987).

Transgenic studies and 2-D gel analyses have demonstrated that *ACE3* itself does not function as an origin of DNA replication (Delidakis and Kafatos, 1989; Heck and Spradling, 1990). However, *ACE3* is necessary in cooperation with Ori β to achieve high levels of gene amplification, as an insulator element (SHWBS) placed between *ACE3* and Ori β in transgenes nearly eliminates amplification. Removal of this insulator element by FLP/FRT-mediated recombination then restores amplification (Lu et al., 2001). Furthermore, elimination of either *ACE3* or Ori β from transgenic constructs that contained the region encompassing *ACE3* and Ori β dramatically reduced amplification levels from the transgene, indicating that together, *ACE3* and Ori β are necessary and sufficient to drive developmental amplification (Lu et al., 2001).

From other transgenic studies, *ACE3* is essential for the amplification of small constructs from the DAFC-66D region, but appears to be dispensable in the amplification of larger transgenes, suggesting a level of redundancy in replication stimulatory elements of the region around *ACE3* (Carminati et al., 1992; Delidakis and Kafatos, 1987; Orr-Weaver et al., 1989). Four such replication stimulatory elements have been discovered, named <u>A</u>mplification <u>E</u>nhancing <u>R</u>egion, *AER-A* through *D*, of which, *AER-D* resides within Ori β (Delidakis and Kafatos, 1987). Within AER-D/Ori β and *AER-C* are ten out

of eleven basepair matches to the Saccharomyces cerevisiae ARS consensus sequence, that serves as an essential part of the origin of replication in yeast (Heck and Spradling, 1990). However, the significance of this sequence similarity has not yet been determined, and notably, the yeast origin ARS1 is incapable of driving amplification when substituted for Ori β on transgenes (J. Tower, unpublished results).

Recently, the sequence requirements of both *ACE3* and Ori β have been further delineated by transgenic deletion and protein-binding studies (J. Tower, unpublished results). Deletion studies of Ori β demonstrate that the 5' 140bp of Ori β plus the adjacent downstream 226bp make up the core region of Ori β, and this core region is necessary and sufficient to induce amplification of transgenes. A further 140bp at the 3' end of the 840 bp may also have a stimulatory effect on amplification. Within the 226bp of the Ori β core is a region (denoted the β region) that has high A/T content and significant homology to a 5'portion of *ACE3* (known as the α region). Deletion studies of *ACE3* have also determined that a 142bp highly evolutionarily conserved "core" region of *ACE3* is responsible for the majority of *ACE3*'s replication stimulatory activity. Within the *ACE3* core, binding sites for the oncoprotein, Myb and Mip120 (Myb Interacting Protein 120, formerly p120) have been identified (Beall et al., 2002). Deletion of the Myb or p120 binding sites within the *ACE3* core from transgenes resulted in nearly no amplification from the transgenes relative to the non-deleted control transgenes (Beall et al., 2002). These results indicate that the Myb and at least one of the p120 binding sites are necessary for amplification.

The Origin Recognition Complex has been shown to preferentially bind to A/T rich sequences in many species, and chromatin immunoprecipitation, in vitro binding,

immunofluorescence, and transgene studies have demonstrated that the replication initiation protein ORC2 binds directly to both *ACE3* and Ori β during gene amplification (Austin et al., 1999; Bosco et al., 2001; Chesnokov et al., 1999; Claycomb et al., 2002; Royzman et al., 1999). It has been suggested that *ACE3* and Ori β serve as nucleating sites for ORC to spread along the chromatin, thus influencing the ability of the region to replicate (Austin et al., 1999; Lu et al., 2001). By immunofluorescence, transgenes of *ACE3* multimers are capable of recruiting ORC2 in vivo (Austin et al., 1999), while insulated transgenes (those buffered by SHWBS) containing only *ACE3*, s18, and Ori β are incapable of recruiting visible ORC2, even though they amplify, and by mutant analysis, this amplification is dependent on the *orc2* gene product (Lu et al., 2001). The addition of extensive sequences at the 5' of *ACE3* and the 3' of Ori β to these insulated transgenes allows the recruitment of ORC2 to visible foci, so it seems that a certain threshold amount of ORC2 must be recruited and spread along the amplified region before it can be detected. This recruitment of ORC2 to visible foci may require the activity of the *chiffon/dbf4-like* gene product, see below (J. Tower, unpublished results).

Cumulatively, these data demonstrate that *ACE3*, the *AER*s, and Ori β are functionally separable, but act cooperatively to drive gene amplification. Although Ori β serves as a robust origin of replication initiation during gene amplification, we should note that it has not been conclusively demonstrated whether the DAFCs contain origins of DNA replication that function during the archetypal cell cycle.

Developmental gene amplification in *Tetrahymena* initiates from the center of the rDNA palindrome and moves outward, toward the telomeres. 2-D gels, mutant analyses, and transgenic studies show that initiation occurs from multiple sites within the 5'NTS,

and the origins of replication reside in the nucleosome free ~430bp repeated elements Domains 1 and 2 (Figure 2C). 2-D gel analysis shows that the preferred sites for initiation are from within or near the conserved type IA and IB elements within these two domains, and deletion analyses show that Domains 1 and 2 are non-redundant (Reischmann et al., 1999; Zhang et al., 1997). Although the promoter of the rRNA genes, located downstream of the 5'NTS, contains Type I and Type II conserved elements that regulate origin firing, it does not function as an origin on its own (Blomberg et al., 1997; Gallagher and Blackburn, 1998; Zhang et al., 1997). During amplification, replication initiates from within the 5'NTS on both sides of the palindrome (Zhang et al., 1997). This is in contrast to replication in vegetatively cycling cells, in which replication initiates from only one of the two 5'NTS regions of the rDNA palindrome. Notably, in both amplification and vegetative, or cell cycle, replication the sites of replication initiation within the 5'NTS are the same (Cech and Brehm, 1981; MacAlpine et al., 1997; Zhang et al., 1997). These data demonstrate the similarities between amplicon arrangement in *Tetrahymena* and *Drosophila*, with regulatory sequences that are not themselves origins mediating effects over distances. There are also similarities between *Tetrahymena* and *Sciara*, in terms of their choice in origin usage throughout various cell cycles.

Cell Cycle Controls on Gene Amplification

Perhaps the most confounding question about gene amplification is how does the cell suspend the rule of replicating DNA once and only once per cell cycle to allow re-firing of amplification origins? The answer may lie in the cell cycle preceding gene

amplification: an endocycle (reviewed in Edgar and Orr-Weaver, 2001). In *Tetrahymena* and the Dipteran flies, the cells in which amplification occurs all undergo an endocycle, a type of modified cell cycle in which synthesis and gap phases alternate with no intervening mitosis, prior to the onset of developmental gene amplification. *Tetrahymena* macronuclei reach a ploidy of approximately 45C, *Sciara* salivary gland cells reach 8192C, and *Drosophila* follicle cells achieve 16C before amplification begins.

In *Drosophila*, entry into the endocycle is dependent on Notch signaling and requires the activity of the anaphase promoting complex activator, fizzy-related/cdh1, to diminish the levels of the mitotic cyclins (Deng et al., 2001; Lopez-Schier and St Johnston, 2001). This serves to eliminate mitosis from the endocycle, and subsequently, the oscillation of Cyclin E and its inhibitor Dacapo causes the periodicity of S and G phases, with a drop in Cyclin E levels being required for the re-setting of origins used in the endocycle (Lilly and Spradling, 1996). Perhaps the down-regulation of mitotic cyclins, known to be inhibitors of origin re-setting and DNA synthesis, during the endocycle sets the stage for gene amplification to occur, and it is only when the mitotic cyclins have been appropriately titrated out of these cells by a series of endocycles that amplification can begin. However, this cannot be the entire story, as not all sequences are re-replicated during gene amplification as they are (except for the peri-centric heterochromatic sequences) in an endocycle, so there must be a way to preferentially promote re-replication at some loci over others.

It has been suggested that Cyclin E may be a part of this regulation. High Cyclin E levels persist in egg chamber stages 9 and 10A, when endocycles cease, and may serve to inhibit the formation of pre-replication complexes at various genomic origins, thus

inhibiting further endocycles (Calvi et al., 1998). In this model, Cyclin E also plays the role of promoting origin firing at the amplified loci, possibly by acting positively on an as yet uncharacterized amplification factor. Recent work shows that follicle cells may also enter gene amplification in a late S/G2-like state, based on having high levels of phospho-histone H1 staining, which appears to be a marker of G2 and Cyclin E/CDK2 activity in *Drosophila* (G. Bosco, unpublished results). This knowledge may provide clues in the future as to what additional factors are involved in overriding replication constraints.

Another important question with regards to gene amplification is: how are the number of firings from each amplified region actually regulated? Again, from *Drosophila*, there may be a link to the endocycle and the factors that control it. Mutations in the transcription factor complex E2F/DP/Rb (known as Rbf in *Drosophila*), a critical regulator of G1/S progression in the canonical cell cycle, have differential effects on gene amplification. Mutation of *rbf* leads to pleiotropic effects, in that follicle cells of a given egg chamber lose their developmental synchrony during gene amplification and develop mosaically (Bosco et al., 2001). Some follicle cells undergo gene amplification to levels higher than normal, others proceed into a full ectopic endocycle and reach a ploidy of 32C (Bosco et al., 2001; Cayirlioglu et al., 2003), while others seem to do both gene amplification and an endocycle concurrently (Bosco et al., 2001). Some mutant cells display ORC2 staining, but no BrdU incorporation, indicating the presence of a gap-like phase and confirming that cells are undergoing a true endocycle and not just a partial ectopic S phase. These data indicate that Rbf is necessary for the switch from endocycles to amplification and for restraining the number of origin

firings during amplification. Although this effect may be partially mediated through the transcriptional activity of Rbf (see below), it is certainly not solely due to transcription, as E2F1, DP, Rbf, and ORC2 are found in a complex at *ACE3* (Bosco et al., 2001).

dp mutants display a phenotype similar to *rbf* mutants, in that they undergo an extra endocycle to reach 32C (Cayirlioglu et al., 2003), with ORC dispersed throughout the nucleus (Royzman et al., 1999). However, *dp* mutants subsequently display decreased levels of gene amplification. *e2f1* mutants in which the DNA-binding domain is disrupted display decreased amplification and no ORC localization, while *e2f1* mutants in which the Rbf (and apparently the ORC2) interaction domain is removed have increased amplification with seemingly normal ORC localization (Bosco et al., 2001; Royzman et al., 1999). From these data, we can conclude that Rb and DP are necessary, perhaps as some sort of copy number counting mechanism, to end the endocycle and properly switch to gene amplification. Furthermore, mutant phenotypes and the fact that these four proteins form a complex at *ACE3* suggests that they inhibit origin firing at the amplicons until the appropriate time in development. In this model, DP and E2F1 may act as part of a switching mechanism, and when not restrained by Rbf, can promote amplification.

E2F2 plays a transcriptional role in regulating amplification. *e2f2* mutants exit the endocycle properly, but display decreased amplification and subsequently slip back into an ectopic S phase that is not a true endocycle (Cayirlioglu et al., 2001; Cayirlioglu et al., 2003). ORC is mislocalized throughout the nucleus and the transcript levels of ORC and a number of replication factors are increased in the mutant, as they are in the *rbf* mutant (Cayirlioglu et al., 2001; Cayirlioglu et al., 2003). A direct interaction between

E2F2 and ORC2 has not yet been found, thus the current conclusion is that the phenotype observed in the *e2f2* mutants is mediated by a transcription repressor function, where E2F2/Rbf complexes maintain replication factors used in amplification at a critical level. In this way, having too much of particular replication factors could lead to an ectopic S phase, as is the case when ORC1 is overexpressed by heat shock (Asano and Wharton, 1999). Futhermore, perhaps the levels of replication initiation factors decrease with more and more initiation events, leading to a depletion of protein pools and a halt to origin firing in stage 11 of egg chamber development.

In contrast to *Drosophila*, where it seems that the endocycle must end for amplification to begin properly and there is a chronological link between the two, *Tetrahymena* seems to be able to undergo amplification without ending or doing a complete endocycle (Kapler and Blackburn, 1994). A starved *Tetrahymena* can arrest its endocycle with 4-8C ploidy, while amplification continues as normal (Allis et al., 1987). This indicates that the two are not interdependent, yet does not preclude the possibility that the endocycle must first start for amplification to initiate.

It is clear that much additional work must be done to delineate the cell cycle controls dictating gene amplification, and to determine the link between an endocycle and developmental gene amplification. The number of cases of gene amplification, however, is too low at this point to be certain that endocycling is a strict pre-requisite for gene amplification.

trans-Factors Known to Play a Role in Developmental Gene Amplication

Genetic, biochemical, and cell biological approaches have clearly demonstrated

that the proteins involved in DNA replication during a normal cell cycle are also involved in replication during gene amplification. A brief summary of replication factors and their functions will be given here, with emphasis on those proteins implicated in gene amplification. For a complete review of the factors involved in eukaryotic DNA replication, see (Bell and Dutta, 2002).

The six member Origin Recognition Complex (ORC) recognizes and binds to specific sequences (the ARS Consensus Sequence or ACS) within the origins of *Saccharomyces cerevisiae* chromosomes (Bell and Stillman, 1992). In metazoans, ORC does not appear to bind in a sequence-specific manner, and is instead recruited to origins by some other means (Remus et al., 2004; Vashee et al., 2003). Once ORC is bound to origins of replication, it then recruits other replication machinery including the CDC6 and DUP/Cdt1 proteins, which in turn load the putative replication fork helicase complex, MCM2-7 (Aparicio et al., 1997; Ishimi, 1997; Labib et al., 2001).

Upon the loading of MCM2-7 onto orgins of replication, the origins are said to be "licensed," or competent for the initiation of replication (Blow and Laskey, 1988). Additional proteins are then recruited to the origins, including CDC45 and MCM10, which are required for origin firing and have been shown to travel with the replication forks (Aparicio et al., 1999; Merchant et al., 1997; Tercero et al., 2000; Wohlschlegel et al., 2002), as are the proteins which function solely at the replication forks, including the single-stranded DNA binding protein RPA, the primase Polα, the clamp loader RFC complex, the polymerase processivity factor PCNA, DPB11, and the replicative polymerases Polδ/ε (see Bell and Dutta, 2002). The activity of CDKs (including Cyclin E/CDK2) and another kinase composed of Cdc7 and Dbf4 (DDK) regulates origin firing

as well, with the possible targets of DDK being MCM2-7 and CDC45 (Lei et al., 1997; Zou and Stillman, 2000). Although this listing of the factors involved in replication and their functions is not complete, the proteins listed are representative of the major players in gene amplification.

As mentioned above, Chromatin Immuno-precipitation (ChIP) demonstrates that the *Sciara* II/9A origin is bound in vivo by a protein that cross-reacts with anti-XlORC2, and is presumably the *Sciara* ORC2 homolog (Bielinsky et al., 2001). This was an important discovery, as it demonstrates that amplification in *Sciara* is likely to be controlled by the same machinery that controls DNA replication in a canonical cell cycle. However, little is known about other factors involved in the process and a dearth of genetic approaches hinders the ability to assess the role of replication factors in *Sciara* amplification.

Recent progress has been made in determining the proteins necessary for gene amplification in *Tetrahymena.* Several proteins or complexes with differential DNA binding activities within the initiation zone have been purified, named TIF1-4 (Type I interacting Factor) (Mohammad et al., 2000), and one of these, *tif1*, has been cloned (Saha et al., 2001). TIFI is a single-stranded Type I element binding factor which possesses limited homology to a transcription factor in plants, p24. TIF1 copurifies with another protein with helicase activity and TIF1's binding to Type I elements has been shown to modulate the activity of other proteins, TIF2 and TIF3 (Saha et al., 2001).

The understanding of *Tetrahymena* DNA replication was advanced when the six-member TIF4 complex, containing a protein that appears to be ORC2, was identified based on its ability to bind Type IB single-stranded origin DNA (the T-rich strand,

specifically) in an ATP and MgCl2-dependent manner (Mohammad et al., 2003). The p69 component of TIF4 was shown to cross-react specifically with anti-ORC2 sera generated to ORC2 of several different species, and this cross-reactive p69 co-purifies with TIF4 activity. Immunofluorescence of p69 in vegetatively cycling cells shows that p69 levels in the nuclei peak when TIF4 activity is at its highest, at a time period when there is greatest overlap between macro- and micronuclear S phases. In the development of the macronucleus after conjugation, p69 levels remain high in the macronuclei as they undergo gene amplification, while staining is lost at the same time from the silent micronuclei. In sum, it appears that *Tetrahymena* possesses a functional homolog of ORC2, that plays a role in both cell cycle DNA replication and amplification, and thus the process of gene amplification in *Tetrahymena* is likely to utilize the same machinery as vegetative or archetypal cell cycle DNA replication.

Genetic, cell biological, and biochemical approaches have all contributed to an understanding of the factors involved in *Drosophila* gene amplification. Thus far, all of the components known to play a role in gene amplification are homologs of conserved replication or transcription factors, indicating once again that gene amplification relies on the archetypal DNA replication machinery. Hypomorphic mutations that lead to female sterility, disrupted eggshells, and severely decreased amplification, as measured by incorporation of BrdU or Quantitative Southern blotting have been identified in several genes encoding replication factors in *Drosophila* (Orr et al., 1984). Among this group are mutations in the *orc2, chif/dbf4-like, pcna/mus209, dup/cdt1*, and *mcm6* genes, all of which encode essential components of the replication machinery (Henderson et al., 2000;

Landis et al., 1997; Landis and Tower, 1999; Schwed et al., 2002; Underwood et al., 1990; Whittaker et al., 2000).

Replication factors and BrdU incorporation can be visualized at loci undergoing gene amplification by immunofluorescence microscopy. Immunofluorescence microscopy studies have shown that ORC1, 2, 5, CDC45, PCNA, MCM2-7, DUP/Cdt1 localize to sites of gene amplification, albeit in different patterns depending on their site of action, at either the origin or replication forks (Asano and Wharton, 1999; Austin et al., 1999; Claycomb et al., 2002; Loebel et al., 2000; Royzman et al., 1999; Whittaker et al., 2000). ORC is the first of the replication initiation factors to localize, and does so in stage 10A (Royzman et al., 1999). By stage 10B other initiation factors have localized, and by stage 11, ORC2 is lost from the amplified loci (Claycomb et al., 2002). Those factors involved in both initiation and elongation, MCM2-7 and DUP/Cdt1, or elongation alone, PCNA, localize first as foci and then resolve into a double bar structure, indicative of replication fork movement.

Some replication factors may have adopted new roles for their involvement in gene amplification. It seems that the Dbf4 homolog in *Drosophila*, Dbf4-like, may play a role in stabilizing the interaction of ORC with the chromatin, as ORC2 does not localize in to foci in the *chiffon/dbf4-like* mutant (J. Tower, unpublished results). This would be a novel role for the Dbf4-like protein, as in other organisms it acts at a step just prior to the initiation of replication and not at an early step in the formation of the pre-RC, such as ORC recruitment (for review, see Bell and Dutta, 2002). Dbf4-like is not the only replication factor that may have a novel function during gene amplification. Interestingly, DUP/Cdt1, known only as a replication initiation factor in all organisms for

which it has been studied, appears to travel with the replication forks during gene amplification, although the functional significance of this has not yet been determined (Claycomb et al., 2002). Both Dbf4-like and DUP/Cdt1 proteins have additional domains which are not present in homologs, suggesting that these domains may contribute to novel functions in gene amplification (J. Tower, unpublished results; Landis and Tower, 1999; Whittaker et al., 2000).

As mentioned above, chromatin immunoprecipitation experiments have shown that ORC2 binds directly to *ACE3* and to Ori β (Austin et al., 1999; Bosco et al., 2001). Additional ChIP and binding studies have demonstrated that the transcription factors E2F1/DP/Rbf (Bosco et al., 2001), and Myb as well as its binding partners Mip120, Mip130, Mip40, and Caf1 p55 also bind to *ACE3* (Beall et al., 2002). Additionally, the Myb complex also binds to Ori β (Beall et al., 2002). Myb is essential for viability and necessary for gene amplification, as somatic follicle cell clones carrying a null mutation of Myb resulted in an absence of BrdU incorporation during gene amplification, although ORC2 and DUP/Cdt1 localized properly (Beall et al., 2002). Knock-out mutants of the *mip130* gene are sterile and lead to BrdU incorporation throughout the nucleus at a time when amplification normally occurs (E. Beall and M. Botchan, unpublished results). These same mutants also have lowered levels of Myb protein, as do *Drosophila* S2 or Kc cells in which Caf1, p55, or Mip120 levels have been knocked down by RNAi, indicating that Myb must be in a complex to be stable. From these data, it seems that Mip130 is a part of a complex with the other Mips that is involved in the repression of replication, while Myb acts as a switch on this complex to stimulate replication. It is likely that Myb is always associated with the Mips, but that Myb becomes activated in some way,

perhaps by phosphorylation (Li and McDonnell, 2002) or other modification to change the activity of the complex. In this way, Myb would be specifically activated at amplification origins to allow the initiation of amplification at the appropriate developmental time.

It has been speculated, based on known roles for Myb and the E2F/Rb complexes in other organisms, that they may function at amplification origins to recruit histone acetyl-transferases (HATs) or histone de-acetylases (HDACs) that would modulate the accessibility of the chromatin at the origin (Beall et al., 2002; Bosco et al., 2001). Recent results demonstrate that histones H3 and H4 at and around *ACE3* are hyperacetylated during gene amplification, and that the lysine residues that are acetylated are associated with replication and not transcription (G. Bosco, unpublished results). Furthermore, the acetylation of H3 and H4 is not the result of histone deposition after replication, as the hyperacetylation is confined to the origins of DAFC-66D and not associated with the replication forks. The functional relevance of histone H3 and H4 hyperacetylation is yet to be demonstrated, but these results suggest that origin firing may be facilitated at amplification origins by a modification of the chromatin state.

The fact that *Drosophila* transgenes carrying origin and replication stimulatory sequences are highly subject to positional effects throughout the genome indicates that the amplified regions are susceptible to chromatin state. Chromatin state and nucleosomal positioning may play a role in gene amplification in *Sciara* and *Tetrahymena*. Histone acetyl-transferases have also been suggested to play a role in *Sciara* gene amplification, as has RNA polII (Clever and Ellgaard, 1970; Lunyak et al., 2002; Mok et al., 2001). Although transcription of the II-9-1 gene does not begin until

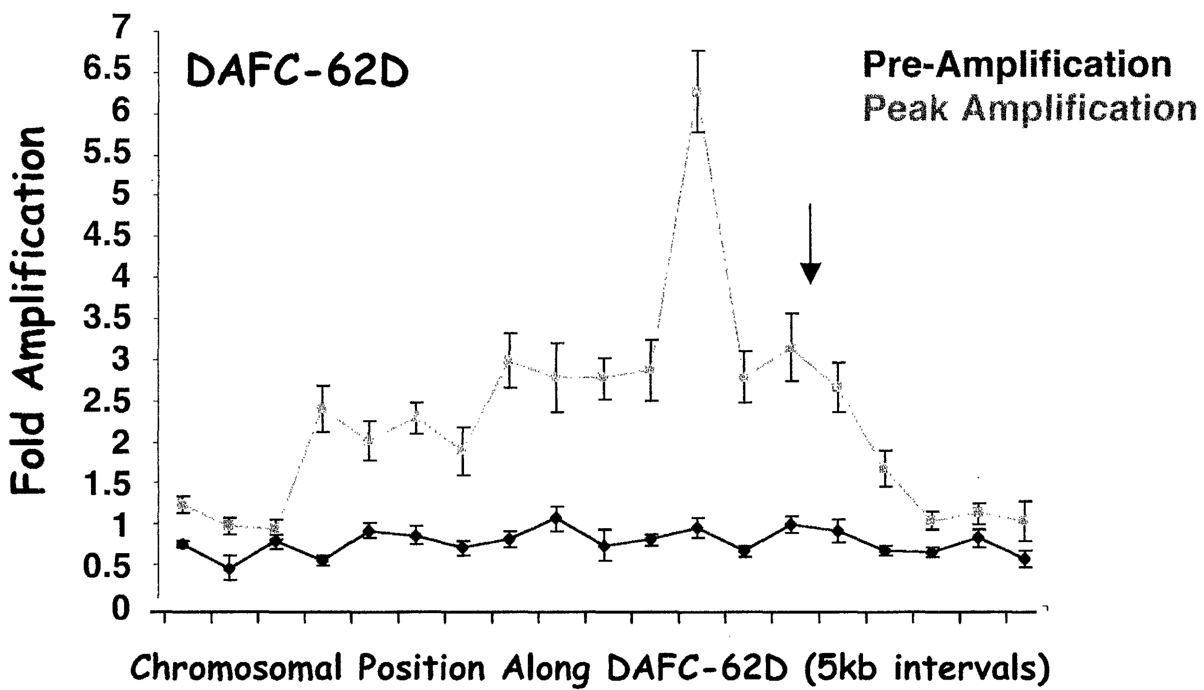
amplification is complete, the promoter of *II-9-1* is occupied by RNApolII during amplification stages, but not during mitotic or endocycles, and it is this presence that is thought to limit the right-hand boundary of the initiation zone during amplification. This effect could be due to RNApolII occupying positions used in previous replication cycles (mitotic, endocycle replication) by ORC, forcing ORC to different positions within the initiation zone, and causing a change in the preferred sites of initiation (Lunyak et al., 2002). Like yeast, the positioning of nucleosomes is precise within the 5'NTS of *Tetrahymena*, and this spacing is necessary for proper replication (Paliulis and Nicklas, 2000; Zhang et al., 1997).

It is clear that gene amplification is under complex control and is the product of cell cycle influences, the sequences of the amplified regions, and the factors actually performing the amplification, as well as those that control the state of the chromatin. However, we must also consider that there may be missing specificity factors or novel functions of known DNA replication proteins which would serve to modulate the use of these origins for gene amplification.

Replication Fork Movement through Amplified Intervals

During gene amplification, various peculiarities to normal replication fork movement have been observed. For instance, the replication forks copying amplified regions of the polytene chromosomes in *Drosophila* move at a particularly slow rate of 50-100bp/min. (Spradling and Leys, 1988) compared to the ~300bp/min. in endocycle polytene replication (Steinemann, 1981) or the 2.6kb/min. in synctial embryos or cell culture (Blumenthal et al., 1973). Perhaps this is due to the complexity of the onionskin

Figure 3. An example of the asymmetric fork movement in amplified regions: DAFC-62D. Real-time PCR analysis of peak amplification DNA and pre-amplification DNA at DAFC-62D in 5kb intervals shows that the fold amplification drops to 1 (non-amplified) within 25kb to the right of the gradient (arrow), as compared to 45kb on the left. This asymmetry may reflect differential replication fork movement. Adapted from Claycomb et al., 2004).



Asymmetry of Amplification Gradients
DAFC-62D
Pre-Amplification
Peak Amplification
Fold Amplification
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Chromosomal Position Along DAFC-62D (5kb intervals)

structure generated, and slow fork movement is necessary to properly traverse all copies of the chromosomes as they are spread out in 3D space. In *Drosophila*, immunofluorescence studies indicate an increase in the width of amplified regions during the elongation-only phase, suggesting that the additional DNA copies must be oriented with a certain minimum distance between each chromosome so as not to become entangled (Claycomb et al., 2002). Interestingly, it has been suggested that a possible role for the initiation factor DUP/Cdt1, at these slow moving replication forks is to adhere or exchange the MCM2-7 putative helicase, or perhaps even some other unknown factor (Claycomb et al., 2002).

The combined characteristics of slow fork movement and terminal differentiation in the salivary glands and follicle cells beg the question, are there specified fork termination sites at the ends of amplified regions or do the forks simply travel as far as they can in the developmental time allotted? In *Sciara*, there must be a means of terminating replication forks, as amplification ends prior to the onset of transcription (Gabrusewycz-Garica, 1971; Wu et al., 1993). In *Drosophila*, neither possibility can be ruled out, as transcription begins during the replication elongation phase for many of the amplified genes studied (Claycomb et al., 2004; Claycomb et al., 2002; Griffin-Shea et al., 1982; Parks et al., 1986). No termination sites have been identified yet and by the end of oogenesis, the forks have traveled the predicted distance based on the average fork rate.

Although there are not termination sites *per se* in *Tetrahymena*, there are certainly barriers to fork movement and replication fork pausing sites, as observed by 2-D gel electrophoresis. It seems that at the center of the rDNA palindrome in the 5' NTS, a

barrier exists during amplification only, that does not allow forks to pass into the opposite side of the palindrome (Zhang et al., 1997). This barrier may be necessary to prevent replication forks from traveling in the opposite direction of transcription, a situation that would be deleterious for the organism. Additionally, directional pause sites in the Type I elements of Domains 1 and 2 serve to modulate fork rate movement through the 5' NTS, but the significance of this is not entirely clear (MacAlpine et al., 1997; Zhang et al., 1997). Interestingly, *Rhynchosciara* and *Drosophila* may possess a similar modulator of replication fork movement, as the gradients of copy number for several amplified regions are not symmetrical (Claycomb et al., 2004; Claycomb et al., 2002; Spradling, 1981; Yokosawa et al., 1999).

The Future of Developmental Gene Amplification

Although previous studies have searched for additional examples of developmentally important gene amplifications, a systematic approach was unavailable until the advent of genomic technologies. The CGH array has been widely and reliably used to find copy-number changes in cancer tissues, and the isolation of two amplified regions in *Drosophila* follicle cells demonstrates that this is a viable approach for isolating new developmental amplicons in a variety of tissues and organisms. At DAFC-30B and 62D, the peak enrichment in copy number is 4 or 6.5-fold, respectively, and although this level of copy number increase seems slight, it is detectable by the array approach. Even more importantly, this seemingly low level of amplification can be biologically significant. These points and the fact that genes encoding enzymes are amplified in and expressed from DAFC-30B and 62D indicate that gene amplification

may be much more pervasive than we once thought. With the genomic tools in hand and a plethora of tissues to examine, the field is poised to expand our knowledge of developmental amplicons, the functions of amplified genes, and the regulation of gene amplification throughout various developmental contexts in the lifecycle of a complex organism.

Summary of Thesis

The work in this thesis attempts to strengthen our understanding of the process of gene amplification as a model for DNA replication and in terms of the developmental importance of this strategy. To refine our understanding of how origins fire at the two chorion gene amplicons, I performed confocal microscopy and quantitative real-time PCR studies, as well as deconvolution microscopy measurements with the help of James Evans, a postdoctoral fellow in the Matsudaira lab. The results of these experiments are summarized in Chapter 2 and Appendix 5. These studies revealed that origin firing occurs during stages 10B and 11 of egg chamber development, when replication factors ORC2, DUP/Cdt1, PCNA, and MCM2-7, as well as BrdU, localize to amplification foci. In stages 12 and 13 of egg chamber development, the foci of staining resolve into double bar structures for BrdU, MCM2-7, PCNA, and even DUP/Cdt1, and ORC2 dissociates from the origins, as only the existing replication forks progress outward and no further origin firings occur. These findings demonstrate that amplification of the chorion genes provides a superb model system for studying the difficult problem of replication elongation *in vivo*, and suggest that the replication initiation factor, DUP/Cdt1, may also play a role in replication elongation.

BrdU incorporation studies (Calvi et al., 1998) had previously shown the existence of four foci in follicle cells during gene amplification stages, leading us to believe that there were at least two additional amplicons. This observation was coupled with the motivation to develop a systematic microarray assay to screen for differentially replicated regions of the genome, and 16C amplifying follicle cell genomic DNA was used in the first array hybridization experiments done by Matt Benasutti, a former Orr-Weaver lab technician. These results are described in Chapter 3, and Appendices 1 and 4. When compared to the 2N embryonic control DNA, a number of genes in four distinct clusters throughout the genome were enriched in copy number. Two of the gene clusters fell within the known amplicons, and two were new clusters of genes at cytological positions 30B and 62D. By Fluorescent *in situ* Hybridization (FISH) and BrdU colabeling and quantitative real-time PCR, I verified that these regions were amplified in the follicle cells during late oogenesis. I determined that genes within each amplicon were robustly expressed, and that amplification was necessary for the expression of these genes. One amplified gene, *yellow-g*, was shown to be necessary for proper eggshell and vitelline membrane formation by mutant analysis. These results indicate that there are additional examples of the use of gene amplification throughout development, and that the microarray technology is capable of isolating previously undiscovered amplicons. Furthermore, the genes found within these amplicons indicate that amplification of enzymes may be an important means of regulating developmental processes.

Although much work has been done to elucidate the factors involved with and regulating gene amplification, the role of the replication initiation factor CDC6 has never been studied in this system, or in any aspect of DNA replication in *Drosophila*. No

mutants are available in *cdc6*, so we turned to overexpression studies using the *UAS*-gal4 system to assess the role of CDC6 in gene amplification. Through the efforts of Rick Austin, a former Bell lab postdoctoral fellow, and Giovanni Bosco, a former Orr-Weaver lab postdoctoral fellow, antibodies to DmCDC6 were generated in guinea pigs, and I performed the characterization of this reagent. These CDC6 results are summarized in Appendix 3.

In trying to better understand the role of DUP/Cdt1 in gene amplification, I performed Co-IP experiments with DUP/Cdt1 and several other replication factors, and I tried to ChIP DUP/Cdt1 at DAFC-66D. Although these experiments never gave conclusive results, they are summarized in Appendix 2. The nature of some of the *dup* mutants isolated in this lab and their complementation to other *dup* alleles was worrisome, given that there is a serine/threonine protein kinase encoded in an overlapping reading frame on the opposite DNA strand from *dup*. We were concerned that some of the *dup* alleles may be affecting the kinase, as well as *dup*, so I generated a *pUASp-dup* transgene to perform rescue experiments on each of the alleles. Unfortunately, the transgene was unable to rescue any of the *dup* alleles, even though both mRNA and protein are produced from it. The lack of rescue may be the result of complex post-translational regulation of DUP/Cdt1 protein. The results of the rescue are summarized in Appendix 2.

References:

Allis, C.D., M. Colavito-Shepanski, and M.A. Gorovsky. 1987. Scheduled and unscheduled DNA synthesis during development in conjugating *Tetrahymena. Dev Biol.* 124:469-80.

Aparicio, O.M., A.M. Stout, and S.P. Bell. 1999. Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proc. Natl. Acad. Sci. USA.* 96:9130-5.

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

Asano, M., and R.P. Wharton. 1999. E2F mediates developmental and cell cycle regulation of ORC1 in *Drosophila. EMBO J.* 18:2435-2448.

Austin, R.J., T.L. Orr-Weaver, and S.P. Bell. 1999. *Drosophila* ORC specifically binds to *ACE3*, an origin of DNA replication control element. *Genes and Dev.* 13:2639-2649.

Beall, E.L., J.R. Manak, S. Zhou, M. Bell, J.S. Lipsick, and M.R. Botchan. 2002. Role for a *Drosophila* Myb-containing protein complex in site-specific DNA replication. *Nature.* 420:833-7.

Bell, S.P., and A. Dutta. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71:333-74.

Bell, S.P., and B. Stillman. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature.* 357:128-134.

Bielinsky, A.K., H. Blitzblau, E.L. Beall, M. Ezrokhi, H.S. Smith, M.R. Botchan, and S.A. Gerbi. 2001. Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.* 11:1427-1431.

Blomberg, P., C. Randolph, C.H. Yao, and M.C. Yao. 1997. Regulatory sequences for the amplification and replication of the ribosomal DNA minichromosome in *Tetrahymena thermophila. Mol Cell Biol.* 17:7237-47.

Blow, J.J., and R.A. Laskey. 1988. A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature.* 332:546-548.

Blumenthal, A., H. Kriegstein, and D. Hogness. 1973. The units of DNA replication in *Drosophila melanogaster* chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* 38:205-223.

Bosco, G., W. Du, and T.L. Orr-Weaver. 2001. DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nature Cell Biol.* 3:289-295.

Brown, D.D., and I.B. Dawid. 1968. Specific gene amplification in oocytes. *Science.* 160:272-280.

Calvi, B.R., M.A. Lilly, and A.C. Spradling. 1998. Cell cycle control of chorion gene amplification. *Genes and Dev.* 12:734-44.

Carminati, J.L., C.J. Johnston, and T.L. Orr-Weaver. 1992. The *Drosophila ACE3* chorion element autonomously induces amplification. *Mol. Cell. Biol.* 12:2444-2453.

Cayirlioglu, P., P.C. Bonnette, M.R. Dickson, and R.J. Duronio. 2001. *Drosophila* E2f2 promotes the conversion from genomic DNA replication to gene amplification in ovarian follicle cells. *Development.* 128:5085-5098.

Cayirlioglu, P., W.O. Ward, S.C. Silver Key, and R.J. Duronio. 2003. Transcriptional repressor functions of *Drosophila* E2F1 and E2F2 cooperate to inhibit genomic DNA synthesis in ovarian follicle cells. *Mol Cell Biol.* 23:2123-34.

Cech, T.R., and S.L. Brehm. 1981. Replication of the extrachromosomal ribosomal RNA genes of *Tetrahymena* thermophilia. *Nucleic Acids Res.* 9:3531-43.

Chesnokov, I., M. Gossen, D. Remus, and M. Botchan. 1999. Assembly of functionally active *Drosophila* origin recognition complex from recombinant proteins. *Genes and Dev.* 13:1289-96.

Claycomb, J.M., M. Benasutti, G. Bosco, D.D. Fenger, and T.L. Orr-Weaver. 2004. Gene amplification as a developmental strategy: isolation of two developmental amplicons in *Drosophila. Dev Cell.* 6:145-55.

Claycomb, J.M., D.M. MacAlpine, J.G. Evans, S.P. Bell, and T.L. Orr-Weaver. 2002. Visualization of replication initiation and elongation in *Drosophila. J. Cell Biol.* 159:225-236.

Clever, U., and E.G. Ellgaard. 1970. Puffing and histone acetylation in polytene chromosomes. *Science.* 169:373-4.

Coelho, P.S., J.C. De Almeida, F. Toledo, G. Buttin, and M.L. Paco-Larson. 1993. DNA puff C4 of Bradysia hygida (Diptera: Sciaridae) contains genes unequally amplified and differentially expressed during development. *Chromosome Res.* 1:121-6.

de Cicco, D., and A. Spradling. 1984. Localization of a cis-acting element responsible for the developmentally regulated amplification of *Drosophila* chorion genes. *Cell.* 38:45-54.

Delidakis, C., and F.C. Kafatos. 1987. Amplification of a chorion gene cluster in *Drosophila* is subject to multiple cis-regulatory elements and to long range position effects. *J. Mol. Biol.* 197:11-26.

Delidakis, C., and F.C. Kafatos. 1989. Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila. EMBO J.* 8:891-901.

Deng, W.M., C. Althauser, and H. Ruohola-Baker. 2001. Notch-Delta signaling induces a transition from mitotic cell cycle to endocycle in *Drosophila* follicle cells. *Development.* 128:4737-46.

Edgar, B.A., and T.L. Orr-Weaver. 2001. Endoreplication cell cycles: more for less. *Cell.* 105:297-306.

Fontes, A.M., J.C. de-Almeida, J.E. Edstrom, and M.L. Paco-Larson. 1992. Cloning of a B10 DNA puff sequence developmentally amplified and expressed in the salivary gland of Bradysia hygida. *Braz J Med Biol Res.* 25:777-80.

Gabrusewycz-Garica, N. 1971. Studies in polytene chromosomes of sciarids. I. The salivary chromosomes of *Sciara* (Lycoriella) pauciseta (II), Felt. *Chromosoma.* 33:421-35.

Gall, J.G. 1968. Differential synthesis of the genes for ribosomal RNA during amphibian oogenesis. *Proc. Natl. Acad. Sci. USA.* 60:553-560.

Gallagher, R.C., and E.H. Blackburn. 1998. A promoter region mutation affecting replication of the *Tetrahymena* ribosomal DNA minichromosome. *Mol Cell Biol.* 18:3021-33.

Gilbert, W., and D. Dressler. 1968. DNA replication: the rolling circle model. *Cold Spring Harb Symp Quant Biol.* 33:473-84.

Glover, D.M., A. Zaha, A.J. Stocker, R.V. Santelli, M.T. Pueyo, S.M. De Toledo, and F.J.S. Lara. 1982. Gene amplification in *Rhynchosciara* salivary gland chromosomes. *Proc. Natl. Acad. Sci. USA.* 79:2947-2951.

Griffin-Shea, R., G. Thireos, and F.C. Kafatos. 1982. Organization of a cluster of four chorion genes in *Drosophila* and its relationship to developmental expression and amplification. *Dev. Biol.* 91:325-336.

Heck, M., and A. Spradling. 1990. Multiple replication origins are used during *Drosophila* chorion gene amplification. *J. Cell Biol.* 110:903-914.

Henderson, D.S., U.K. Wiegand, D.G. Norman, and D.M. Glover. 2000. Mutual correction of faulty PCNA subunits in temperature-sensitive lethal *mus209* mutants of *Drosophila melanogaster*. *Genetics.* 154:1721-1733.

Hourcade, D., D. Dressler, and J. Wolfson. 1973. The amplification of ribosomal RNA genes involves a rolling circle intermediate. *Proc Natl Acad Sci U S A.* 70:2926-30.

Ishimi, Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J. Biol. Chem.* 272:24508-24513.

Kapler, G.M., and E.H. Blackburn. 1994. A weak germ-line excision mutation blocks developmentally controlled amplification of the rDNA minichromosome of *Tetrahymena thermophila. Genes Dev.* 8:84-95.

Kapler, G.M., D.L. Dobbs, and E.H. Blackburn. 1996 DNA replication in *Tetrahymena*, p.915-932. *In* M.L. DePamphilis (ed.), DNA replication in eukaryotic cells Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Kapler, G.M., E. Orias, and E.H. Blackburn. 1994. *Tetrahymena thermophila* mutants defective in the developmentally programmed maturation and maintenance of the rDNA minichromosome. *Genetics.* 137:455-66.

Labib, K., S.E. Kearsey, and J.F. Diffley. 2001. MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G1-phase and are required to establish, but not maintain, the S-phase checkpoint. *Mol Biol Cell.* 12:3658-67.

Laicine, E.M., M.A. Alves, J.C. de Almeida, E. Rizzo, W.C. Albernaz, and H. Sauaia. 1984. Development of DNA puffs and patterns of polypeptide synthesis in the salivary glands of Bradysia hygida. *Chromosoma.* 89:280-4.

Landis, G., R. Kelley, A.C. Spradling, and J. Tower. 1997. The *k43* gene, required for chorion gene amplification and diploid cell chromosome replication, encodes the *Drosophila* homolog of yeast origin recognition complex subunit 2. *Proc. Natl. Acad. Sci. USA.* 94:3888-3892.

Landis, G., and J. Tower. 1999. The *Drosophila chiffon* gene is required for chorion gene amplification, and is related to the yeast dbf4 regulator of DNA replication and cell cycle. *Development.* 126:4281-4293.

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

Li, X., and D.P. McDonnell. 2002. The transcription factor B-Myb is maintained in an inhibited state in target cells through its interaction with the nuclear corepressors N-CoR and SMRT. *Mol Cell Biol.* 22:3663-73.

Liang, C., and S.A. Gerbi. 1994. Analysis of an origin of DNA amplification in *Sciara coprophila* by a novel three-dimensional gel method. *Mol Cell Biol.* 14:1520-9.

Liang, C., J.D. Spitzer, H.S. Smith, and S.A. Gerbi. 1993. Replication initiates at a confined region during DNA amplification in *Sciara* DNA puff II/9A. *Genes and Dev.* 7:1072-1084.

Lilly, M.A., and A.C. Spradling. 1996. The *Drosophila* endocycle is controlled by Cyclin E and lacks a checkpoint ensuring S-phase completion. *Genes and Dev.* 10:2514-2526.

Loebel, D., H. Huikeshoven, and S. Cotterill. 2000. Localisation of the DmCdc45 DNA replication factor in the mitotic cycle and during chorion gene amplification. *Nucleic Acids Res.* 28:3897-903.

Lopez-Schier, H., and D. St Johnston. 2001. Delta signaling from the germ line controls the proliferation and differentiation of the somatic follicle cells during *Drosophila* oogenesis. *Genes Dev.* 15:1393-405.

Lu, L., Z. Hongjun, and J. Tower. 2001. Functionally distinct, sequence specific replicator and origin elements are required for *Drosophila* chorion gene amplification. *Genes and Dev.* 15:134-146.

Lunyak, V.V., M. Ezrokhi, H.S. Smith, and S.A. Gerbi. 2002. Developmental changes in the *Sciara* II/9A initiation zone for DNA replication. *Mol Cell Biol.* 22:8426-37.

MacAlpine, D.M., Z. Zhang, and G.M. Kapler. 1997. Type I elements mediate replication fork pausing at conserved upstream sites in the *Tetrahymena thermophila* ribosomal DNA minichromosome. *Mol Cell Biol.* 17:4517-25.

Merchant, A.M., Y. Kawasaki, Y. Chen, M. Lei, and B.K. Tye. 1997. A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 17:3261-71.

Mohammad, M., S. Saha, and G.M. Kapler. 2000. Three different proteins recognize a multifunctional determinant that controls replication initiation, fork arrest and transcription in *Tetrahymena. Nucleic Acids Res.* 28:843-51.

Mohammad, M., R.D. York, J. Hommel, and G.M. Kapler. 2003. Characterization of a novel origin recognition complex-like complex: implications for DNA recognition, cell cycle control, and locus-specific gene amplification. *Mol Cell Biol.* 23:5005-17.

Mok, E.H., H.S. Smith, S.M. DiBartolomeis, A.W. Kerrebrock, L.J. Rothschild, T.S. Lange, and S.A. Gerbi. 2001. Maintenance of the DNA puff expanded state is independent of active replication and transcription. *Chromosoma.* 110:186-96.

Monesi, N., M.A. Fernandez, A.M. Fontes, L.R. Basso, Jr., Y. Nakanishi, B. Baron, G. Buttin, and M.L. Paco-Larson. 1995. Molecular characterization of an 18 kb segment of DNA puff C4 of Bradysia hygida (Diptera, sciaridae). *Chromosoma.* 103:715-24.

Orr, W., K. Komitopoulou, and F.C. Kafatos. 1984. Mutants supressing in *trans* chorion gene amplification in *Drosophila. Proc. Natl. Acad. Sci. USA.* 81:3773-3777.

Orr-Weaver, T.L., C.G. Johnston, and A.C. Spradling. 1989. The role of *ACE3* in *Drosophila* chorion gene amplification. *EMBO J.* 8:4153-4162.

Osheim, Y.N., O.L. Miller, and A.L. Beyer. 1988. Visualization of *Drosophila melanogaster* chorion genes undergoing amplification. *Mol. Cell. Biol.* 8:2811-2821.

Paliulis, L.V., and R.B. Nicklas. 2000. The reduction of chromosome number in meiosis is determined by properties built into the chromosomes. *J. Cell Biol.* 150:1223-1231.

Parks, S., B. Wakimoto, and A. Spradling. 1986. Replication and expression of an X-linked cluster of *Drosophila* chorion genes. *Dev. Biol.* 117:294-305.

Penalva, L.O., J. Yokosawa, A.J. Stocker, M.A. Soares, M. Graessmann, T.C. Orlando, C.E. Winter, L.M. Botella, A. Graessmann, and F.J. Lara. 1997. Molecular characterization of the C-3 DNA puff gene of Rhynchosciara americana. *Gene.* 193:163-72.

Reischmann, K.P., Z. Zhang, and G.M. Kapler. 1999. Long range cooperative interactions regulate the initiation of replication in the *Tetrahymena thermophila* rDNA minichromosome. *Nucleic Acids Res.* 27:3079-89.

Remus, D., E.L. Beall, and M.R. Botchan. 2004. DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC-DNA binding. *Embo J.* 23:897-907.

Rochaix, J.D., A. Bird, and A. Barkken. 1974. Ribosomal RNA gene amplification by rolling circles. *J Mol Biol.* 87:473-87.

Royzman, I., R.J. Austin, G. Bosco, S.P. Bell, and T.L. Orr-Weaver. 1999. ORC localization in *Drosophila* follicle cells and the effects of mutations in *dE2F* and *dDP*. *Genes and Dev.* 13:827-840.

Saha, S., A. Nicholson, and G.M. Kapler. 2001. Cloning and biochemical analysis of the *Tetrahymena* origin binding protein TIF1: competitive DNA binding in vitro and in vivo to critical rDNA replication determinants. *J Biol Chem.* 276:45417-26.

Santelli, R.V., G.M. Machado-Santelli, M.T. Pueyo, L.D. Navarro-Cattapan, and F.J. Lara. 1991. Replication and transcription in the course of DNA amplification of the C3 and C8 DNA puffs of Rhynchosciara americana. *Mech Dev.* 36:59-66.

Schwed, G., N. May, Y. Pechersky, and B.R. Calvi. 2002. *Drosophila* minichromosome maintenance 6 is required for chorion gene amplification and genomic replication. *Mol. Biol. Cell.* 13:607-620.

Spradling, A.C. 1981. The organization and amplification of two clusters of *Drosophila* chorion genes. *Cell.* 27:193-201.

Spradling, A.C. 1993. Developmental genetics of oogenesis. *In* The development of *Drosophila melanogaster.* Vol. 1. M. Bate and A. Martinez Arias, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1-70.

Spradling, A.C., D.V. de Cicco, B.T. Wakimoto, J.F. Levine, L.J. Kalfayan, and L. Cooley. 1987. Amplification of the X-linked *Drosophila* chorion gene cluster requires a region upstream from the *s38* chorion gene. *EMBO J.* 6:1045-1053.

Spradling, A.C., M.E. Digan, A.P. Mahowald, M. Scott, and E.A. Craig. 1980. Two clusters of genes for major chorion proteins of *Drosophila melanogaster. Cell.* 19:905-14.

Spradling, A.C., and E. Leys. 1988. Slow replication fork movement during *Drosophila* chorion gene amplification. *In* Cancer Cells. Vol. 6. T. Kelly and B. Stillman, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor. 305-309.

Stark, G.R., M. Debatisse, E. Giulotto, and G.M. Wahl. 1989. Recent progress in understanding mechanisms of mammalian DNA amplification. *Cell.* 57:901-8.

Steinemann, M. 1981. Chromosomal replication in *Drosophila* virilis. III. Organization of active origins in the highly polytene salivary gland cells. *Chromosoma.* 82:289-307.

Swimmer, C., M.G. Fenerjian, J.C. Martinez-Cruzado, and F.C. Kafatos. 1990. Evolution of the autosomal chorion cluster in *Drosophila* III. Comparison of the *s*18 gene in evolutionarily distant species and interspecific control of chorion gene amplification. *J. Mol. Biol.*:In Press.

Tercero, J.A., K. Labib, and J.F. Diffley. 2000. DNA synthesis at individual replication forks requires the esential initiation factor Cdc45p. *EMBO J.* 19:2082-93.

Underwood, E.M., A.S. Briot, K.Z. Doll, R.L. Ludwiczak, D.C. Otteson, J. Tower, K.B. Vessey, and K. Yu. 1990. Genetics of 51D-52A, a region containing several maternal-effect genes and two maternal-specific transcripts in *Drosophila. Genetics.* 126:639-650.

Urnov, F.D., C. Liang, H.G. Blitzblau, H.S. Smith, and S.A. Gerbi. 2002. A DNase I hypersensitive site flanks an origin of DNA replication and amplification in *Sciara. Chromosoma.* 111:291-303.

Vashee, S., C. Cvetic, W. Lu, P. Simancek, T.J. Kelly, and J.C. Walter. 2003. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* 17:1894-908.

Ward, J.G., P. Blomberg, N. Hoffman, and M.C. Yao. 1997. The intranuclear organization of normal, hemizygous and excision-deficient rRNA genes during developmental amplification in *Tetrahymena thermophila. Chromosoma.* 106:233-42.

Whittaker, A.J., I. Royzman, and T.L. Orr-Weaver. 2000. *Drosophila* double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes Dev.* 14:1765-1776.

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

Wu, N., C. Liang, S.M. DiBartolomeis, H.S. Smith, and S.A. Gerbi. 1993. Developmental progression of DNA puffs in *Sciara coprophila*: amplification and transcription. *Dev. Biol.* 160:73-84.

Yao, M.C., A.R. Kimmel, and M.A. Gorovsky. 1974. A small number of cistrons for ribosomal RNA in the germinal nucleus of a eukaryote, *Tetrahymena* pyriformis. *Proc Natl Acad Sci U S A.* 71:3082-6.

Yokosawa, J., M.A. Soares, P.A. Dijkwel, A.J. Stocker, J.L. Hamlin, and F.J. Lara. 1999. DNA replication during amplification of the C3 puff of Rhynchosciara americana initiates at multiple sites in a 6 kb region. *Chromosoma.* 108:291-301.

Zhang, Z., D.M. Macalpine, and G.M. Kapler. 1997. Developmental regulation of DNA replication: replication fork barriers and programmed gene amplification in *Tetrahymena thermophila. Mol Cell Biol.* 17:6147-56.

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

Chapter Two

Visualization of Replication Initiation and Elongation in *Drosophila*

Julie M. Claycomb[1,2], David M. MacAlpine[2*], James G. Evans[1*], Stephen P. Bell[2], and Terry L. Orr-Weaver[1,2]

[1]Whitehead Institute for Biomedical Research and [2]Dept. of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

*These authors contributed equally to this work.

This chapter was published in the Journal of Cell Biology 159(2): 225-236 (2002).

Julie M. Claycomb performed all of the immunolocalization experiments and confocal microscopy for ORC2, BrdU, PCNA, DUP/Cdt1, MCM2-7, and *ACE3* FISH, did all of the quantitative real-time PCR experiments for the third chromosome chorion locus, and assisted James Evans in collecting the confocal images for deconvolution microscopy.

ABSTRACT

Chorion gene amplification in the ovaries of *Drosophila melanogaster* is a powerful system for the study of metazoan DNA replication *in vivo*. Using a combination of high resolution confocal and deconvolution microscopy and quantitative realtime PCR, we found that initiation and elongation occur during separate developmental stages, thus permitting analysis of these two phases of replication *in vivo*. BrdU, ORC, and the elongation factors MCM2-7 and PCNA were precisely localized and the DNA copy number along the third chromosome chorion amplicon was quantified during multiple developmental stages. These studies revealed that initiation takes place during stages 10B and 11 of egg chamber development, whereas only elongation of existing replication forks occurs during egg chamber stages 12 and 13. The ability to distinguish initiation from elongation makes this an outstanding model to decipher the roles of various replication factors during metazoan DNA replication. We utilized this system to demonstrate that the pre-RC component, DUP/Cdt1 not only is necessary for proper MCM2-7 localization but unexpectedly is present during elongation.

INTRODUCTION

Studies in the yeast *Saccharomyces cerevisiae* have provided insight into the mechanism and control of eukaryotic DNA replication. Yeast possess specific, well-defined origins of DNA replication onto which complexes of replication factors assemble. Generally, yeast origins are 200bp or less and consist of an 11bp A-T rich ARS Consensus Sequence (ACS), as well as the B1 and B2 elements. The pre-replication complex (pre-RC) assembles onto these regions during the G1 phase of the cell cycle, resulting in origins that are competent to initiate DNA replication and serving as a molecular beacon to recruit the replication fork machinery (Bell and Dutta, 2002; Bielinsky et al., 2001 for reviews) .

A combination of approaches in *S. cerevisiae* has identified components of the pre-RC and the replication fork machinery (Dutta and Bell, 1997; Bell and Dutta, 2002 for reviews). The six-member Origin Recognition Complex (ORC) was identified as a pre-RC component by its ability to bind to yeast replication origins (Bell and Stillman, 1992). ORC binds to the ACS and B1 elements, then recruits the pre-RC factors, Cdc6/Cdc18 and DUP/Cdt1. DUP/Cdt1 and Cdc6/Cdc18 in turn load the hexameric MCM2-7 complex onto pre-RCs. MCM2-7 are necessary for initiation, but are also required for elongation and travel with replication forks (Aparicio et al., 1997; Labib et al., 2000). Furthermore, MCM 4, 6, and 7 have helicase activity *in vitro*, suggesting that they function as the replicative helicase (Ishimi, 1997; Labib et al., 2000).

Once MCM2-7 are loaded, additional replication factors are recruited to origins and replication initiates. Cdc45 and Mcm10 are two other factors necessary for both initiation and elongation that travel with replication forks (Merchant et al., 1997;

Aparicio et al., 1999; Tercero et al., 2000; Wohlschlegel et al., 2002). CDK and Cdc7-Dbf4 kinase activity are required for initiation, with MCM2-7 and Cdc45 as potential targets (Lei et al., 1997; Zou and Stillman, 2000). Replication fork components must also be recruited for origin firing. These include the single-stranded DNA binding protein RPA, Polα primase, the clamp loader Replication Factor C (RFC), the sliding clamp Proliferating Cell Nuclear Antigen (PCNA), DPB11, and the replicative polymerases Polδ/ε (Waga and Stillman, 1998; Bell and Dutta, 2002 for reviews).

Although the pre-RC and replication fork components are structurally conserved in metazoans (Donaldson and Blow, 1999), analysis of replication initiation and elongation is limited by the lack of model replicons. Using cells and extracts from humans, Xenopus, or *Drosophila*, pre-RCs can assemble on model templates and DNA replication can initiate *in vitro*, giving results consistent with the yeast paradigm of pre-RC and replication fork composition and activity (Chesnokov et al., 1999; Mendez and Stillman, 2000; Blow, 2001). However, obstacles such as multiple potential initiation sites and complex *cis*-regulatory sequences have hindered the progress of *in vivo* replication initiation studies (DePamphilis, 1999; Bielinsky and Gerbi, 2001 for reviews). In addition, a lack of genetic assays has made it difficult to study the precise localization and properties of the *trans*-factors necessary for replication. Thus the available models in vertebrates have yielded information about either *cis*-elements or *trans*-factors necessary for replication, but a single system has not provided information about both.

In contrast, amplification in the Dipteran flies *Drosophila melanogaster* and *Sciara coprophila* has provided the framework to study DNA replication in which the *cis*-regulatory sequences are well defined and *trans*-acting replication factors can be

examined (Calvi and Spradling, 1999; Bielinsky et al., 2001). In *Sciara* the replication start site within an amplified salivary puff origin, *ori II/9A*, is understood at the single nucleotide level and displays similarities to the yeast ARS. Furthermore, *Drosophila* ORC has been shown to bind to an 80bp region adjacent to this replication start site (Bielinsky et al., 2001). In *Drosophila*, amplification of the chorion gene clusters provides another powerful system for the study of metazoan DNA replication. The ovarian follicle cells, somatic cells that surround the developing oocyte, synthesize and secrete the chorion, or eggshell. In response to developmental signals at stages 9-10 of egg chamber development, the follicle cells end genomic DNA replication and begin to amplify several clusters of genes throughout the genome, including two clusters of chorion genes (Calvi et al., 1998).

Amplification of the chorion clusters occurs via a bi-directional replication mechanism, in which initiation occurs repeatedly from defined origins and forks progress outward to approximately 50kb on either side of the origins (Spradling and Mahowald, 1981; Spradling, 1981; Osheim et al., 1988; Delidakis and Kafatos, 1989; Heck and Spradling, 1990). By stage 13 of egg chamber development a gradient of copy number results, with the origins and chorion genes located at the central, maximally amplified region. Quantitative Southern blots detect a maximum copy level of 16-20 fold for the *X* chromosome chorion cluster and 60-100 fold for the third chromosome chorion cluster (Spradling, 1981; Delidakis and Kafatos, 1989;). P-element mediated transformation of DNA fragments from the third chromosome cluster defined the *cis*-regulatory element, Amplification Control Element on 3 (*ACE3*), which is required for high levels of amplification and sufficient for low levels of amplification (de Cicco and Spradling,

1984; Carminati et al., 1992). Two dimensional gel analysis demonstrated that repeated firings occur from a preferred origin, *oriβ*, about 1.5kb downstream of *ACE3* (Delidakis and Kafatos, 1989; Heck and Spradling, 1990). Further transformation experiments showed that *ACE3* interacts with *oriβ* (Lu et al., 2001).

Genetic studies took advantage of female-sterile mutations to demonstrate an essential role for known replication factors in chorion amplification. Females mutant for *orc2, dbf4-like, mcm6,* and *dup/cdt1* lay eggs with thin or otherwise abnormal eggshells due to defects in chorion amplification (Underwood et al., 1990; Landis et al., 1997; Landis and Tower, 1999; Whittaker et al., 2000; Schwed et al., 2002).

In addition to genetic approaches, the process of chorion amplification can be visualized directly. BrdU incorporation at amplicons can be detected throughout the amplification process, from stages 10B to 13 (Calvi et al., 1998). The replication proteins ORC2, ORC1, ORC5, DUP/Cdt1, and CDC45 localize specifically to amplicons during chorion amplification in follicle cells (Asano and Wharton, 1999; Austin et al., 1999; Royzman et al., 1999; Loebel et al., 2000; Whittaker et al., 2000). In this study, we use a cell biology approach, coupled with quantitative realtime PCR, to decipher the dynamics of DNA replication at the chorion loci in a developmental context. We find that all initiation at chorion origins occurs during one part of amplification, while in subsequent stages, only the existing replication forks elongate. We also observe that the localization pattern of several replication factors during amplification correlates with the roles of these proteins in initiation or elongation.

RESULTS

Localization patterns of ORC2 and BrdU throughout chorion amplification

We performed high-resolution deconvolution microscopy to analyze the pattern of ORC localization with respect to BrdU incorporation at the third chromosome chorion locus throughout amplification. Previous observations showed that ORC2 localizes to amplified regions for only a portion of amplification, from egg chamber stages 10A to 11 (Royzman et al., 1999). In contrast, BrdU incorporation begins in stage 10B and persists until stage 13 of egg chamber development (Calvi et al., 1998; Royzman et al., 1999; Calvi and Spradling, 2001). These differences in localization patterns suggest that DNA replication continues in the absence of ORC2 at chorion loci; that is elongation exclusively may occur during stages 12 and 13.

When BrdU incorporation became detectable early in stage 10B follicle cell nuclei, ORC2 localized to the *X* and third chromosome chorion clusters and was coincident with BrdU (unpublished results). As stage 10B continued, ORC2 no longer localized to the *X* chromosome cluster, but persisted at the third chromosome cluster, coincident with BrdU (Fig. 1A-B, Supplemental Movie 1). At this time ORC2 was present at origin sequences, as the ORC2 signal colocalized with that of a FISH probe spanning *ACE3* and *oriβ* on the third chromosome (Fig. 1C-E). Additionally, chromatin immunoprecipitation experiments have shown that *in vivo*, ORC is bound in the vicinity of *ACE3* and *oriβ* in amplifying stage 10 follicle cells (Austin et al., 1999). Furthermore, our observations are consistent with previous results obtained by Calvi, localizing the same FISH probe relative to BrdU incorporation (Calvi et al., 1998; Calvi and Spradling,

2001). It should also be noted that even though the follicle cells are polyploid (16C), the fact that there is a single BrdU spot (or set of double bars, see below) for each amplicon demonstrates that all the chromosome copies must be tightly aligned as polytene chromosomes (Calvi and Spradling, 2001). These data demonstrate that ORC2 is at chorion origins when they fire and begin to incorporate BrdU.

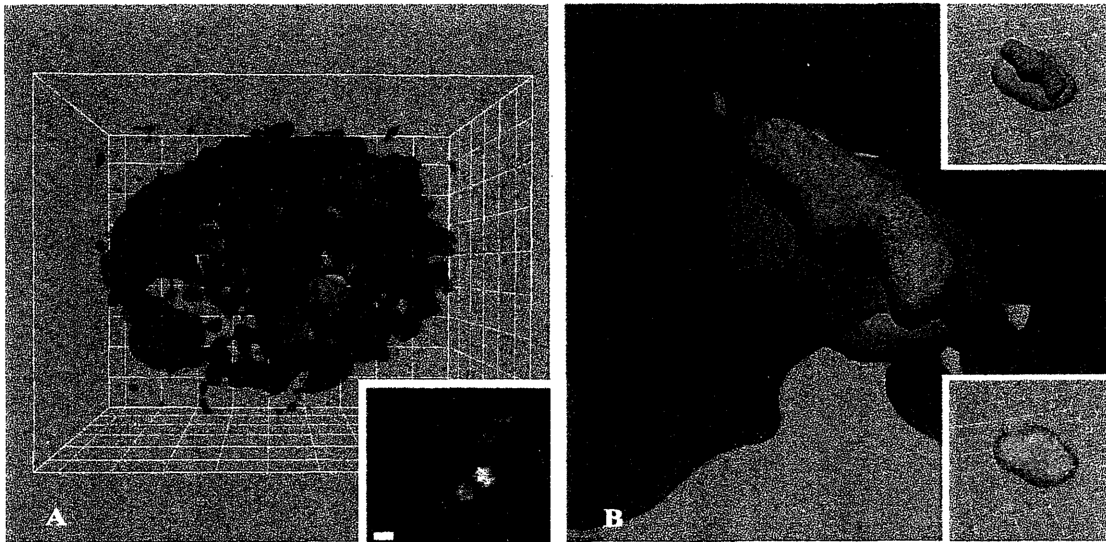
As chorion amplification proceeded, deconvolution microscopy revealed that the pattern of BrdU incorporation diverged from that of ORC2. In stage 11 egg chambers, the BrdU staining pattern resolved into a coffee-bean like structure, with bands of BrdU incorporation flanking ORC2 present at the origins (Fig. 1 F-G). Furthermore, and consistent with the results of Royzman, during stage 11 ORC2 dissipated from the origins and a higher level of diffuse nuclear and cytoplasmic ORC2 staining was observed (Royzman et al., 1999). While ORC2 staining was indetectable at chorion loci after stage 11, BrdU incorporation continued, and during stages 12 and 13, the BrdU pattern resolved into a double bar structure (Fig. 3E and Fig. 4E for BrdU, Fig. 5 C-D for lack of ORC2). Similar results were observed for ORC1 (unpublished results).

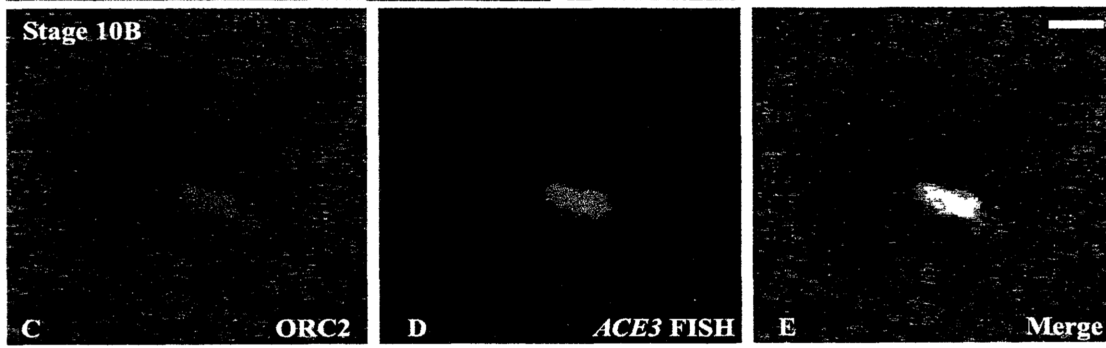
Deconvolution microscopy enabled us to measure the dimensions of the fluorescent signals at the third chromosome amplicon from stages 10B to stage 13. We examined the gap from the inside of one BrdU (or DUP, see below) signal to the inside of the second BrdU signal, the length of the bars, and the depth of each of the bars (Fig. 1I). Based on the onionskin or reinitiation model of chorion amplification (Botchan et al., 1979; Osheim et al., 1988) (Fig 1H, I) the gap should represent the extent of replication fork progression; the length, the number of origin firings; and the depth, the complexity of the onionskin as replication forks progress outward and are arranged in three-

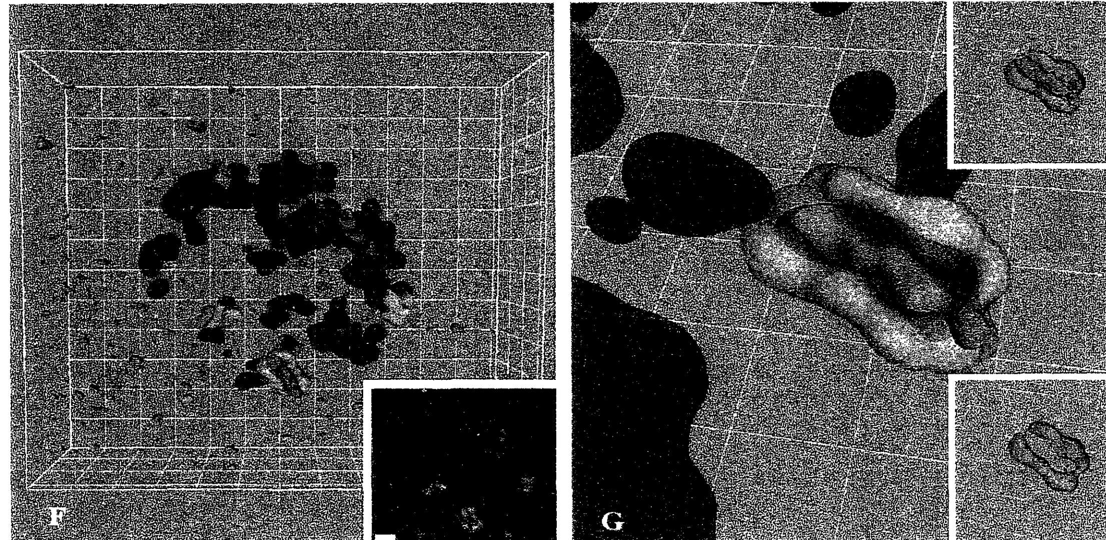
Figure 1.

ORC2 is present at chorion origins during amplification initiation but is lost from origins as initiation ends. (A, B) Deconvolution microscopy and volume rendering shows that in stage 10B follicle cell nuclei, ORC2 (green) partially colocalizes with BrdU (red) at the third chromsome chorion cluster. DNA is in blue (TOTO). BrdU foci without ORC2 localized correspond to uncharacterized sites of amplification throughout the genome, and the focus next to the third chromsome amplicon is likely the *X* cluster. Inset (A) shows the immunofluorescence image from which (A, B) were created. Insets in (B) show a close-up view of BrdU and ORC2 without DNA (top), and a close-up view of BrdU only (bottom). See also, Supplemental Movie 1. (C-E) Confocal microscopy shows that in stage 10B follicle cell nuclei, ORC2 (green) colocalizes with FISH signal from a 3.8kb third chromosome chorion probe (red) that spans *ACE3* and *oriβ*. (E) Shows the merged image; all images are in a single plane. (F, G) In stage 11 follicle cell nuclei, ORC2 (green) remains localized to origin regions of the third chromosome chorion locus, while BrdU (red) signal begins to resolve into bars as forks move outward. DNA is in blue (TOTO). Inset (F) shows the immunofluorescence image from which the images in F and G were created. Insets in (G) show a close-up view of BrdU and ORC2, without DNA (top), and a close-up view of BrdU only (bottom). (H) The onionskin/reinitiation model of chorion amplification representing the localization of ORC2 (olive) and incorporation of BrdU (salmon) in stage 10B and 11 follicle cells as initiation and limited elongation occur. (I) The onionskin/reinitiation model representing amplification by stages 12 and 13, when ORC2 is no longer localized and no further initiation events occur. Only existing replication forks move out and BrdU (salmon)

incorporated at these replication forks is seen as double bars. The dimensions used for deconvolution measurements are shown in (I). All scale bars represent 1 μm, and all grid bars, 1 μm^2. Online Supplemental Materials are available at: http://www.jcb.org/cgi/content/full/jcb.200207046/DC1. Movie 1 accompanies Fig. 1, A and B, and shows a three-dimensional volume rendering of ORC2 (green) in relation to BrdU (red) and DNA (blue).


A
B


Stage 10B
C
ORC2
D
ACE3 FISH
E
Merge


F
G

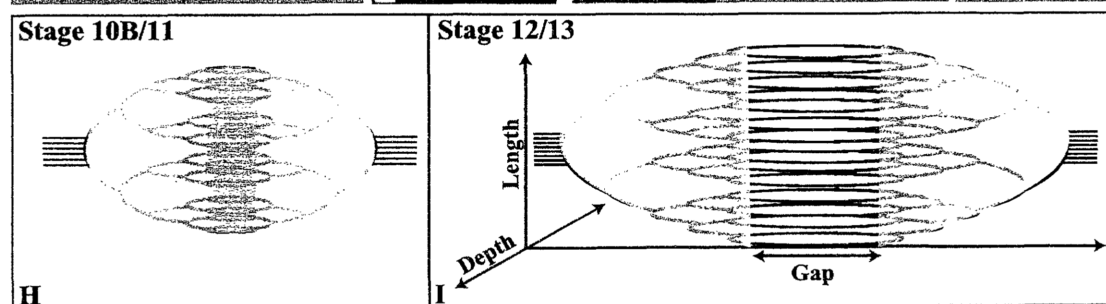

Stage 10B/11
H
Stage 12/13
Length
Depth
Gap
I

Table I. Deconvolution Microscopy Measurements of Chorion Amplicons

	Gap (nm)	Length (nm)	Depth (nm)
Stage 10B	300 +/- 30	1280 +/- 100	400 +/- 50
Stage 11	550 +/- 130	1760 +/-250	770 +/- 100
Stage 12	740 +/- 70	1740 +/-20	1040 +/- 170

Measurements were made based on 10-20 follicle cell nuclei at each stage, stained for either BrdU or DUP/Cdt1. For a more detailed description of the dimensions in reference to our model, see Figure 1I.

dimensions. The dimensions during stages 10B, 11, and 13 are summarized in Table 1. The length of the bars remained constant after stage 11 (1760 nm in stage 11 and 1740 nm in stage 13), suggesting that the maximum number of origin firings occurred by stage 11. The depth measurement increased dramatically throughout the later stages of amplification, from 400 nm in stage 10B to 1040 nm in stage 13. The gap measurement increased from 300 nm in stage 10B to 740 nm in stage 13. The gap measurement can be used to calculate the distance in kilobases the forks have progressed at a particular stage, with the conversion factor of 100nm~10kb. This conversion factor was calculated based on data by Calvi (Calvi and Spradling, 2001), in which the distance of two FISH probes 46kb apart and flanking *ACE3* was measured to be about 480 nm, giving the conversion factor of 480nm~46kb, or approximately 100nm~10kb. Thus, in stage 10B, replication forks have traveled a total distance of 30kb (an average of 15kb on either side of *ACE3*) and by stage 13 they have moved out across a 74kb total region (an average of 37kb on each side).

Considering the lack of ORC at chorion loci after stage 11, the essential role ORC plays in initiation, and the microscopy measurements, we propose that amplification can be separated into two phases. The first phase of amplification occurs during stages 10B and 11, is ORC-dependent, and involves initiation coupled with elongation (Fig. 1H). After this discrete period of initiation, ORC is lost from chorion origins and only the existing replication forks progress outward, in an elongation-only phase, to give the double bar structure seen in stages 12 and 13 (Fig. 1I).

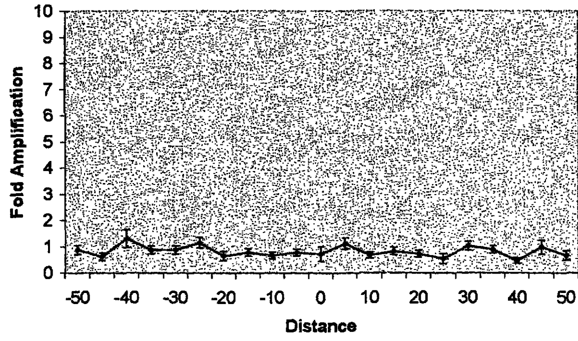
Quantitative realtime PCR measurement of DNA copy number along the third chromosome chorion amplicon

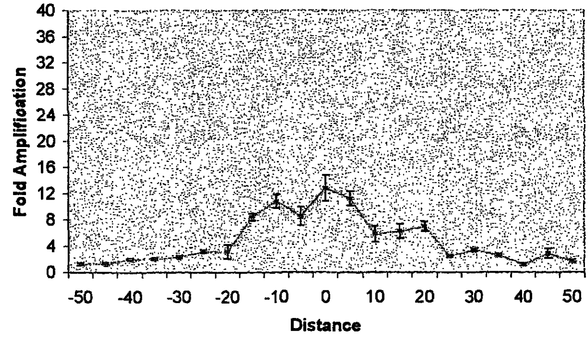
The immunofluorescence studies suggested that if the relative DNA copy number along the amplified regions were measured, a maximum copy number at origin sequences would be detected by stage 11. Furthermore, as replication forks progress outward by stages 12 and 13, we would expect to see a sequential increase in copy number of the loci proximal and distal to origins. To test this model, we used realtime PCR to quantify copy number in 5kb intervals along the third chromosome chorion locus during each stage of egg chamber development. (See Materials and Methods for a detailed description of Quantitative Realtime PCR).

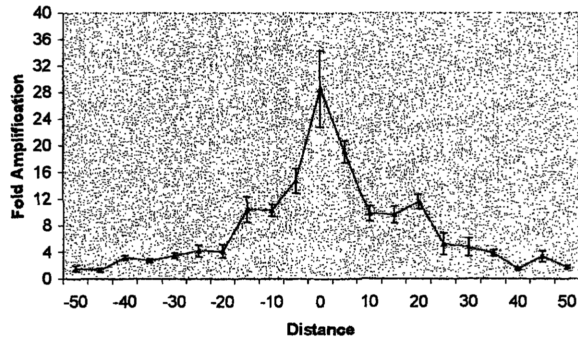
Quantitation of fold amplification in each of the stages allowed us to measure both inititiation and elongation events. In stage 1-8 (pre-amplification) egg chambers, no amplification was observed (Fig. 2A). Stage 10B egg chambers, in which chorion amplification has begun, showed increase in copy number at and around *ACE3*, (from 25 to –15kb) with a maximum of 15 fold amplification at 0kb (Fig. 2B). Loci proximal and distal to *ACE3*, from 25 to 35kb and –20 to –40kb, also showed some amplification during stage 10B (2-4 fold). This suggests that a subset of forks had replicated the entire amplicon. By stage 11, 30 fold amplification was observed at *ACE3*, as further rounds of initiation occurred. We did not observe integral doublings of copy number at *ACE3* between stages 10B and 11, probably because pools of egg chambers were used, and the result obtained represents the average of the pool. An increase in copy number from approximately 25 to -20kb also was detected in stage 11 (Fig. 2C).

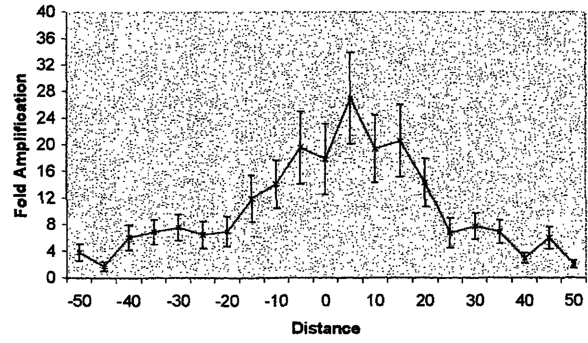
Figure 2.

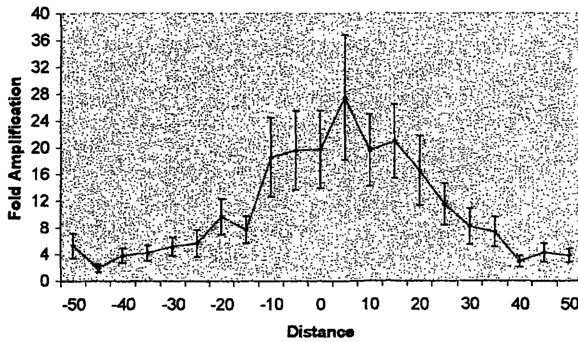
Quantitative Realtime PCR performed on staged egg chamber DNA confirms the timing of initiation and elongation. DNA from egg chambers prior to chorion amplification, stages 1-8, and during amplification, stages 10B, 11, 12, and 13 was used in quantitative realtime PCR reactions. Primer sets used for chorion loci spanned the third chromosome 50kb on either side of *ACE3* (denoted as 0 distance), in 5 kb intervals, and control primer sets (nonamplified) were to an intergenic region on chromosome arm 3R. The Y axis represents fold amplification, measured as the ratio of the chorion locus to the 3R locus and errors are the standard deviation of the sample. The X axis represents distance along the chorion locus in kilobases, with the major origin, *oriβ* located between 0 and 5kb. (A) In stage 1-8 egg chambers, no chorion amplification has occurred and the ratio of chorion to control loci is centered about 1. Note that the scale in 1-8 is different from the scale in (B-F). (B) By stage 10B, chorion gene amplification has initiated and there is an increase in fold amplification over approximately 35kb total. (C) By stage 11, additional initiation has occurred at the origins, as fold amplification increases to approximately 30. (D) During stage 12, no further increases in copy number are detected at origins, but an increase in fold amplification both proximal and distal to origins is detected. (E) By stage 13, replication forks have progressed out further, as an increase in fold amplification is detected out to about 35 and –40kb. No further initiation events occurred. The stage 13 reactions were performed on two separate samples of stage 13 DNA and similar results were observed in both trials (unpublished results). (F) A composite graph of (A-E) showing fold amplification at the third chromosome chorion locus throughout egg chamber development.

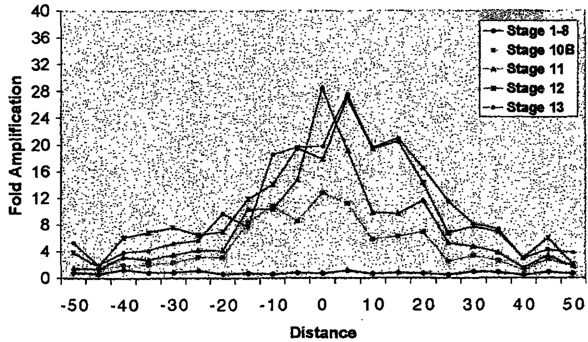

A.
Stages 1-8
Fold Amplification
Distance


B.
Stage 10B
Fold Amplification
Distance


C.
Stage 11
Fold Amplification
Distance


D.
Stage 12
Fold Amplification
Distance


E.
Stage 13
Fold Amplification
Distance


F.
Composite
Stage 1-8
Stage 10B
Stage 11
Stage 12
Stage 13
Fold Amplification
Distance

Strikingly, in stage 12 and 13 reactions (Figs. 2D and 2E, respectively), no further increase in copy number was detected at *ACE3*, with 29 and 27 fold amplification, respectively. This indicates that no further initiation occurred. In contrast, at loci proximal and distal to *ACE3*, an increase in copy number was detected as the existing replication forks progressed outward to about –40 and 35kb. For example, at 35kb, 9 fold chorion amplification was detected in stage 13 and 7.5 fold chorion amplification was detected in stage 12, as compared to 4 and 2 fold in stages 11 and 10B, respectively. We observed only half the maximum level of amplification detected by Spradling's original quantitative Southern blots (~30 versus 64 fold) (Spradling, 1981), probably because of the increased sensitivity of fluorescent PCR detection and the uniformity of the intervals used to measure amplification here.

When data from all stages are compared (Fig. 2F), it is clear that the final rounds of initiation occur between stages 10B and 11 and the copy number of flanking regions increases throughout subsequent stages. The results in Figure 2 were obtained using the 3R non-amplified control to determine fold amplification, and similar results were observed using the *ry* control (unpublished results).

Localization patterns of PCNA and MCM2-7 during chorion amplification

Both lines of data described above indicate that initiation and elongation occur simultaneously during one phase of chorion amplification, while only elongation occurs during a separate developmental phase. As an additional test of this hypothesis, we

studied the localization patterns of replication factors known to travel with the replication forks, PCNA and MCM2-7.

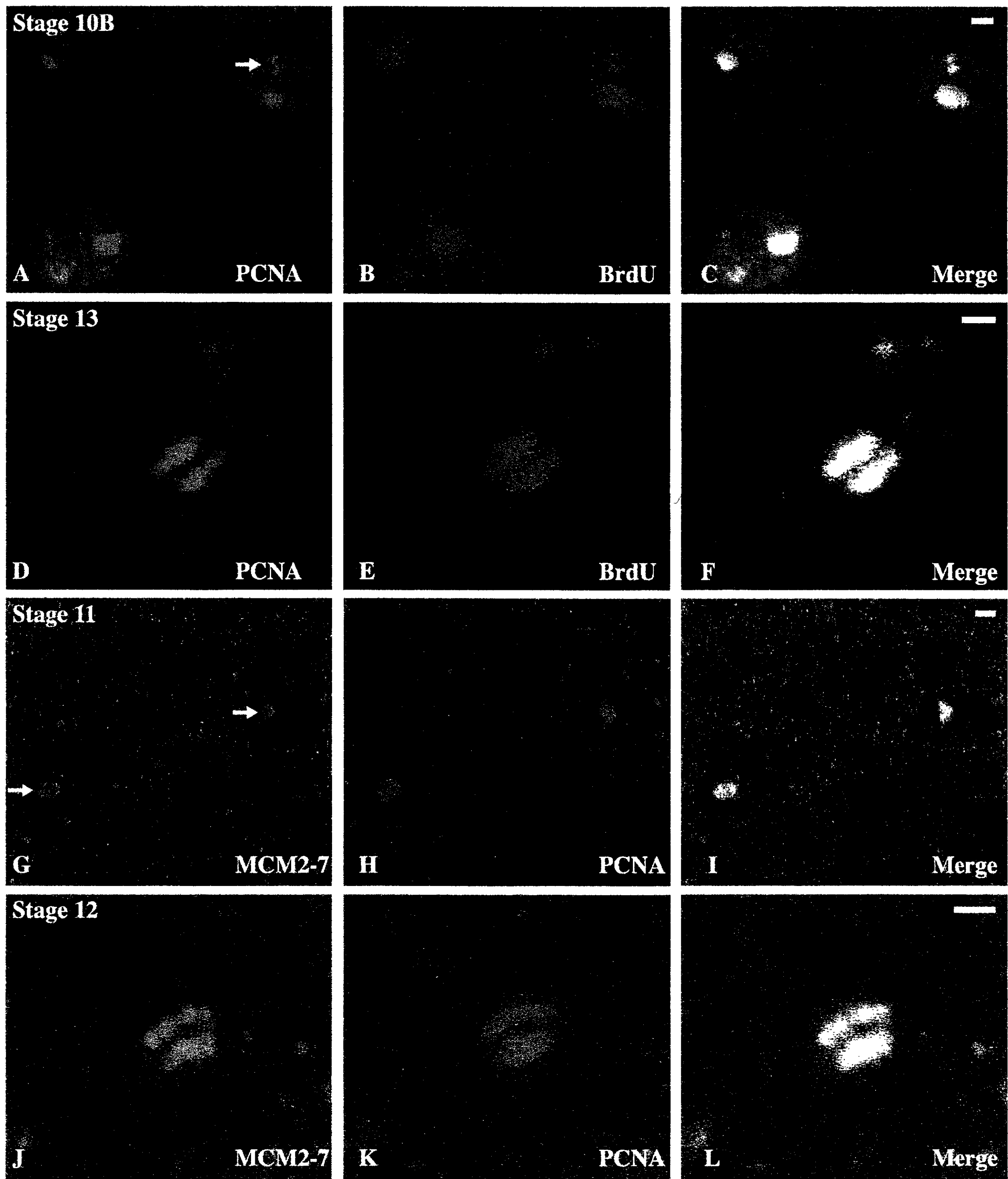
We observed a compelling pattern of PCNA localization in follicle cell nuclei. PCNA was nuclear throughout stages 1-9 (unpublished results), but by stage 10B foci of PCNA staining were detected above faint nuclear staining (Fig. 3A, C). As chorion amplification proceeded, PCNA remained localized and resolved into the double bar structure (Fig. 3D, F). To ensure that PCNA was localized to chorion regions, co-labeling with BrdU was performed, and PCNA was shown to colocalize with BrdU (Fig. 3 A-C. and D-F). These data support the idea that the double bar structure arises from fronts of bidirectional replication fork movement.

Previously, polyclonal antibodies raised against MCM2, 4, and 5 (Su and O'Farrell, 1997; Su et al., 1997; Su and O'Farrell, 1998) showed nuclear staining with no localization to chorion foci (Royzman et al., 1999). This was true even when egg chambers were treated with a high salt, high detergent buffer in an attempt to remove non-chromatin bound MCMs from the nucleus (Schwed et al., 2002). We re-examined the localization of the MCM2-7 complex during amplification using a monoclonal antibody that recognizes an epitope present in all six MCM subunits (Jayson Bowers and Stephen Bell, personal communication), thereby enhancing the sensitivity of detection. Additionally, egg chambers were washed with a high salt, high detergent buffer to remove non-chromatin bound MCM proteins from the nuclei.

Under these conditions we saw MCM2-7 localized throughout amplification. MCM2-7 localization first became visible as foci in stage 10B (Fig. 3G, I) and progressed to the double bar structure by stages 12 and 13 (Fig. 3J, L). To confirm that MCM2-7

Figure 3.

PCNA and MCM2-7 staining patterns coincide with BrdU incorporation throughout amplification. In panels A-C and D-F, PCNA is in red, BrdU is in green. (A-C) Show several stage 10B follicle cell nuclei, in which initiation of amplification is coupled with elongation. In such nuclei, PCNA is present and colocalizes with BrdU incorporation at the *X* and the third chorion loci. The third chromosome is the larger of the foci (Calvi, 1998), and the *X* chromosome cluster (arrow) has already resolved into the double bar structure by this stage. In addition to being at the chorion loci, PCNA is diffusely present throughout the nucleus during this stage. (D-F) A single follicle cell nucleus from a stage 13 egg chamber shows this pattern of PCNA and BrdU staining, which is characteristic of replication fork movement. The 2 smaller foci of staining in this image may be the *X* chromosome amplicon. In (G-I) and (J-L), MCM2-7 are in red and PCNA is in green. (G-I) MCM2-7 and PCNA colocalize in stage 11 follicle cell nuclei (arrows represent third chromosome clusters in two nuclei). (J-L) MCM2-7 staining, like PCNA, persists throughout chorion amplification and resolves into the double bar structure by stage 12. One stage 12 nucleus is shown. All scale bars represent 1μm.


Stage 10B
A
PCNA
B
BrdU
C
Merge
Stage 13
D
PCNA
E
BrdU
F
Merge
Stage 11
G
MCM2-7
H
PCNA
I
Merge
Stage 12
J
MCM2-7
K
PCNA
L
Merge

were localized to the chorion regions, we co-stained with PCNA and observed colocalization throughout all stages of amplification (Fig. 3G-I and J-L). Thus, MCMs are present at chorion amplicons during initiation and persist throughout amplification, presumably moving with the replication forks. The correlation of MCM2-7, PCNA, and BrdU staining patterns supports our model for chorion amplification.

The localization pattern of DUP/Cdt1 during chorion amplification

We then characterized the properties of the pre-RC component, DUP/Cdt1 (Maiorano et al., 2000; Nishitani et al., 2000; Whittaker et al., 2000; Devault et al., 2002; Tanaka and Diffley, 2002;). DUP/Cdt1 requires ORC2 to localize to chorion origins (Whittaker et al., 2000) and DUP/Cdt1 homologs in yeast and Xenopus have been shown to interact with Cdc6/18 to load MCM2-7 onto origins (Maiorano et al., 2000; Nishitani et al., 2000; Tada et al., 2001; Devault et al., 2002; Tanaka and Diffley, 2002). In Xenopus extracts, fission yeast, and budding yeast, Cdt1 is dispensable after initiation (Maiorano et al., 2000; Nishitani et al., 2000; Devault et al., 2002; Tanaka and Diffley, 2002). Furthermore, Cdt1 appears to be lost from chromatin or the nucleus at the onset of S phase (Maiorano et al., 2000; Tanaka and Diffley, 2002). These data suggest that Cdt1 is not necessary after performing its role in pre-RC formation. In contrast, the initial description of DUP/Cdt1 staining during amplification showed that DUP/Cdt1 localized to chorion loci throughout amplification, and was present during stage 13 in the double bar structure (Whittaker et al., 2000). Therefore, we examined the localization pattern of DUP/Cdt1 during amplification in relation to BrdU and ORC2, using confocal and

deconvolution microscopy, to investigate whether DUP/Cdt1 could be traveling with replication forks.

DUP/Cdt1 colocalized with BrdU throughout amplification. In stage 10B, DUP/Cdt1 staining was detected as foci (Fig. 4A) that overlapped completely with BrdU staining (Fig. 4B, C). By stage 13, DUP/Cdt1 staining resolved into the double bar structure (Fig. 4 D) and was coincident with BrdU (Fig. 4. E, F). The fact that DUP/Cdt1 remained localized to chorion regions throughout the elongation phase suggests that DUP/Cdt1 travels with the replication forks.

We precisely localized DUP/Cdt1 with respect to ORC2 by deconvolution microscopy, and in contrast to the colocalization of DUP/Cdt1 and BrdU, the ORC2 and DUP/Cdt1 staining patterns diverged as amplification proceeded. In early stage 10B, ORC2 and DUP/Cdt1 staining overlapped (unpublished results), similar to the results with ORC2 and BrdU co-staining (Fig. 1 A, B). By late stage 10B and stage 11, however, DUP/Cdt1 staining became fainter at the origins and resolved into a coffee bean-like structure (Fig. 5 A, B and Supplemental Movie 2). This change in the DUP/Cdt1 localization pattern occurred while ORC2 remained bound to origins. By stage 13, DUP/Cdt1 was detected in the double bar structure, with no evidence of ORC2 staining at origins (Fig. 5 C, D). Similar results were seen for DUP/Cdt1 and ORC1 (unpublished results). The pattern of DUP/Cdt1 localization in relation to BrdU and the fact that DUP/Cdt1 clears from origin sequences while ORC2 remains bound strongly indicate that DUP/Cdt1 travels with elongating replication forks.

Figure 4.

DUP/Cdt1 colocalizes with BrdU throughout chorion amplification. In panels A-C and D-F, DUP is in red and BrdU is in green. (A-C) In a stage 10B egg chamber, DUP colocalizes at sites of chorion amplification with BrdU. Two follicle cell nuclei are shown. (D-F) The DUP staining pattern colocalizes with that of BrdU throughout subsequent stages of chorion amplification and resolves into the double bar structure by stage 13, as seen in this follicle cell nucleus. All scale bars equal 1 μm.

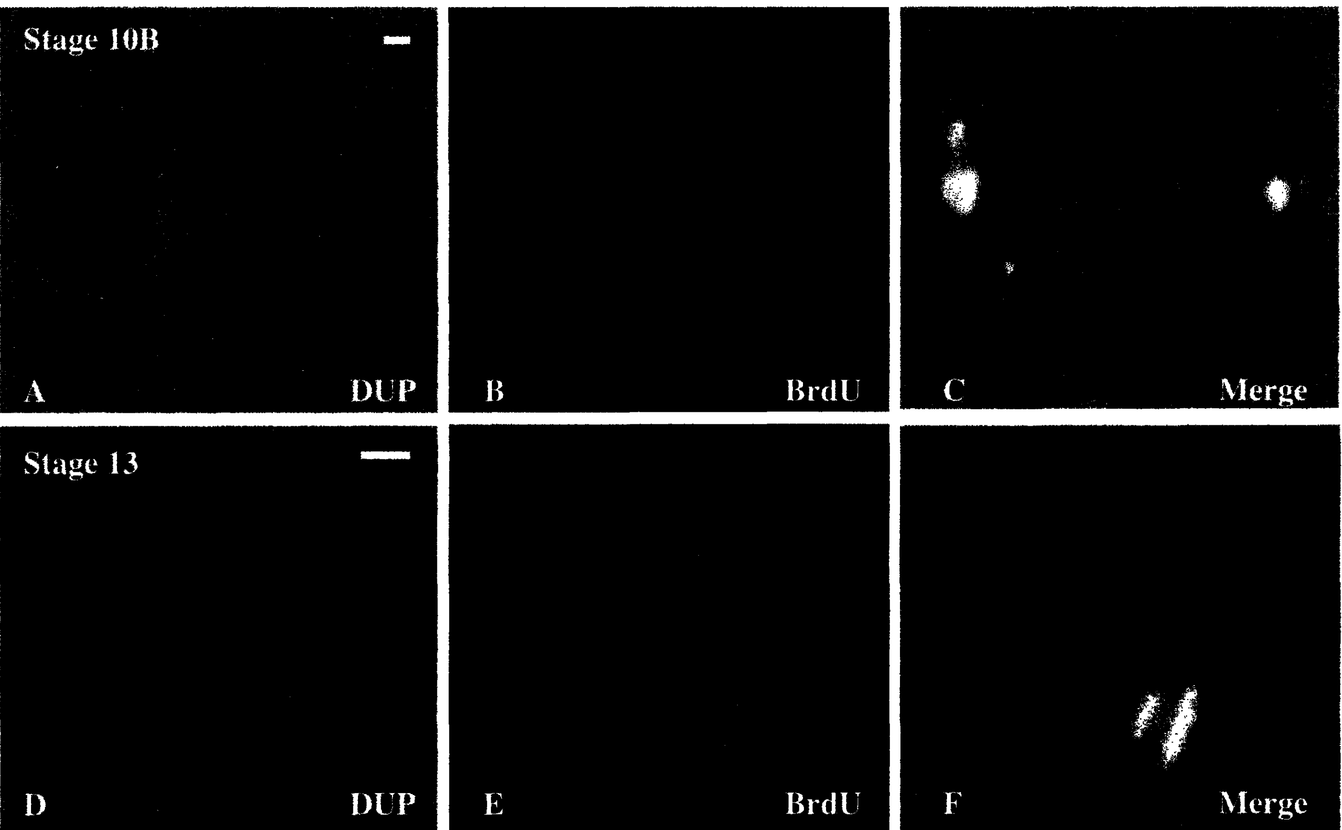
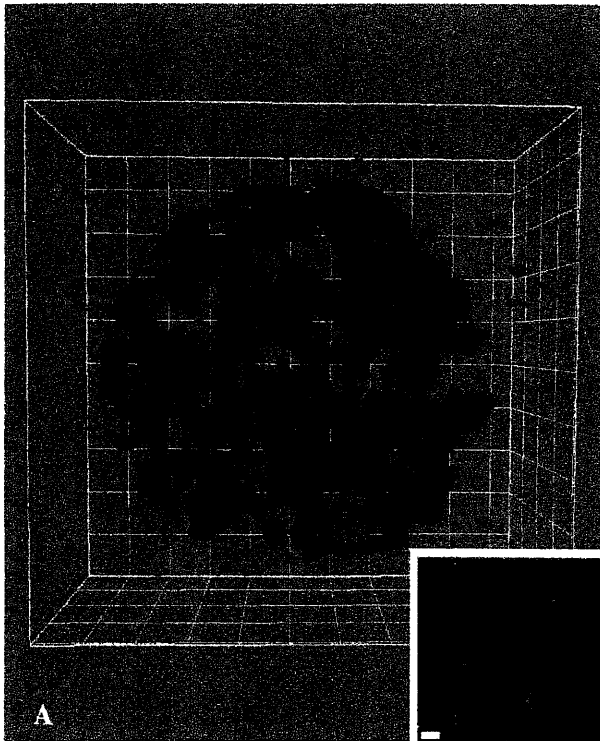
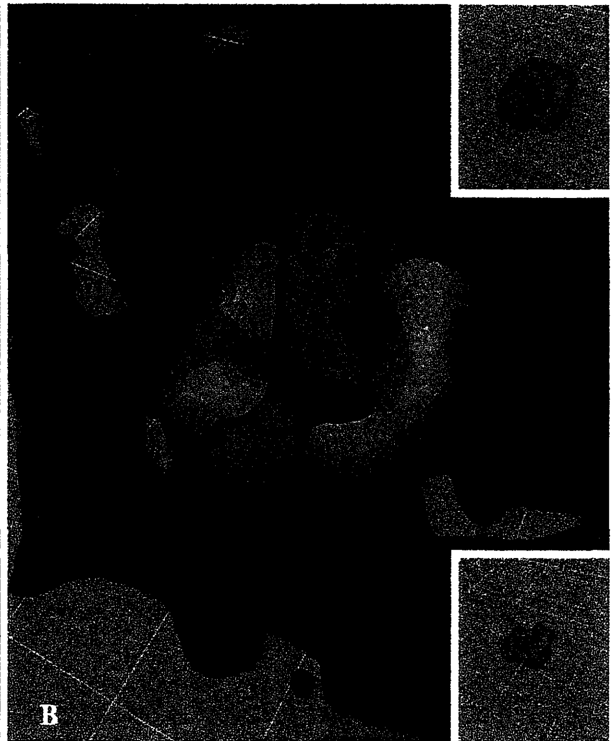
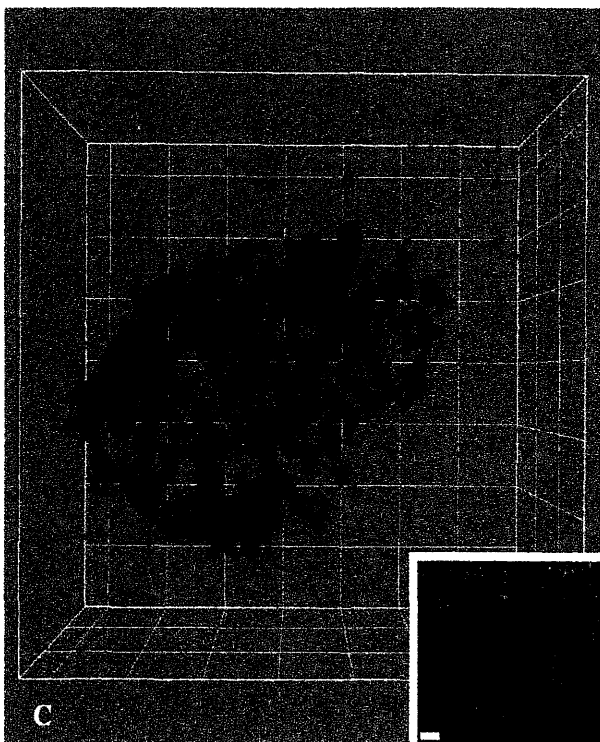

Stage 10B
A
DUP
B
BrdU
C
Merge
Stage 13
D
DUP
E
BrdU
F
Merge

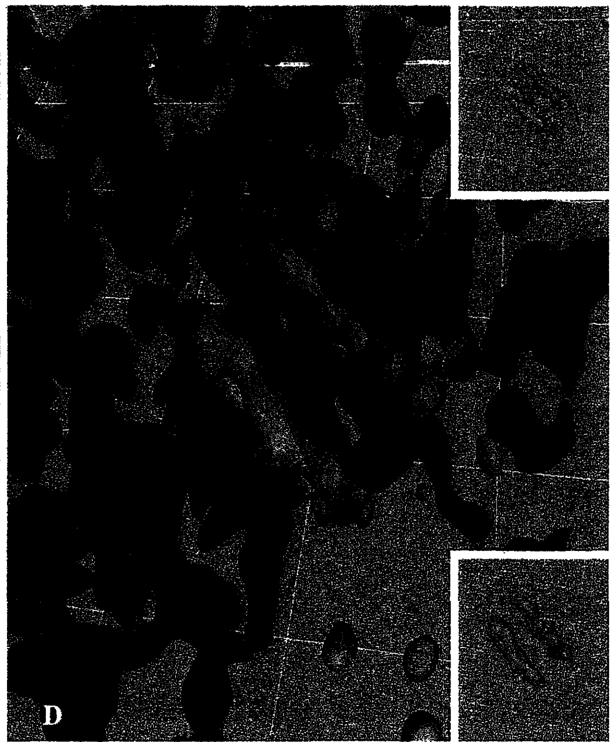
Figure 5.

The pattern of DUP and ORC2 localization indicates that DUP travels with replication forks. (A-B) Deconvolution microscopy and volume rendering of a stage 10B follicle cell nucleus shows that the patterns of DUP/Cdt1 (red) and ORC2 (green) slightly overlap at origins. DNA is in blue (TOTO). The relative amount of DUP/Cdt1 at the origins is less than the amount of DUP/Cdt1 in regions corresponding to fronts of replication fork movement. The inset in (A) shows the fluorescence image from which (A and B) were developed. The insets in (B) show a close-up of DUP/Cdt1 and ORC2 without the DNA (top), and a close-up view of DUP/Cdt1 alone (bottom). See also Supplemental Movie 2. (C-D) By stage 13, deconvolution microscopy and volume rendering shows that ORC2 (green) has been lost from origins, while DUP/Cdt1 (red) persists and resolves into the double bar structure. The inset in (C) shows the fluorescence image used to make (C-D) and the insets in (D) show a close-up view of the DUP/Cdt1 double bars in relation to ORC2 signal (top) and DUP/Cdt1 only (bottom). All scale bars represent 1 μm, and grid boxes, 1 μm^2. Online Supplemental Materials are available at: http://www.jcb.org/cgi/content/full/jcb.200207046/DC1. Movie 2 accompanies Fig. 5, A and B, and shows a three-dimensional volume rendering of DUP/Cdt1 (red) in relation to ORC2 (green) and DNA (blue).


A


B


C


D

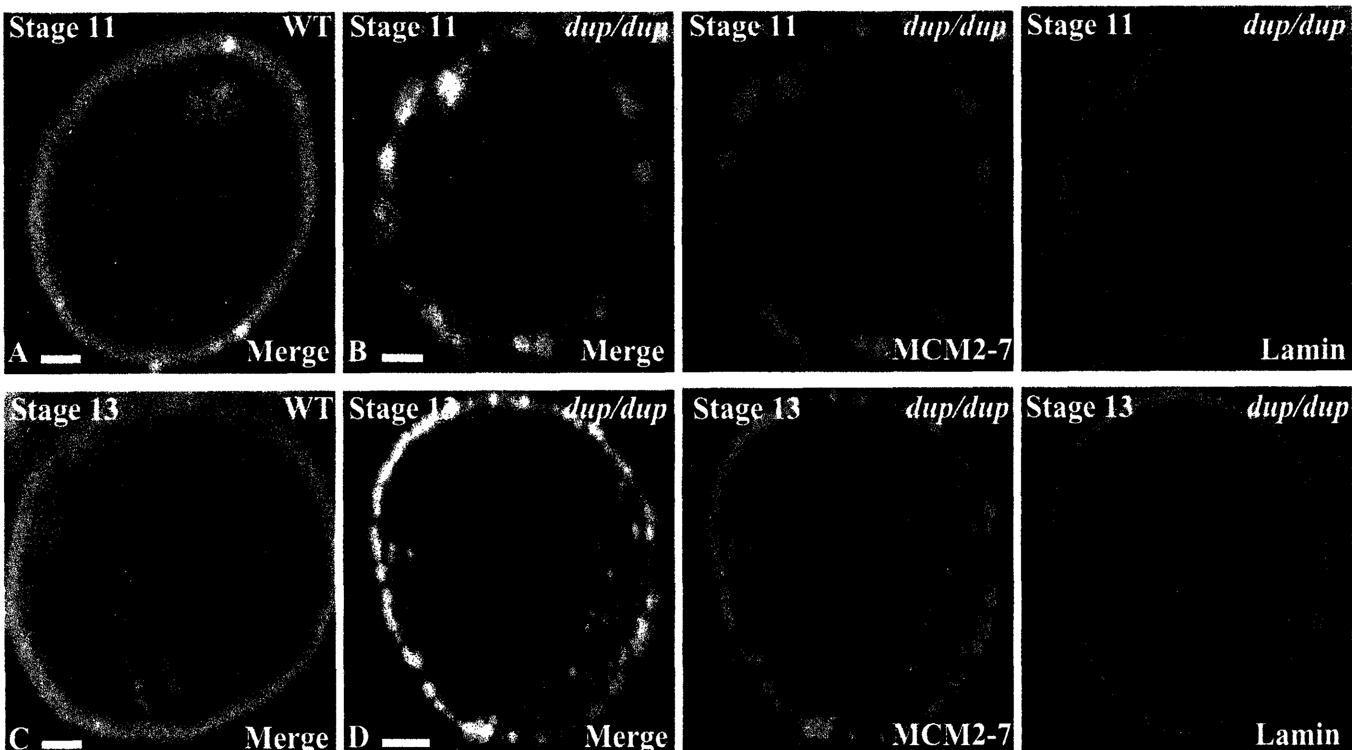
DUP/Cdt1 is necessary to localize MCM2-7 during amplification

Given the unexpected presence of DUP/Cdt1 during elongation, we wanted to know if DUP/Cdt1 functioned in this system to load MCM2-7 during initiation. To test this, we studied the localization pattern of MCM2-7 in the dup^{PA77} female-sterile mutant ovaries. These mutants have thin eggshells and decreased and delayed BrdU incorporation during amplification (Underwood et al., 1990; Whittaker et al., 2000).

In dup^{PA77} homozygous mutant ovaries, we did not detect the localization of MCM2-7 to chorion loci at any stage of amplification (Fig. 6). Furthermore, MCM2-7 appeared to cluster at the nuclear envelope, where it colocalized with nuclear lamins. These data indicate that DUP/Cdt1 is necessary to localize MCM2-7 to origins during chorion amplification, the same as the role of DUP/Cdt1 orthologs. The clustering of MCM2-7 at the nuclear periphery suggests that DUP/Cdt1 may be necessary for the nuclear transport of MCM2-7, consistent with the findings in *S. cerevisiae* that Cdt1 and MCM2-7 display an interdependence for nuclear trafficking (Tanaka and Diffley, 2002).

Figure 6.

MCM2-7 localization to chorion loci is disrupted in *dup*^{PA77}/*dup*^{PA77} mutants. In (A-C) MCM2-7 are in red and lamin is in green. (A, C) In wild-type follicle cells, MCM2-7 localize to chorion foci throughout the process of chorion amplification. (B, D) In contrast, in the *dup* female-sterile mutant, localization of MCM2-7 to chorion loci is not observed during any stage of amplification, and MCM2-7 cluster at the nuclear envelope. For (B) and (D), the MCM2-7 and lamin staining are shown separately to the right. At this level of resolution it is impossible distinguish whether MCM2-7 are trapped inside or outside of the nucleus. All images were captured at the same exposure for comparison. All scale bars represent 1 μm.


Stage 11
WT
A
Merge
Stage 11
dup/dup
B
Merge
Stage 11
dup/dup
MCM2-7
Stage 11
dup/dup
Lamin
Stage 13
WT
C
Merge
dup/dup
D
Merge
Stage 13
dup/dup
MCM2-7
Stage 13
dup/dup
Lamin

DISCUSSION

We demonstrated by three independent lines of evidence that initiation and the bulk of elongation at a chorion amplicon occur during two separate developmental periods. First, deconvolution microscopy shows that ORC and BrdU initially colocalize at origins and then diverge, as ORC is lost in stage 11 and BrdU resolves into a double bar structure. Second, elongation factors PCNA and MCM2-7 follow the same pattern as BrdU, resolving from foci early in amplification to a double bar structure by stage 12-13. Third, quantitative realtime PCR shows a peak increase in DNA copy number at the origins by stage 11, with increases in flanking sequences becoming substantial in stages 12 and 13. Thus initiation ends by stage 11, and during stages 12 and 13 only the existing forks progress outward. Furthermore, these observations led to the unanticipated conclusion that DUP/Cdt1 travels with replication forks.

Our realtime PCR and immunofluorescence data are remarkably consistent. First, both methods restrict initiation to stages 10B and 11, and elongation to stages 12 and 13. Between stages 10B and 11 the maximum fold amplification was detected at *ACE3* by realtime PCR, ORC localized to origins, and the deconvolution showed a maximum increase in bar length. During stages 12 and 13, increases in fold amplification were detected only proximal and distal to *ACE3*, and ORC no longer localized to origins while BrdU incorporation resolved into the double bar structure. Second, the distances of fork movement are consistent. Deconvolution measurements predicted that forks were maximally 30 +/- 3kb apart in stage 10B, and this correlates with the 40kb span of peak copy number detected by realtime PCR. In stage 11, forks were measured to have progressed across a 55 +/- 13kb region by deconvolution and across a 45kb region by

realtime PCR. By stage 13, deconvolution showed that replication forks were maximally separated by 74 +/- 7kb, whereas realtime PCR measured a 75kb span.

The convergence of the three lines of data argues against two alternative explanations for the immunofluorescence results. One alternate hypothesis is that ORC remains localized after stage 11, yet is not detectable because protein levels drop below detectable limits or epitopes become inaccessible. This is unlikely, as we observed the elongation factors PCNA, MCM2-7 and even DUP/Cdt1 change from a focus to double bar structure without a change in staining intensity. In contrast, during stage 11 ORC staining intensity decreased at origins, concomitant with a rise in nuclear and cytoplasmic levels. A second alternate hypothesis is that the double bar structures do not represent fork movement but result from firings of unidentified origins to either side of the *ACE3/oriβ* origin region. If this were the case, initiation events after stage 11 would occur independently of ORC, and the gradient profile from realtime PCR would be much different. As a result of these additional origins firing, the stage 12 graphs would show peaks of increased copy to either side of *ACE3/oriβ*, and by stage 13 the new forks would broaden the area of maximum copy number into a plateau.

The quantitative analysis of the amplification gradient provides insight into mechanisms affecting fork movement and termination and suggests that the onionskin structure (Botchan et al., 1979; Osheim et al., 1988) impedes fork movement. We calculated the maximal rate of fork movement during amplification to be 90 bp/min. on average, well within the 50-100bp/min. range calculated previously (Spradling and Leys, 1988). (By quantitative realtime PCR, the furthest a replication fork could travel is 40kb between stages 10B and 13, a period of 7.5 hours.) In comparison, replication forks in

the polytene larval salivary glands travel at approximately 300 bp/min (Steinemann, 1981), whreas rates of fork movement in both diploid *Drosophila* cell culture and embryo syncytial divisions are approximately 2.6 kb/min (Blumenthal et al., 1973). From these rates, it seems that polyteny hinders replication fork movement, an effect even more pronounced in amplification, given that the chorion cluster has a rate of fork movement three times less than polytene salivary glands. The fact that by stage 13 there is a gradient of copy number, and not a plateau further demonstrates the inefficiency of fork movement along the chorion cluster.

There do not seem to be specific termination sites to stop forks either along or at the ends of the chorion region, but fork movement may display some sequence or chromatin preference. The gradient of decreasing copy number implies that forks stop at a range of sites, as we would expect the presence of specific termination points along the region to cause steep drops in copy number. Despite this lack of specific termination sites, during stages 12 and 13 we see a greater increase in copy number to one side of *ACE3* (the right side in Fig. 2 graphs), and often observe by immunofluorescence that one of the two bars is shorter. This suggests that the sequence or chromatin structure to the other side of *ACE3* hinders fork movement, and as fewer forks move out, less BrdU incorporation occurs and a shorter bar results.

In contrast to other systems (Maiorano et al., 2000; Nishitani et al., 2000; Tada et al., 2001; Devault et al., 2002; Tanaka and Diffley, 2002), our results reveal that DUP/Cdt1 travels with replication forks during amplification. Although it could be argued that DUP/Cdt1 simply spreads along the chromatin as amplification proceeds, this is unlikely. DUP/Cdt1 and ORC2 colocalization studies show that although ORC2

remains at origins, the DUP/Cdt1 signal decreases at origins and subsequently flanks the ORC2 signal. Furthermore, during elongation DUP/Cdt1 does not spread across the entire chorion region. Rather, there is a gap between the double bars of DUP/Cdt1 staining which increases from 300 +/- 30 nm in stage 10B to 740 +/- 70 nm in stage 13.

The presence of DUP/Cdt1 at forks during elongation strongly suggests it has a role in this phase of replication. Why might DUP/Cdt1 be required during elongation in this system? Chorion amplification is unique because replication forks chase forks, instead of converging as in normal eukaryotic replication. Given this peculiarity of amplification, and considering the steric constraints that arise and impede forks, DUP/Cdt1 may be necessary to maintain MCM2-7 at these lethargic forks. DUP/Cdt1 could function as a processivity factor for the MCM2-7 complex, holding it on the DNA, or it could continuously re-load new MCM2-7 as they fall off the progressing replication forks. It is formally possible that although DUP/Cdt1 travels with the forks it does not perform a function. DUP/Cdt1 could simply not be expelled from the replication machinery upon initiation and then be dragged along during elongation. Although we do not favor this possibility, definitively proving that the DUP/Cdt1 at forks is necessary for elongation will require the use of a currently unavailable conditional allele. Such a mutation would permit inactivation of DUP/Cdt1 after initiation and allow a functional test for a role in elongation.

These studies highlight the complex regulation of chorion gene amplification. How are the number of origin firings restricted to the proper developmental time? It is known that the number of rounds of origin firing at the chorion amplicons is limited by the action of Rb, E2F1, and DP (Bosco et al., 2001). Perhaps DUP and MCM2-7 are also

a part of this regulation, with origins firing only when MCM2-7 are properly loaded. It will also be interesting to decipher the regulation of DUP/Cdt1 during amplification. Recent studies have demonstrated that a *Drosophila* homolog of the metazoan re-replication inhibitor, Geminin, exists and interacts biochemically and genetically with DUP/Cdt1 (Mihaylov et al., 2002; Quinn et al., 2001). Female-sterile mutations in *geminin* result in increased BrdU incorporation during amplification (Quinn et al., 2001), raising the possibility that Geminin acts on DUP/Cdt1 at the chorion loci to limit origin firing. In addition to permitting the delineation of the regulatory circuitry controlling origin firing, the ability to distinguish initiation from elongation developmentally provides a powerful tool for the analysis of the properties of metazoan replication factors *in vivo*.

MATERIALS AND METHODS

Fly Strains

Ovary stainings were performed on the Oregon-R wild-type strain unless otherwise noted. The *dup* mutant allele, *dup*^{PA77} was described previously (Underwood et al., 1990; Whittaker et al., 2000).

Immunofluorescence and BrdU labeling

Double labeling of *Drosophila* ovaries with anti-ORC2 and BrdU was performed as described previously (Royzman et al., 1999), with the following changes: BrdU was used at 6.4 μg/ml; secondary detection of ORC2 was with donkey anti-rabbit Rhodamine-RedX at 1:200; secondary detection of BrdU was with goat anti-mouse FITC at 1:200; and ovaries were mounted in Slowfade (Molecular Probes).

Double labeling of *Drosophila* ovaries with anti-PCNA (Henderson et al., 2000) and BrdU was performed as per anti-ORC2/BrdU, but incubating ovaries with anti-PCNA at 1:1000 and mounting in Vectashield (Vector Labs).

Labeling of *Drosophila* ovaries with anti-MCM2-7 was performed by first washing ovaries for 30 minutes in high salt buffer (50mM HEPES pH 7.5, 100mM NaCl, 1mM EDTA, 0.5% Triton X-100, 0.1% Na deoxycholate), then fixing with 8% EM grade formaldehyde, and processing as described for anti-ORC2 (Royzman et al., 1999). Ovaries were incubated with 1:200 anti-MCM2-7 overnight, and secondary detection was with donkey anti-mouse Cy-3 at 1:250. The anti-MCM2-7 is a monoclonal antibody, clone number AS1.1, which recognizes a conserved epitope in all MCM2-7 subunits (Klemm and Bell, 2001). When anti-PCNA, anti-MCM2-7 co-staining was performed,

ovaries were treated as described for anti-MCM2-7 labeling alone, and anti-PCNA was used at 1:1000 with anti-MCM2-7 overnight. Secondary detection of PCNA was with goat anti-rabbit FITC at 1:200, and ovaries were mounted in Vectashield. When anti-Dm_0Lamin (Gruenbaum et al., 1988), anti-MCM2-7 double labeling was performed, ovaries were treated as described, and anti-Dm_0Lamin was added at 1:200 in the primary incubation. Secondary detection of Dm_0Lamin was with goat anti-mouse FITC at 1:150 and ovaries were mounted in Vectashield.

Double labeling of *Drosophila* ovaries with anti-DUP and BrdU was performed as per anti-ORC2/BrdU labeling (above), but incubating ovaries with anti-DUP (Whittaker et al., 2000) at 1:1000 for 48 hours at 4°C. Secondary detection of DUP was performed with donkey anti-guinea pig Rhodamine-RedX at 1:200. Slides were mounted in Vectashield.

Anti-DUP, anti-ORC2 double labeling was performed as described previously (Whittaker et al., 2000), but with the following changes: the primary antibody incubation was performed for 48 hours at 4°C, and secondary detection was with donkey anti-guinea pig Rhodamine-RedX at 1:200 for anti-DUP and goat anti-rabbit FITC at 1:200 for anti-ORC2. Ovaries were mounted in Slowfade.

For some ovary samples, TOTO (Molecular Probes) was used to stain the DNA. These samples were treated as described above, but were treated with 1mg/ml RNase A (Sigma) for 1 hour at room temperature, then were incubated with a 1:2000 dilution of TOTO (Molecular Probes) for 10 minutes.

All secondary antibodies were from Jackson Immunoresearch.

All confocal imaging was performed using a Ziess Axiovert 100M with LSM510 software, using 63X Plan Neofluar or 100X Plan Neofluar objectives and with filters set according to the manufacturer's parameters.

Deconvolution Microscopy

Fluorescence data was collected using a Zeiss Axiovert 100M Meta confocal microscope with LSM510 software. Excitation of FITC, Rhodamine, and TOTO-1 dyes used the 488, 543 and 633 nm lasers, respectively. Emission filters were tuned to minimize bleedthrough between channels. Voxels were collected at 45 nm lateral and 1000 nm axial intervals. Deconvolution was carried out using the cMLE algorithm of Huygens2.3-professional (Scientific Volume Imaging) on an Silicon Graphics Origin 3400 server (SGI). Rendering and analysis of 3D data was carried using the MeasurementPro module of Imaris3 Surpass 3.2 (Bitplane).

Fluorescent *In Situ* Hybridization

ORC2 and *ACE3*-FISH co-labeling was performed as follows: Ovaries were stained for ORC2 as described (Royzman et al., 1999). Secondary detection of ORC2 was with donkey anti-rabbit Cy3 at 1:250. After staining for ORC2, ovaries were fixed (as per the ORC2/BrdU double labeling protocol), and were then processed for whole mount FISH as described (Calvi et al., 1998). The probe used for the third chromosome chorion locus was a 3.8kb *Sal*I fragment from the plasmid pT2, containing both *ACE3* and *oriβ*. The hybridized probe was detected with goat anti-DIG FITC at 1:200. Samples were mounted in Vectashield.

Isolation of *Drosophila* DNA for Quantitative Realtime PCR

Egg chamber staging was performed based on morphological markers as described (Spradling, 1993). Pools of approximately 400 or 500 egg chambers of each stage 10B, 11, 12, 13, and 130 ovaries of stage 1-8 were isolated from fattened Oregon-R females. DNA was isolated from the pools of egg chambers as described (Royzman et al., 1999), with the addition of RnaseA treatment (1mg/sample; Sigma) during the Proteinase K step.

Embryo genomic DNA was generated for use as standard curves in the realtime PCR reactions according to standard techniques (Ashburner, 1989).

Quantitative Realtime PCR

Quantitative Realtime PCR was performed using the ABI Prism 7000 Sequence Detection System with Qiagen SYBR Green PCR mix. Thermo-cycling was done for 35 cycles.

Primer sets spanning 50kb on either side of *ACE3* (denoted as distance 0) at 5kb intervals, primers to the non-amplified *rosy* (*ry*) locus, and primers to another non-amplified intergenic region on chromosome arm 3R (located approximately at cytological position 93F2, about 25kb upstream of the *polα* locus) were generated using Primer 3 software. Primers were designed to be 22bp on average, with an optimum Tm of 65°C, and yielding products of 85bp on average. Primers were supplied by IDT, and primer sequences are available upon request.

Each experimental reaction (per egg chamber stage, per primer set) was performed in triplicate, alongside four ten-fold dilutions of standard DNA (embryo genomic DNA) and no-template control reactions (all in triplicate). The same embryo genomic DNA samples were used in all control reactions for internal consistency. Each experimental reaction contained DNA from approximately one to one half of an egg chamber, and was done in 25μl total volume (12.5μl SYBR Green 2x Master Mix, 10μl dH20, 2μl DNA, 0.25μl each 25nmolar primer). Relative fluorescence was measured per sample in comparison to standard curves and standard deviations of the triplicate reactions were calculated by the ABI Prism 7000 software. Fold amplification was calculated by dividing relative fluorescence for one of the third chromosome amplicon products by the relative fluorescence of either the *ry* or the 3R non-amplified control product for a given stage. Error is expressed in terms of standard deviation, where the

Standard deviation of the ratio A/C= (FA/FC) *{ [(SA/FA)^2 + (SC/FC)^2]^.5}

A=amplicon locus

C=control locus

FA=relative fluorescence from amplicon locus

FC=relative fluorescence from control locus (*ry* or 3R)

SA=standard deviation from same amplicon locus

SC=standard deviation from same control locus

ACKNOWLEDGEMENTS

We are grateful to Anthony Schwaca for supplying the monoclonal MCM2-7 antibody and to Jayson Bowers for sharing unpublished results about that antibody. We thank Paul Fisher, Daryl Henderson, and Maki Asano for donations of anti-lamin, anti-PCNA, and anti-ORC1 antibodies, respectively. We are grateful for the assistance of Milan DeVries in analyzing the realtime PCR data, and we thank Aaron Aslanian, Giovanni Bosco, Alex Ensminger, Irena Ivanovska, Gwen Wilmes, and Leah Vardy for their comments on this manuscript. The microscopy was conducted using the W.M. Keck Foundation Biological Imaging Facility at the Whitehead Institute.

D. M. M. is a fellow of the Damon Runyon-Walter Winchell Cancer Fund. This work was funded by NIH grant GM58701 and support from the Howard Hughes Medical Institute to S. P. B., and by NIH grant GM57960 to T. O.-W.

REFERENCES

Aparicio, O.M., A.M. Stout, and S.P. Bell. 1999. Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proc. Natl. Acad. Sci. USA*. 96:9130-9135.

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

Asano, M., and R.P. Wharton. 1999. E2F mediates developmental and cell cycle regulation of ORC1 in *Drosophila*. *EMBO J*. 18:2435-2448.

Ashburner, M. 1989. *Drosophila*: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Austin, R.J., T.L. Orr-Weaver, and S.P. Bell. 1999. *Drosophila* ORC specifically binds to *ACE3*, an origin of DNA replication control element. *Genes and Dev*. 13:2639-2649.

Bell, S.P., and A. Dutta. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71:333-374.

Bell, S.P., and B. Stillman. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*. 357:128-134.

Bielinsky, A.K., H. Blitzblau, E.L. Beall, M. Ezrokhi, H.S. Smith, M.R. Botchan, and S.A. Gerbi. 2001. Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.* 11:1427-1431.

Bielinsky, A.K., and S.A. Gerbi. 2001. Where it all starts: eukaryotic origins of DNA replication. *J. Cell Sci*. 114:643-651.

Blow, J.J. 2001. Control of chromosomal DNA replication in the early Xenopus embryo. *Embo J.* 20:3293-3297.

Blumenthal, A., H. Kriegstein, and D. Hogness. 1973. The units of DNA replication in *Drosophila melanogaster* chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* 38:205-223.

Bosco, G., W. Du, and T.L. Orr-Weaver. 2001. DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nature Cell Biol.* 3:289-295.

Botchan, M., W. Topp, and J. Sambrook. 1979. Studies on simian virus 40 excision from cellular chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* 43 Pt 2:709-719.

Calvi, B., and A. Spradling. 2001. The nuclear location and chromatin organization of active chorion amplification origins. *Chromosoma.* 110:159-172.

Calvi, B.R., M.A. Lilly, and A.C. Spradling. 1998. Cell cycle control of chorion gene amplification. *Genes and Dev.* 12:734-744.

Calvi, B.R., and A.C. Spradling. 1999. Chorion gene amplification in *Drosophila*: A model for origins of DNA replication and S phase control. *In* Genetic approaches to eukaryotic replication and repair. Vol. 18. P. Fisher, editor. Academic Press, New York, New York. 407-417.

Carminati, J.L., C.J. Johnston, and T.L. Orr-Weaver. 1992. The *Drosophila ACE3* chorion element autonomously induces amplification. *Mol. Cell. Biol.* 12:2444-2453.

Chesnokov, I., M. Gossen, D. Remus, and M. Botchan. 1999. Assembly of functionally active *Drosophila* origin recognition complex from recombinant proteins. *Genes and Dev.* 13:1289-1296.

de Cicco, D., and A. Spradling. 1984. Localization of a cis-acting element responsible for the developmentally regulated amplification of *Drosophila* chorion genes. *Cell.* 38:45-54.

Delidakis, C., and F.C. Kafatos. 1989. Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila. EMBO J.* 8:891-901.

DePamphilis, M.L. 1999. Replication origins in metazoan chromosomes: fact or fiction. *BioEssays.* 21:5-16.

Devault, A., E.A. Vallen, T. Yuan, S. Green, A. Bensimon, and E. Schwob. 2002. Identification of Tah11/Sid2 as the Ortholog of the Replication Licensing Factor Cdt1 in *Saccharomyces cerevisiae. Curr. Biol.* 12:689-694.

Donaldson, A.D., and J.J. Blow. 1999. The regulation of replication origin activation. *Curr. Opin. Genet. and Dev.* 9:62-68.

Dutta, A., and S.P. Bell. 1997. Initiation of DNA replication in eukaryotic cells. *Annu. Rev. Cell. Dev. Biol.* 13:293-332.

Gruenbaum, Y., Y. Landesman, B. Drees, J.W. Bare, H. Saumweber, M.R. Paddy, J.W. Sedat, D.E. Smith, B.M. Benton, and P.A. Fisher. 1988. *Drosophila* nuclear lamin precursor Dm0 is translated from either of two developmentally regulated mRNA species apparently encoded by a single gene. *J. Cell Biol.* 106:585-596.

Heck, M., and A. Spradling. 1990. Multiple replication origins are used during *Drosophila* chorion gene amplification. *J. Cell Biol.* 110:903-914.

Henderson, D.S., U.K. Wiegand, D.G. Norman, and D.M. Glover. 2000. Mutual correction of faulty PCNA subunits in temperature-sensitive lethal *mus209* mutants of *Drosophila melanogaster. Genetics.* 154:1721-1733.

Ishimi, Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J. Biol. Chem.* 272:24508-24513.

Klemm, R.D., and S.P. Bell. 2001. ATP bound to the origin recognition complex is important for preRC formation. *Proc. Natl. Acad. Sci. USA*. 98:8361-8367.

Labib, K., J.A. Tercero, and J.F. Diffley. 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science*. 288:1643-1647.

Landis, G., R. Kelley, A.C. Spradling, and J. Tower. 1997. The *k43* gene, required for chorion gene amplification and diploid cell chromosome replication, encodes the *Drosophila* homolog of yeast origin recognition complex subunit 2. *Proc. Natl. Acad. Sci. USA*. 94:3888-3892.

Landis, G., and J. Tower. 1999. The *Drosophila chiffon* gene is required for chorion gene amplification, and is related to the yeast dbf4 regulator of DNA replication and cell cycle. *Development*. 126:4281-4293.

Lei, M., Y. Kawasaki, M.R. Young, M. Kihara, A. Sugino, and B.K. Tye. 1997. Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes and Dev*. 11:3365-3374.

Loebel, D., H. Huikeshoven, and S. Cotterill. 2000. Localisation of the DmCdc45 DNA replication factor in the mitotic cycle and during chorion gene amplification. *Nucleic Acids Res*. 28:3897-3903.

Lu, L., Z. Hongjun, and J. Tower. 2001. Functionally distinct, sequence specific replicator and origin elements are required for *Drosophila* chorion gene amplification. *Genes and Dev*. 15:134-146.

Maiorano, D., J. Moreau, and M. Mechali. 2000. XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. *Nature*. 404:622-625.

Mendez, J., and B. Stillman. 2000. Chromatin association of human origin recognition complex, Cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* 20:8602-8612.

Merchant, A.M., Y. Kawasaki, Y. Chen, M. Lei, and B.K. Tye. 1997. A lesion in the DNA replication initiation factor Mcm10 induces pausing of elongation forks through chromosomal replication origins in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 17:3261-3271.

Mihaylov, I.S., T. Kondo, L. Jones, S. Ryzhikov, J. Tanaka, J. Zheng, L.A. Higa, N. Minamino, L. Cooley, and H. Zhang. 2002. Control of DNA replication and chromosome ploidy by geminin and cyclin A. *Mol. Cell. Biol.* 22:1868-1880.

Nishitani, H., Z. Lygerou, T. Nishimoto, and P. Nurse. 2000. The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature*. 404:625-628.

Osheim, Y.N., O.L. Miller, and A.L. Beyer. 1988. Visualization of *Drosophila melanogaster* chorion genes undergoing amplification. *Mol. Cell. Biol.* 8:2811-2821.

Quinn, L.M., A. Herr, T.J. McGarry, and H. Richardson. 2001. The *Drosophila* Geminin homolog: roles for Geminin in limiting DNA replication, in anaphase and in neurogenesis. *Genes and Dev.* 15:2741-2754.

Royzman, I., R.J. Austin, G. Bosco, S.P. Bell, and T.L. Orr-Weaver. 1999. ORC localization in *Drosophila* follicle cells and the effects of mutations in *dE2F* and *dDP*. *Genes and Dev.* 13:827-840.

Schwed, G., N. May, Y. Pechersky, and B.R. Calvi. 2002. *Drosophila* minichromosome maintenance 6 is required for chorion gene amplification and genomic replication. *Mol. Biol. Cell.* 13:607-620.

Spradling, A. 1981. The organization and amplification of two clusters of *Drosophila* chorion genes. *Cell.* 27:193-202.

Spradling, A.C. 1993. Developmental genetics of oogenesis. *In* The development of *Drosophila melanogaster*. Vol. 1. M. Bate and A. Martinez Arias, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1-70.

Spradling, A.C., and E. Leys. 1988. Slow replication fork movement during *Drosophila* chorion gene amplification. *In* Cancer Cells. Vol. 6. T. Kelly and B. Stillman, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor. 305-309.

Spradling, A.C., and A.P. Mahowald. 1981. A chromosome inversion alters the pattern of specific DNA replication in *Drosophila* follicle cells. *Cell.* 27:203-209.

Steinemann, M. 1981. Chromosomal replication in *Drosophila* virilis. III. Organization of active origins in the highly polytene salivary gland cells. *Chromosoma.* 82:289-307.

Su, T.T., and P.H. O'Farrell. 1997. Chromosome association of minichromosome maintenance proteins in *Drosophila* mitotic cycles. *J. Cell Biol.* 139:13-21.

Su, T.T., and P.H. O'Farrell. 1998. Chromosome association of minichromosome maintenance proteins in *Drosophila* endoreplication cycles. *J. Cell Biol.* 140:451-460.

Su, T.T., N. Yakubovich, and P.H. O'Farrell. 1997. Cloning of *Drosophila* MCM homologs and analysis of their requirement during embryogenesis. *Gene.* 192:283-289.

Tada, S., A. Li, D. Maiorano, M. Mechali, and J.J. Blow. 2001. Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat. Cell Biol.* 3:107-113.

Tanaka, S., and J.F. Diffley. 2002. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2--7 during G1 phase. *Nat. Cell Biol.* 4:198-207.

Tercero, J.A., K. Labib, and J.F. Diffley. 2000. DNA synthesis at individual replication forks requires the esential initiation factor Cdc45p. *EMBO J.* 19:2082-2093.

Underwood, E.M., A.S. Briot, K.Z. Doll, R.L. Ludwiczak, D.C. Otteson, J. Tower, K.B. Vessey, and K. Yu. 1990. Genetics of 51D-52A, a region containing several maternal-effect genes and two maternal-specific transcripts in *Drosophila*. *Genetics*. 126:639-650.

Waga, S., and B. Stillman. 1998. The DNA replication fork in eukaryotic cells. *Annu Rev Biochem.* 67:721-751.

Whittaker, A.J., I. Royzman, and T.L. Orr-Weaver. 2000. *Drosophila* double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. *Genes and Dev.* 14:1765-1776.

Wohlschlegel, J.A., S.K. Dhar, T.A. Prokhorova, A. Dutta, and J.C. Walter. 2002. Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45. *Mol. Cell.* 9:233-240.

Zou, L., and B. Stillman. 2000. Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Mol. Cell. Biol.* 20:3086-3096.

Chapter Three

Gene Amplification as a Developmental Strategy: Isolation of Two Developmental Amplicons in *Drosophila*

Julie M. Claycomb, Matt Benasutti, Giovanni Bosco[1], Douglas D. Fenger[2], and Terry L. Orr-Weaver*

Whitehead Institute and Dept. of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

1 Present Address: University of Arizona Cancer Center, Tucson, AZ 85719
2 Present Address: Amylin Pharmaceuticals, San Diego, CA 92121

This chapter was published in Developmental Cell 6(1): 145-155 (2004).

Julie M. Claycomb performed the DAFC-30B and 62D real-time PCR, FISH/BrdU immunofluorescence and confocal microscopy, and the RNA *in situ* hybridizations for the amplified genes. She found the *yellow-g* mutant in the Baylor database and performed RNA *in situ* hybridization and neutral red staining on the mutant. She also performed RNA *in situ* hybridization and BrdU/FISH labeling on replication factor mutant ovaries.

ABSTRACT

Gene amplification is known to be critical for upregulating gene expression in a few cases, but the extent to which amplification is utilized in the development of diverse organisms remains unknown. By quantifying genomic DNA hybridization to microarrays to assay gene copy number, we identified two additional developmental amplicons in the follicle cells of the *Drosophila* ovary. Both amplicons contain genes which, following their amplification, are expressed in the follicle cells, and the expression of three of these genes becomes restricted to specialized follicle cells late in differentiation. Genetic analysis establishes that at least one of these genes, *yellow-g*, is critical for follicle cell function, because mutations in *yellow-g* disrupt eggshell integrity. Thus, during follicle cell differentiation the entire genome is overreplicated as the cells become polyploid, and subsequently specific genomic intervals are overreplicated to facilitate gene expression.

INTRODUCTION

Some tissues require the production of massive amounts of particular gene products during periods of development so brief that increased transcription alone is insufficient. One mechanism by which sufficient gene expression can be achieved is via amplification of the genes prior to their transcription, leading to an increase in the amount of template available for transcription. Such developmentally-regulated gene amplification is employed for the ribosomal RNA genes in amphibian oocytes to facilitate stockpiling of the oocyte with ribosomes (Brown and Dawid, 1968; Gall, 1968), as well as during the production of the macronucleus in *Tetrahymena* (for review see Prescott (Prescott, 1994)). In the Sciarid flies, the salivary gland rapidly synthesizes structural proteins for cocoons by amplifying these genes (Glover et al., 1982; Rudkin and Corlette, 1957; Wu et al., 1993). In the follicle cells of *Drosophila melanogaster*, the genes for the structural proteins of the eggshell (chorion) are amplified prior to their transcription (Spradling, 1981). In the latter case, amplification is essential for adequate levels of gene expression because mutations that reduce amplification cause thin eggshells and female sterility (Orr et al., 1984).

Genomic technologies now provide the opportunity to determine the global use of gene amplification during development. The last developmental amplicon was identified over 20 years ago. Early identification of amplified DNA relied on cytological evidence such as DNA puffs in the case of the Sciarid flies or extra chromosomal DNA circles in oocytes. Subsequently, amplicons were identified by testing DNA clones encoding developmentally expressed genes for increased gene copy number during differentiation,

but this methodology was employed only sporadically (for review see Spradling and Orr-Weaver (Spradling and Orr-Weaver, 1987)).

The developmental requirement for chorion gene amplification and its role in follicle cell differentiation are understood (for review see Orr-Weaver (Orr-Weaver, 1991)). The genes encoding six of the major structural components of the eggshell are clustered on the *X* chromosome at cytological location 7F and on the third chromosome at 66D. The follicle cells first synthesize and secrete the vitelline membrane proteins onto the oocyte surface, then secrete the chorion proteins to build a multi-layered eggshell. Prior to the transcription of the major chorion protein genes, the genomic intervals containing these genes are amplified to increase the amount of template available for transcription. The genes encoding the vitelline membrane proteins as well as those encoding other minor chorion proteins, however, are not amplified in the follicle cells (Higgins et al., 1984; Popodi et al., 1988). Instead, they are transcribed over a longer period of approximately fifteen hours (Mahowald and Kambysellis, 1980), as compared with the major chorion protein genes that must be transcribed over approximately two to three hours (Parks and Spradling, 1987).

In addition to being an intriguing developmental paradigm, analysis of insect amplicons has provided key insights into the regulation of metazoan DNA replication. At these sites, gene amplification occurs by repeated firing of replication origins within the gene clusters and movement of replication forks to produce a gradient of amplified DNA. In *Drosophila*, cis-acting control elements have been delineated by transformation experiments (Carminati et al., 1992; de Cicco and Spradling, 1984; Lu et al., 2001), and in both *Drosophila* and *Sciara* the positions of the replication origins used in

amplification have been mapped (Bielinsky et al., 2001; Delidakis and Kafatos, 1989; Heck and Spradling, 1990). Initiation factors and proteins utilized for normal genomic replication are essential for amplification (for review see Bosco and Orr-Weaver (Bosco and Orr-Weaver, 2002)). In fact, the Origin Recognition Complex (ORC) was first demonstrated to bind specific metazoan genomic sequences at the third chromosome chorion amplicon (Austin et al., 1999). ORC also binds to key replication elements in the *Sciara* amplicon (Bielinsky et al., 2001). Mutations that disrupt amplification have led to the identification of new replication factors (Landis and Tower, 1999; Whittaker et al., 2000). In *Drosophila* follicle cells, all of the replication initiation events for the chorion amplicons occur in two specific stages of egg chamber development. During subsequent stages replication forks elongate in the absence of additional initiation events (Claycomb et al., 2002). This developmental separation of initiation and elongation permits distinct roles for proteins in these two processes to be distinguished.

The adaptation of microarrays to measure gene copy number provides a means to screen for gene amplification events across the genome throughout development. The recovery of additional amplicons provides model replicons and serves as a means to identify genes whose functions are crucial at particular developmental points. It is of interest to learn what types of proteins, in addition to ribosomal proteins or structural proteins of the eggshell or cocoon, require amplification in some contexts. Moreover, given the role that amplification of oncogenes plays in cell transformation (Gray and Collins, 2000), it is conceivable that amplification of cell division or tumor suppressor genes could regulate replication or division during differentiation. Here we use

microarrays to identify two follicle cell amplicons and demonstrate that these encode genes expressed in and essential for follicle cell function.

RESULTS

Identification of Amplified Genes on Microarrays

We developed a technique to identify sites of DNA amplification by copy number comparison using microarrays that were simultaneously hybridized with experimental genomic DNA labeled with one fluorochrome and control genomic DNA labeled with another. We focused our studies on gene-encoding regions of the genome by producing microarrays of cDNAs from the *Drosophila* Gene Collection. This collection contains unique full length cDNAs for 5928 genes and represents 42% of the predicted protein-coding genes in the *Drosophila* genome (Rubin et al., 2000). In the two known follicle cell amplicons the peak amplification levels are at the sites of the chorion genes, but there are gradients of increased DNA copy number extending about 50kb to either side (Spradling, 1981). Thus we reasoned that, based on the average gene density in *Drosophila*, we were likely to detect genes within an amplified domain, even if all of the predicted genes were not present on the microarrays. The cDNA inserts for each clone were PCR amplified, and the PCR products were spotted onto slides to generate the arrays. The arrays were simultaneously hybridized with Cy-5 labeled control genomic DNA from 0-2 hour (2C) embryos and Cy-3 labeled genomic DNA from FACS sorted 16C follicle cell nuclei. The follicle cells become polyploid before undergoing chorion gene amplification, thus the 16C population is enriched for amplified DNA.

The ratio of the hybridization signal between the follicle cell genomic DNA probe and the control embryonic genomic DNA probe revealed increases in copy number of specific genes in follicle cells. The experiment was repeated three times to ensure reproducibility. As a positive control, the array included five clones from the third

chromosome chorion amplicon and three from the *X* amplicon. These genes flank the maximally amplified regions containing the chorion genes, and most displayed copy number increases of greater than two-fold in our assay (Table 1). Nine of the other clones that were significantly amplified (see Experimental Procedures) were striking because six were localized together within 65kb at cytological interval 30B10, and three were clustered within 10kb at 62D5 (Table 1) (Figure 1C, E). In addition to these eight clones, other genes within the 30B and 62D regions had increased copy numbers close to the significance cut off (Table 1). These observations strongly suggested that the microarray experiments had identified two additional follicle cell amplicons, and we named them DAFC (*Drosophila* Amplicon in Follicle Cells)-30B and 62D. For consistency we will refer to the chorion amplicons as DAFC-7F (*X* chromosome) and DAFC-66D (third chromosome).

Confirmation that DAFC-30B and DAFC-62D are Follicle Cell Amplicons

To validate the microarray data and verify that DAFC-30B and 62D are amplified in the follicle cells during late oogenesis, we used the approaches of quantitative real-time PCR and Fluorescent *in situ* Hybridization (FISH) with Bromodeoxyuridine (BrdU) colabeling of the follicle cells. We have previously employed quantitative real-time PCR to determine DNA copy number across the amplified domains DAFC-66D and 7F ((Claycomb et al., 2002) and data not shown). Genomic DNA was isolated from distinct populations of egg chambers: 1) stage 1-8 egg chambers, developmental stages prior to the onset of gene amplification in the follicle cells at stage 9; 2) stage 13 egg chambers at the peak of gene amplification; or 3) 16C amplifying follicle cell nuclei (as described

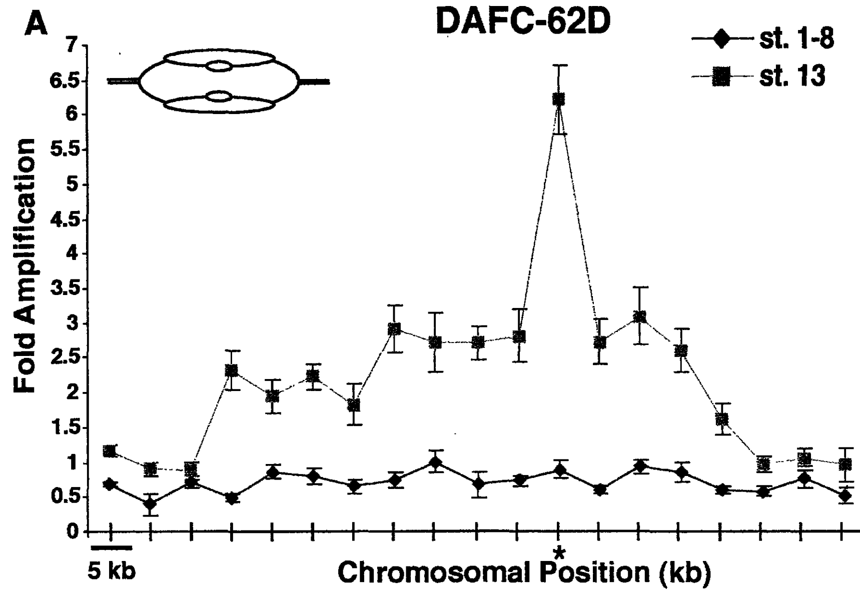
Table 1. Microarray experiments were repeated three times. Clones flanking the known chorion amplicons at DAFC-66D and 7F were positive controls. Threshold for significance is described in Experimental Procedures. ND-Not determined. Gene homologies or predicted functions are as listed in Flybase or determined by BLAST.

Table 1. Microarray experiments reveal clusters of genes with increased copy number.

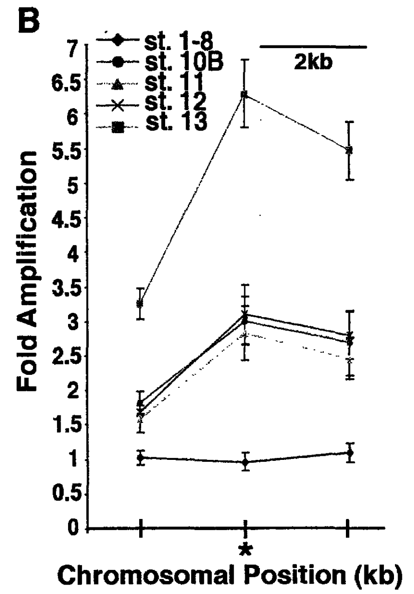
Gene (Clone ID)	Fold Amplified Trial 1	Fold Amplified Trial 2	Fold Amplified Trial 3	Gene Function/ Homology
DAFC-66D				
srpRβ (GM04779)	16.56	10.33	17.63	Signal Recognition Particle Receptor-β
prm (GH17893)	21.16	13.93	14.90	Paramyosin
prm (GH14085)	17.45	16.08	15.08	
CG32030 (LD24110)	2.64	4.28	2.50	Actin binding
CG32030 (SD08909)	ND	2.66	3.96	
DAFC-7F				
sptr (GH04031)	5.47	4.68	7.90	Sepiapterin Reductase
es2 (SD03464)	ND	2.93	6.33	Sepiapterin Reductase
CG12123 (GH02722)	1.34	ND	2.51	Novel
CG1440 (LD46760)	ND	2.13	2.47	Cysteine-type Endopeptidase
DAFC-62D				
CG1275 (LD36721)	1.59	1.40	1.73	Vesicle Electron Transporter
oxt (LD43716)	ND	ND	2.29	Glycosyl Transferase; Core-2/I-Branching Enzyme
CG5714 (GH14368)	3.11	1.84	1.25	Novel
CG32302 (LP11057)	1.73	2.54	1.98	Chitin Binding
DAFC-30B				
CG3811 (GH04717)	2.30	3.18	3.18	Transporter; Kazal-type Serine Protease Inhibitor
CG31883 (GH13755)	2.72	2.61	ND	Chitin Binding
Gdi (LD46767)	1.35	0.88	1.52	GDP-dissociation Inhibitor; Synaptic Vesicle Fusion
CG3838 (LD04047)	1.57	1.73	1.09	Novel
CG3838 (LD21447)	ND	ND	1.93	
CG4389 (GH12558)	1.54	1.85	1.56	Long Chain Enoyl-CoA Hydratase
CG18419 (GM07803)	2.18	2.12	2.36	Ca2+ Transporting ATPase
jp (GH28348)	ND	1.93	1.67	Junctophilin Matrix Protein

Figure 1

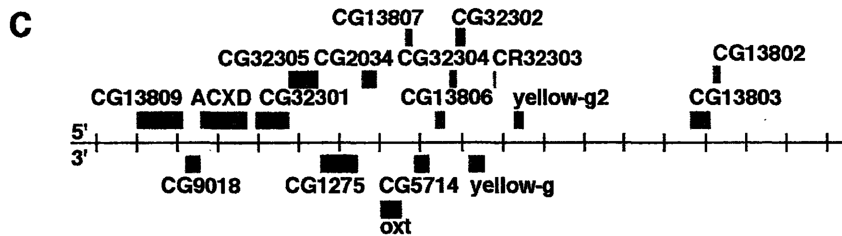
Real-time PCR demonstrates that the genes in DAFC-62D and 30B are amplified during late oogenesis. In (A) and (D), the fold amplification of each region, relative to a non-amplified portion of the genome and to 2C embryo genomic DNA standards, was determined in 5kb intervals by quantitative real-time PCR performed on preamplification- (st.1-8, blue diamonds) and amplification- (st. 13, pink boxes) stage whole egg chamber genomic DNA. Error bars are the standard deviations of triplicate reactions (Claycomb et al., 2002). (B) The peak of amplification at DAFC-62D, (A, asterisk), was analyzed by quantitative real-time PCR in 2kb intervals. In (C) and (E) the locations of the genes in 5kb intervals along the DAFC-62D and 30B amplicons, respectively, are diagrammed. Those genes represented on the microarray are shown in blue; others are in green. Tick marks in (A), (D) correspond to those in (C), (E), and the fold amplification for each gene can be examined by tracing upward to the graph. The insets in (A) and (D) show models for the gradients of amplification.



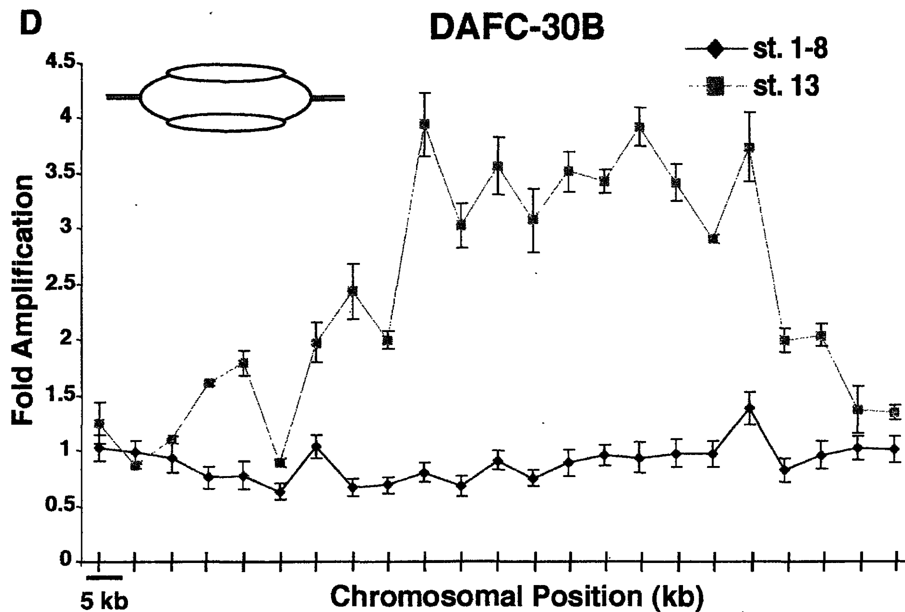
A
DAFC-62D
st. 1-8
st. 13
Fold Amplification
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
5 kb
*
Chromosomal Position (kb)



B
st. 1-8
st. 10B
st. 11
st. 12
st. 13
2kb
Fold Amplification
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
*
Chromosomal Position (kb)



C
CG13807
CG32302
CG32305
CG2034
CG32304
CR32303
CG13802
CG13809
ACXD
CG32301
CG13806
yellow-g2
CG13803
5'
3'
CG9018
CG1275
CG5714
yellow-g
oxt



D
DAFC-30B
st. 1-8
st. 13
Fold Amplification
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
5 kb
Chromosomal Position (kb)

E
CG31883
CG17855
CG13113
CG3838
CG13112
CG18419
CG3811
CG3818
CG13114
5'
3'
Gdl
jp
CG31709
CG4389

above). These genomic DNAs were used as templates for primers spaced at 5kb intervals along the putative amplification domain. We observed reproducibly that the 30B and 62D genomic intervals were amplified, with copy number increases extending across 75kb for DAFC-62D and 100kb for DAFC-30B (Figure 1A, D).

DAFC-62D is amplified a maximum of six-fold at an intergenic region, a lower level than the 14-fold amplification at DAFC-7F and 30-fold at DAFC-66D ((Claycomb et al., 2002) and Appendix 5). The peak of amplification in the DAFC-62D gradient suggests the position of an origin of DNA replication. To map this peak more precisely, we performed real-time PCR in 2kb intervals at the maximally amplified region. This confirmed the peak to be approximately 1.5kb from the 3' side of the *yellow-g2* gene (Figure 1B). We also investigated the developmental timing of replication initiation at the amplification peak by measuring copy number changes in stage 10B, 11, 12 and 13 egg chambers. We showed previously that DAFC-7F and 66D completed initiation by stage 11 and in subsequent stages existing replication forks elongated ((Claycomb et al., 2002) and Appendix 5). In contrast, DAFC-62D undergoes a late round of initiation between stages 12 and 13 (Figure 1B). It appears that the forks from this last initiation do not progress far, resulting in a small region of increased amplification in stage 13 egg chambers (model inset, Figure 1A).

The peak levels of amplification for DAFC-30B are 4-fold, and the maximum copy number increase is distributed over a 50kb region (Figure 1D). We examined the developmental timing of initiation at this amplicon but found that the initiation events were completed by stage 10B (data not shown). Thus the breadth of the amplification peak most likely results from elongation of these forks during stages 11-13 without

additional rounds of initiation (model inset, Figure 1D), an amplification profile similar to that of DAFC-66D and 7F.

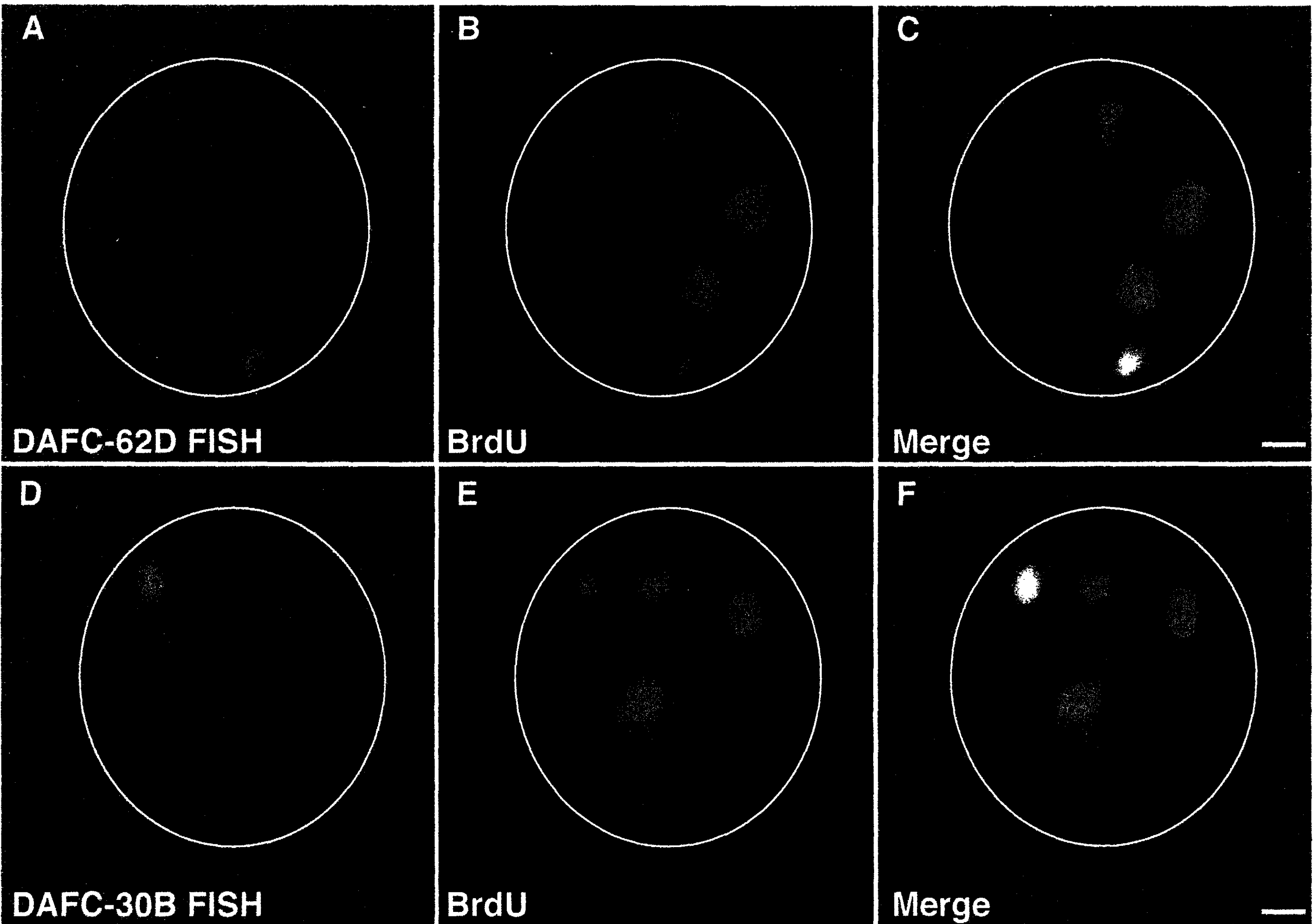
As a second approach to confirm that DAFC-30B and 62D were amplicons we directly observed replication patterns in follicle cells. The presence of two additional amplicons in follicle cells was suggested by the pattern of BrdU labeling during amplification stages (Calvi et al., 1998). At developmental times when genomic replication has ceased, follicle cells show BrdU incorporation in four foci. The larger two foci were shown by FISH to be the chorion gene clusters, but the identity of the smaller two foci remained unknown (Calvi et al., 1998). To test if DAFC-30B and DAFC-62D were replicated during amplification stages, we performed double labeling with FISH and BrdU. The FISH probe for each amplicon colocalized to one of the two small foci of BrdU incorporation in stage 10B follicle cells (Figure 2). The combined results of the microarray analysis, real-time PCR, and FISH-BrdU labeling experiments establish that DAFC-30B and 62D are two follicle cell amplicons.

Predicted Amplified Gene Products

Because our goal was to find additional examples of gene amplification that are necessary for proper development, we needed to determine the developmental relevance of the amplified genes in DAFC-30B and 62D. Our first step to evaluate the developmental significance of DAFC-30B and 62D amplification was to examine the homologies of the genes encoded in these regions. There are a variety of genes in DAFC-30B and 62D, none of which have been previously associated with mutant phenotypes or homologies that implicate them as functioning in oogenesis or eggshell

Figure 2

Fluorescent *in situ* Hybridization and BrdU labeling reveals that DAFC-30B and 62D correspond to sites of amplification in follicle cells during late oogenesis. A single representative follicle cell nucleus is shown, probed with a 10kb fragment from DAFC-62D (A, green), or with DAFC-30B (D, green), and labeled with BrdU (B, and D red). Colocalization is yellow (C) and (E). The larger BrdU spots in each nucleus correspond to DAFC-7F and 66D. The scale bar represents 1 μm.



A
DAFC-62D FISH
B
BrdU
C
Merge
D
DAFC-30B FISH
E
BrdU
F
Merge

formation. However, we found it notable that there were at least two groups of genes in the amplicons encoding proteins that could potentially function in egg production.

The maximally amplified genes in DAFC-62D, *yellow-g* and *yellow-g2*, are members of the *yellow* gene family that are predicted to encode secreted proteins (Drapeau, 2001; Maleszka and Kucharski, 2000). The family shares homology with the Major Royal Jelly Protein Family in honeybees (*Apis mellifera*), involved in the specification of the queen bee (Albert et al., 1999; Maleszka and Kucharski, 2000). The founding member of the Yellow family, Yellow-y, is known to play a role in mating behavior and in the melanization and hardening of the adult cuticle. Other Yellow family members have been shown to act as dopachrome-conversion enzymes that catalyze a key reaction in the melanization process (Han et al., 2002; Sugumaran, 2002). Interestingly, a similar process is used in the hardening of the egg chorion in mosquitos (Li, 1994) and suggests that Yellow-g and Yellow-g2 may play a catalytic role in the crosslinking of the chorion and/or underlying vitelline membrane proteins in *Drosophila.*

A second group of genes encodes proteins with chitin-binding motifs that could function in egg production. Genes of this type are present in both amplicons, with DAFC-62D containing two such genes and DAFC-30B containing one. Chitin-binding domains serve an antimicrobial function in a variety of plants and marine invertebrates. Homologs of marine invertebrate proteins, such as tachycitin, could provide the egg with protection against microbes (Kawabata et al., 1996). Alternatively, chitin, a structural polysaccharide found in many organisms, could also be a component of the eggshell, and interaction with the chitin binding proteins might contribute to eggshell integrity. In both DAFC-30B and 62D there are also a number of genes whose role in follicle cells is not

yet clear. These include both genes encoding proteins without known sequence motifs and genes whose products are predicted to have the enzymatic activities of adenylate cyclases, membrane transporters, calcium-transporting ATPases, GTP dissociation inhibitors, and others (Table 1).

Expression Patterns of the Amplified Genes

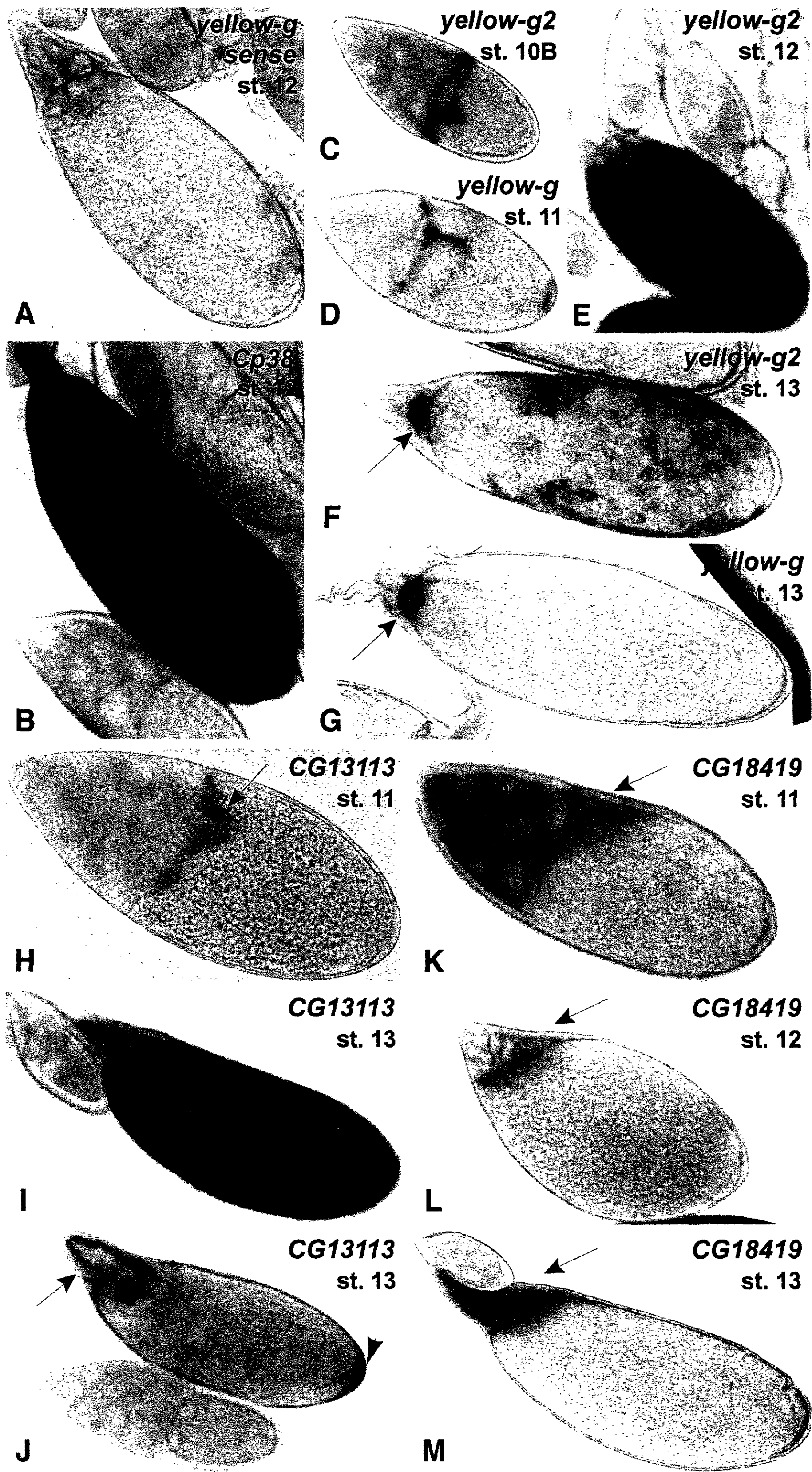
Gene homologies suggested it likely that at least some of the amplified genes would play a role in oogenesis or eggshell formation. Furthermore, we predicted that if amplification was required to achieve optimal levels of expression by these genes, they would be highly expressed in the follicle cells during late oogenesis, after amplification had initiated. To assess this hypothesis, we examined gene expression in the ovaries by RNA *in situ* hybridization.

The *yellow-g* and *yellow-g2* transcripts from DAFC-62D were initially detected in a subset of stage 10B follicle cells at the anterior end of the oocyte, concentrated at the dorsal side (Figure 3C, D). In egg chamber stages 11 and 12, all the follicle cells, except those around the nurse cells, robustly expressed the transcript (Figure 3E). In stage 13, expression was restricted to the follicle cells that produce the micropyle, a hollow tunnel in the eggshell through which the sperm enters (Figure 3F, G). The specificity of this final expression suggests a role for these gene products in vitelline membrane or eggshell formation.

Several other genes from the amplicons were expressed during follicle cell differentiation, when the vitelline membrane and eggshell are forming. The *CG13113* transcript from DAFC-30B, encoding a protein of unknown function, initiates expression

Figure 3

RNA *in situ* hybridization shows that genes in DAFC-62D and 30B are highly expressed in differentiating follicle cells. (A) The sense *yellow-g* probe shows no hybridization signal in any egg chamber stage. (B) The chorion gene *Cp38* is robustly expressed during egg chamber stages 11 and 12. *yellow-g2* (C) and *yellow-g* (D) initiate expression during late stage 10B, and transcripts accumulate over the egg chamber during stage 12 (E, *yellow-g2* shown). During stage 13, expression decreases (F, *yellow-g2*, slightly earlier stage 13 than in G, *yellow-g*), and the mRNA is concentrated in the follicle cells around the micropyle (arrows). *CG13113* expression begins during late stage 10B (H, arrow, stage 11 shown), and accumulates over the egg chamber in early stage 13 (I). (J) In later stage 13 egg chambers, *CG13113* transcripts are restricted to the follicle cells covering dorsal appendages (J, arrow), and at the posterior end (J, arrowhead). *CG18419* is expressed in the anterior dorsal follicle cells beginning in stage 10B (K, arrow, stage 11 shown), and into stage 12 (L, arrow). In stage 13, *CG18419* mRNA accumulates over the entire anterior of the egg chamber, including the dorsal appendages (M, arrow). Although it appears that *CG18419* is also expressed from the nurse cells in stage 11, we observed the same level of expression with the sense probe for this gene over the nurse cells in comparable stages of egg chambers (data not shown). Anterior is left.



yellow-g
sense
st. 12
A
yellow-g2
st. 10B
C
yellow-g
st. 11
D
yellow-g2
st. 12
E
Cp38
B
yellow-g2
st. 13
F
G
CG13113
st. 11
H
CG18419
st. 11
K
CG13113
st. 13
I
CG18419
st. 12
L
CG13113
st. 13
J
CG18419
st. 13
M

in dorsal follicle cells in late stage 10B, and is expressed across the follicle cell layer until stage 13. Late stage 13 egg chambers display high expression in a subset of follicle cells surrounding and building the dorsal appendages, structures that allow for gas exchange and respiration of the embryo, and in the posterior follicle cells (Figure 3H-J). The DAFC-30B transcript *CG18419*, with homology to a calcium transporting ATPase, is expressed throughout the follicle cell layer, but at highest levels in the dorsal follicle cells from stage 10B throughout later stages (Figure 3K-M). We also observed that, from stage 10 onward, the transcripts from the *CG3811*, *CG3818*, *CG13803*, and *CG5714* genes are present in the nurse cells and at low levels throughout the follicle cell layer (data not shown).

yellow-g is Essential for Proper Eggshell Formation

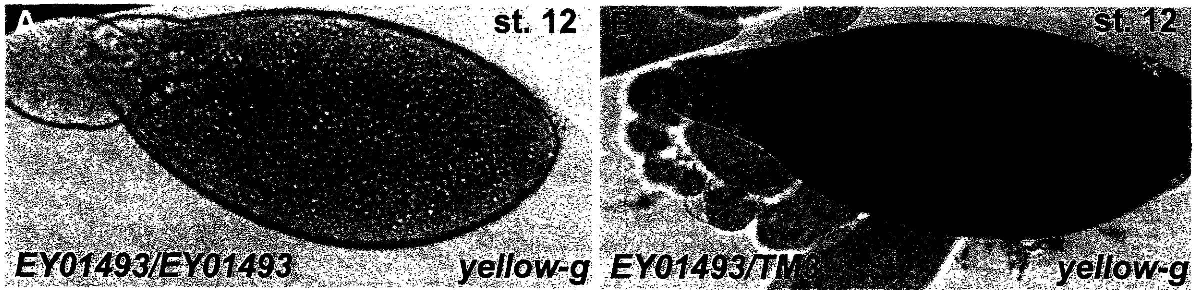
Although the homologies and expression patterns of the amplified genes are consistent with a role in eggshell or vitelline membrane formation or oogenesis in general, we sought to evaluate directly the necessity for amplified genes in these processes. In particular, the expression pattern of *yellow-g* and *yellow-g2* suggested that these genes would play a role in overall eggshell formation, or perhaps in the formation of eggshell substructures, such as the micropyle. To evaluate whether *yellow-g* was essential for follicle cell function we analyzed the phenotype caused by a P element transposon insertion that disrupts the *yellow-g* gene.

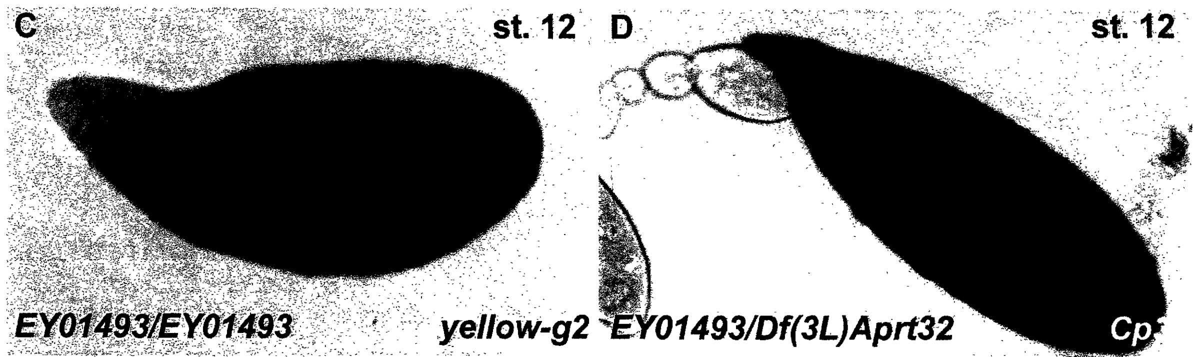
In the *EY01493* line there is a P element inserted in the 3' exon of *yellow*-g (Bellen, 2003; Spradling et al., 1999). This mutation disrupted *yellow-g* expression, whereas the expression of *yellow-g2* and *Cp-38* was unaffected (Figure 4A-D). These

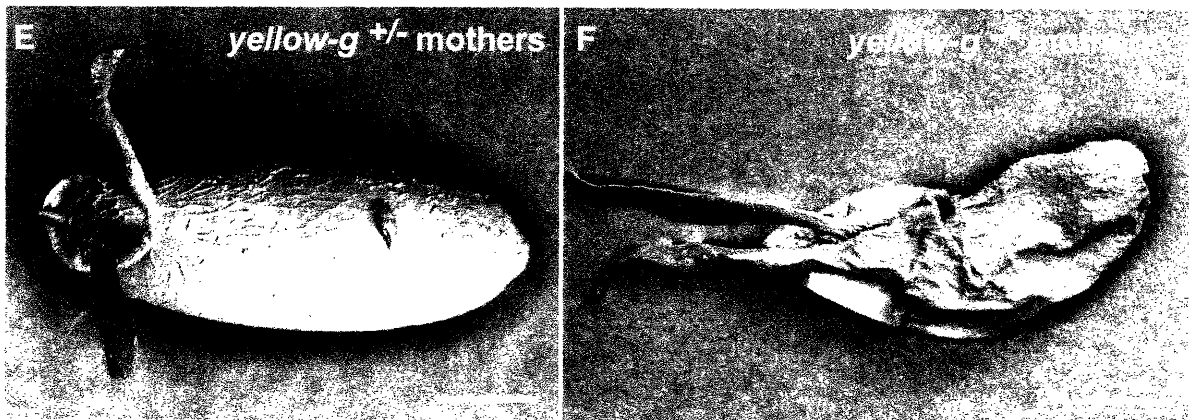
Figure 4

Characterization of the *yellow-g* mutant. RNA *in situ* hybridization shows that *yellow-g* mutant females (A, genotype: *EY01493/EY01493*) have decreased levels of the *yellow-g* transcript in comparison with control siblings (B, genotype: *EY01493/TM3*). However, RNA *in situ* hybridization to mutant ovaries for *yellow-g2* (C) or *Cp38* (D) shows that females homozygous for the P element insertion *EY01493* (C, genotype: *EY01493/EY01493*) or transheterozygous for the P element and the deficiency (D, genotype: *EY01493/Df(3L)Aprt32*) display no changes in transcript levels. Anterior is left.

SEM was performed on embryos from mothers with a wild-type copy of *yellow-g* (E, genotypes: *EY01493/TM6B* or *Df(3L)Aprt32/TM3*), or from *yellow-g* mutant mothers (F, genotype: *EY01493/Df(3L)Aprt32*). Embryos from the *yellow-g* mutant mothers appear to have normal chorion and dorsal appendages, but eggs spontaneously collapse when laid. Anterior is left and dorsal is up. Scale bars are 100 μm, both images are magnified 180x.


A
st. 12
EY01493/EY01493
yellow-g
st. 12
EY01493/TM3
yellow-g


C
st. 12
EY01493/EY01493
yellow-g2
D
st. 12
EY01493/Df(3L)Aprt32


E
yellow-g +/- mothers
F

mutant females were sterile, yet male fertility was not affected. Oogenesis proceeded normally in mutant females, but mature stage 14 oocytes often had indentations in the vitelline membrane, and at these sites the yolk was displaced (data not shown). Eggs laid by these mutant mothers were defective and collapsed, although the exochorion and dorsal appendages appeared normal (Figure 4E, F), indicative of defects in the vitelline membrane (Savant and Waring, 1989; Waring, 2000). These phenotypes show that *yellow-g* is needed for proper egg formation, possibly for the production of a structurally sound vitelline membrane, or to catalyze the crosslinking of eggshell layers for the rigidity of the egg.

Amplification is Necessary for Gene Expression

Amplification of the chorion genes is required for high levels of expression, as mutations that disrupt DNA replication factor genes such as *double parked* (*dup/cdt1*), *origin recognition complex subunit 2* (*orc2*), *chiffon* (*chif, dbf4-like*), *proliferating cell nuclear antigen* (*pcna, mus209*), or *minichromosome maintenance factor 6 (mcm6)*, result in decreased amplification and thin eggshells (Henderson et al., 2000; Landis et al., 1997; Landis and Tower, 1999; Schwed et al., 2002; Underwood et al., 1990). These mutants also display decreased BrdU incorporation at the four amplified loci (Calvi et al., 1998; Schwed et al., 2002; Whittaker et al., 2000). To determine if amplification of DAFC-30B and 62D was necessary for adequate levels of gene expression, we performed RNA *in situ* hybridization to the *yellow-g, yellow-g2,* and *CG13113* transcripts in the

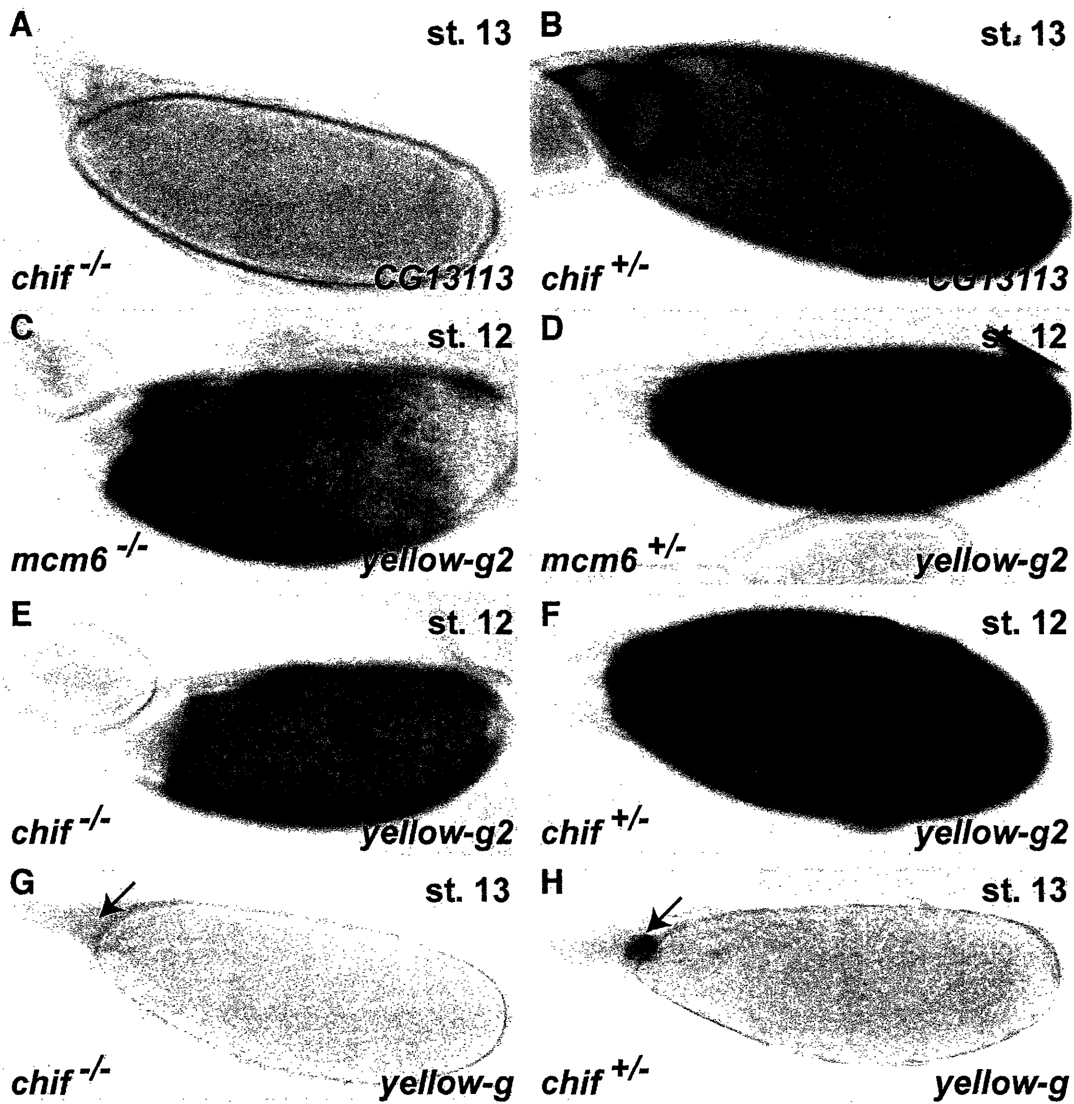
mcm6 and *chiffon* mutants. We found that transcript levels for all three amplified genes tested were reduced, but not eliminated, in the mutant ovaries (Figure 5A-H).

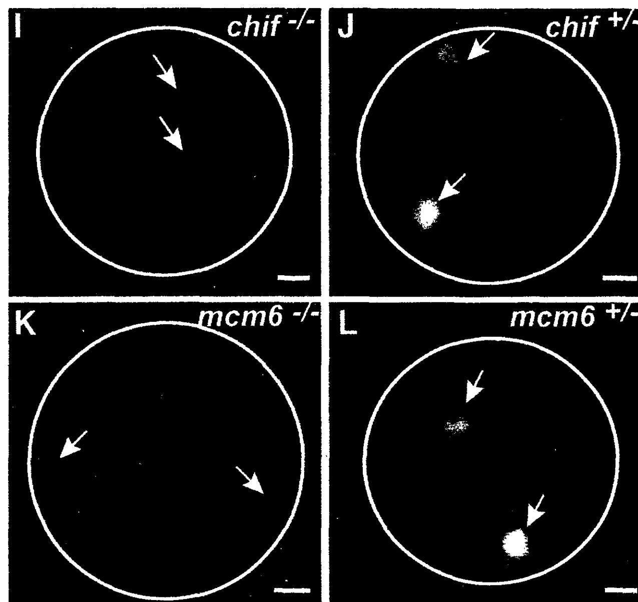
Although *mcm6* and *chiffon* mutants have been reported to have decreased BrdU incorporation at the amplifying loci (Calvi et al., 1998; Schwed et al., 2002), we wanted to test directly whether DAFC-30B and 62D specifically were amplified in the mutants. Thus, we performed FISH and BrdU colabeling on *mcm6* and *chiffon* mutant and sibling control ovaries. These experiments verified that DAFC-30B and 62D were not amplified to any significant degree, as could be detected by BrdU incorporation, in the majority of follicle cells (Figure 5I-L). These data demonstrate that amplification of DAFC-30B and 62D is necessary for high levels of expression and reiterate that these amplicons rely on the normal replication machinery for their amplification.

Figure 5

Amplified genes are poorly expressed in replication factor mutants that result in decreased amplification. RNA *in situ* hybridization for *CG13113* indicates that transcript levels are reduced in the *chiffon* female-sterile mutant (A, genotype: $chif^{QW16}/chif^{WD18}$) as compared to heterozygous siblings with one wild-type copy of *chiffon* (B, genotype: $chif^{QW16}/TM3$ or $chif^{WD18}/TM3$). RNA *in situ* hybridization for *yellow-g2* shows that transcript levels are reduced in the *mcm6* female-sterile mutant (C, genotype: $mcm6^{fs(1)k1214}/mcm6^{fs(1)k1214}$), compared to heterozygous siblings (D, genotype: $mcm6^{fs(1)k1214}/FM6$). *yellow-g2* (E) and *yellow-g* (G, arrow) transcript levels are reduced in the *chiffon* female-sterile mutant (genotype: $chif^{QW16}/chif^{WD18}$), compared to heterozygous siblings (F and H, genotype: $chif^{QW16}/TM3$ or $chif^{WD18}/TM3$). Anterior is left.

FISH to DAFC-30B and 62D concurrently (green, arrows) with BrdU colabeling (red) in *chiffon* mutant egg chambers shows that DAFC-30B and 62D are not amplified to a significant extent (I, genotype: $chif^{QW16}/chif^{WD18}$), as compared with heterozygous siblings (J, genotype: $chif^{QW16}/TM3$ or $chif^{WD18}/TM3$). Similar results were observed for the *mcm6* mutant females (K, genotype: $mcm6^{fs(1)k1214}/mcm6^{fs(1)k1214}$) and their sibling controls (L, genotype: $mcm6^{fs(1)k1214}/FM6$). A single follicle cell nucleus is shown for each. The scale bar represents 1 μm.


A
st. 13
chif -/-
CG13113
B
st. 13
chif +/-
CG13113
C
st. 12
mcm6 -/-
yellow-g2
D
st. 12
mcm6 +/-
yellow-g2
E
st. 12
chif -/-
yellow-g2
F
st. 12
chif +/-
yellow-g2
G
st. 13
chif -/-
yellow-g
H
st. 13
chif +/-
yellow-g


I
chif -/-
J
chif +/-
K
mcm6 -/-
L
mcm6 +/-

DISCUSSION

There are several mechanisms by which organisms can fulfill a need for bursts of gene expression, including carrying stable duplications of the highly expressed genes, up-regulating transcription, up-regulating translation, or by developmentally-regulated amplification of specific genes. To date, only a handful of developmental amplicons have been examined, and the isolation of new amplicons has mainly relied on visual detection of amplified DNA. Thus, it remains to be seen how widely gene amplification is used as a developmental strategy for robust gene expression across different species.

We have established a methodology for the systematic analysis of gene amplification as a developmental strategy, and in doing so we have identified two additional developmentally-regulated amplicons in the *Drosophila* follicle cells. The recovery of these amplicons validates the microarray approach to survey DNA copy number and provides additional model replicons to study. Additionally, the power of *Drosophila* genetics affords us a system to evaluate the functions of amplified genes in particular developmental processes. The process of amplification in the follicle cells reveals a progressive restriction of increased gene copy number in the genome. Initially the entire euchromatin is increased in copy number as the follicle cells become polyploid. Later in follicle cell differentiation only four specific genomic regions are amplified.

None of the amplified genes we identified in DAFC-30B and 62D had been previously implicated in eggshell formation, and thus recovery of additional amplicons also highlights developmental activities of the amplified genes. We showed that the *yellow-g* gene is essential for a rigid eggshell, and the predicted gene products of the

yellow-g and *yellow-g2* genes suggest a molecular explanation for these mutant defects. The eggshell is composed of several layers, including the outermost exochorion, the endochorion, the inner chorion layer, and the vitelline membrane, which is the innermost structure that also contacts the oocyte (for reviews see Spradling and Waring (Spradling, 1993; Waring, 2000).) The collapsed embryos and disrupted vitelline membranes that result from mutation of *yellow-g* indicate that *yellow-g* is necessary for the structural integrity of the eggshell. At the level of the light microscope, the exochorion of embryos laid by mutant mothers appears normal. The collapsed embryos are reminiscent of vitelline membrane defects (Savant and Waring, 1989) leading us to hypothesize that *yellow-g* is necessary for proper vitelline membrane formation.

We propose that Yellow-g and Yellow-g2 act to crosslink the vitelline membrane, or perhaps the inner chorion layer. The Yellow family members, Yellow-f and Yellow-f2, are capable of catalyzing the conversion of dopachrome to dihydroxyindole, a limiting step in the melanization pathway, during larval, pupal and adult stages (Han et al., 2002). The enzymatic events leading to the crosslinking of the vitelline membrane are not well understood, but seem to involve one phase of disulfide bond formation and a subsequent disulfide bond-independent phase (Waring, 2000). Additionally, the *α methyl dopa resistant* (*amd*) gene product, which acts in the conversion of dopamine during the polymerization of the adult cuticle, is required in the follicle cells for proper vitelline membrane crosslinking (Konrad et al., 1993). This suggests that a similar set of dopamine conversion reactions catalyzed by Yellow-g and Yellow-g2 may be necessary for the crosslinking of the vitelline membrane just prior to egg laying. Consistent with this hypothesis, we observed that eggs laid by homozygous *yellow-g* mutant females are

highly sensitive to sodium hypochlorite (bleach), and the majority of these embryos burst upon brief exposure (See Appendix 4). Of the remaining, intact embryos, 100% were permeable to the dye neutral red (See Appendix 4, performed as described (LeMosy and Hashimoto, 2000)), which has been used to assay vitelline membrane defects (Degelmann et al., 1990; Komitopoulou et al., 1983; Konrad et al., 1993). These results are indicative of a failure to crosslink the vitelline membrane and further implicate *yellow-g* in the crosslinking process. However, this hypothesis does not explain the specific expression of the *yellow-g* and *yellow-g2* genes in the follicle cells producing the micropyle late in egg chamber development. It is possible that crosslinking of the vitelline membrane or inner chorion layer within this specialized structure requires distinct regulation or timing. A more detailed analysis of the eggshell defect and biochemical studies of Yellow-g and Yellow-g2 will help us to better understand the steps necessary for vitelline membrane crosslinking and will uncover any specialized micropyle functions.

DAFC-30B and DAFC-62D provide insights into the use of amplification as a developmental strategy. All of the previously characterized amplified genes play a purely structural role in eggshell formation; no enzymes necessary for proper eggshell formation have been examined. None of the genes of DAFC-30B and DAFC-62D encode known structural components of the eggshell. However, several of the amplified genes that are highly expressed in follicle cells, including *CG18419* and the *yellow-g* genes, encode products predicted to possess enzymatic, signal transduction, or transporting activities. Furthermore, at least *yellow-g* is essential for proper egg formation, thus revealing an additional function of amplification: to increase the levels of

enzymes needed to catalyze developmentally important reactions. Thus, the identification of additional amplicons highlights genes likely to be crucial in developmental events and opens the possibility that other tissues employ amplification to maximize gene expression during differentiation. It is surprising that a four to six-fold increase in gene copy number would affect gene product levels in a developmentally significant manner. It is possible, however, that copy number increases are considerably higher in subsets of follicle cells, or that the replication process itself facilitates transcription.

The follicle cell amplicons serve as superb model metazoan replicons, permitting delineation of *cis*-regulatory elements, identification of replication proteins, and clarifying the developmental control of the initiation and elongation. Developmental distinctions between DAFC-62D and the previously studied DAFCs provide clues into how origin firing can be linked to developmental signals. Previously, we showed by real-time PCR that replication initiates at DAFC-66D and 7F, coupled with replication fork movement, during egg chamber stages 10B and 11. Subsequently (stages 12 and 13), origins cease firing and only existing replication forks move bidirectionally to produce a gradient of copy number that extends over 100kb (Claycomb et al., 2002). Furthermore, the replication initiation factor ORC2 only localizes to amplification origins during the initiation phase and dissociates at the onset of the elongation phase. Replication factors involved in multiple steps of DNA replication, such as MCM2-7 and PCNA, co-localize with BrdU throughout amplification (Claycomb et al., 2002; Royzman et al., 1999; Spradling, 1981; Whittaker et al., 2000).

DAFC-62D behaves differently from these amplicons and from DAFC-30B. There is a final increase in copy number at a very precise region of the amplicon, about 1.5kb downstream of *yellow-g2*, during stage 13. As it is the peak of amplification, this region is likely to possess a replication origin. Understanding how DAFC-62D can undergo a final initiation hours after ORC is no longer detectable at origins by immunofluorescence will provide insights into the control of replication initiation. The additional replication in stage 13 may occur in only subsets of follicle cells and ORC could persist specifically at DAFC-62D in these cells. For example, additional gene copies could permit optimal levels of expression of the *yellow-g* genes in the follicle cells building the micropyle.

We initiated these studies to devise a systematic approach for finding developmental amplicons. We have demonstrated that the microarray assay is sensitive and can detect low levels of gene amplification, and we have shown that amplification levels as low as four-fold can be developmentally important. Thus, we believe our approach will be invaluable in surveying for gene amplification in a number of tissues and in a variety of organisms where amplification has not been detected. Not only has the microarray strategy identified additional amplicons, but when coupled with the power of a genetic organism, it has proven to be a functional genomics approach for highlighting genes involved in specific developmental pathways.

MATERIALS AND METHODS

Quantitation of DNA Copy Number on Microarrays

Drosophila Gene Collection strains were grown in deep-well 96 well plates with 1 ml media. Plasmid minipreps were done using Millipore MultiScreen and yields quantitated using a Tecan GENios microplate fluorometer and Picogreen (Molecular Probes). The cDNA inserts were PCR-amplified, and primer sets used are available upon request. PCR products were isopropanol precipitated, analyzed to be the predicted size and to have an average concentration of 370 μg/ml. Microarrays were printed using a Cartesian Technologies arrayer on Corning (experiment 1) or Ultragap slides (experiments 2 and 3) and crosslinked with a 2400UV Stratalinker at 300mJ.

Genomic DNA was isolated from embryos or FACS sorted 16C follicle cells from a *y; cn bw sp* stock as described previously (Lilly and Spradling, 1996). 200 ng of each was digested with Rsa I and labeled by random priming with either Cy3 or Cy5-dCTP (Pollack et al., 1999) and hybridized to the slides in 3.4X SSC, 0.3% SDS with 30μg human Cot 1 DNA and 100μg yeast tRNA in 50μl. The hybridization was at 55°C overnight, and slides were washed in 0.1X SSC at room temperature.

Fluorescent hybridization was detected on an Axon Instruments GenePix 4000A microarray scanner, with manual adjustment of the scan area for each feature. Spots having an intensity less than 100 were discarded. The raw ratios (16C follicle cell /embryo) were calculated using the background substracted median intensities of the remaining features. The ratios were normalized by dividing each raw ratio by the mean raw ratio. Clones with a ratio higher than two standard deviations from the mean were scored as significantly amplified.

In the first experiment 3,643 of the clones were scored. The significance cut off was an amplification level of 2.1 or higher, and 13 clones were significantly amplified. In the second experiment 5,568 clones were scored. 63 had ratios higher than the cut off amplification value of 1.8 or higher. In the third experiment 5,929 clones were scored. These were done in duplicate on the same slides; Table 1 shows the average values. 28 clones showed amplification values of 1.8 or higher, the significance cut off in this experiment.

Quantitative Real-Time PCR

Quantitative real-time PCR was performed as described (Claycomb et al., 2002) except that: ABI SYBR Green Master Mix was used (Applied Biosystems), and primers in 5kb intervals across 30B10 and 62D5, and in 2kb intervals across the 62D5 amplification peak were supplied by IDT and Genelink. Primer sequences are available upon request. All experimental PCR reactions at the amplicons were compared to nonamplified control loci on the same chromosome arm to calculate fold amplification. For DAFC-30B, the nonamplified control locus was located at 30B2 and for DAFC-62D, the nonamplified control locus was located at 62C5. The relative fluorescence of control loci was initially compared to values obtained with the nonamplified control *ry* primer set (Claycomb et al., 2002), to assure that they were valid nonamplified loci.

Fluorescent *in situ* Hybridization, BrdU Labeling, and Confocal Microscopy

Ovaries were dissected, labeled with 4 μg/ml BrdU and prepared as described (Calvi et al., 1998; Royzman et al., 1999). BrdU was detected with donkey anti-mouse

Rhodamine-RedX (Jackson Immunoresearch) at 1:250. After BrdU detection, ovaries were re-fixed (Royzman et al., 1999), and FISH was performed as described (Calvi et al., 1998). The probes used were generated from 10kb PCR products (Clontech Advantage 2 PCR Kit, BD Biosciences). The DAFC-30B probe spans genes *CG18419* to *CG31883*, and the DAFC-62D probe covers from *CG5714* to the intergenic region between the *yellow-g* genes. Primers are available upon request.

Templates for the PCR were BACR07D23 for DAFC-30B and BACR22J16 for DAFC-62D (BAC PAC/CHORI). FISH probes were detected with goat anti-DIG FITC at 1:200 (Enzo), and samples were mounted in Vectashield (Vector Labs) or Slowfade (Molecular Probes). All images were collected on a Zeiss Axiovert 100M Meta confocal microscope with LSM51 Software using a 100x Plan Aprochromat objective and the filters set according to the manufacturer's parameters.

RNA *in situ* Hybridization

Templates for *in situ* probes were generated by PCR of each gene's largest predicted exon from *Oregon-R* genomic DNA. Primers used were 30-35mers and added a 5' EcoRI restriction site or a 3' XhoI site to the PCR product. Primers were supplied by Genelink. Sequences are available upon request. PCR products were purified with the Qiagen PCR Clean-up kit, digested with EcoRI and XhoI (NEB), then cloned into pBluescript SK+ using T4 DNA ligase (NEB). Sense and antisense probes were made as described (Royzman et al., 1997), and hybridizations were done at 55° C on ovaries as described (Royzman et al., 2002). The images in (Figure 3C, D, H, K) were captured using the 25X Zeiss Neofluar objective, water immersion. A Plan Neofluar 20X

objective was used for all others. A Zeiss Axiophot microscope with a SPOT RT CCD camera and software was used to capture all images.

yellow-g mutant analysis and Scanning Electron Microscopy

The line *EY01493* contains an EPgy2 P element in the 3' exon of *yellow-g*. The line was generated by the P-element Screen /Gene Disruption Project of the Bellen/Rubin/Spradling labs (Bellen, 2003; Spradling et al., 1999) and obtained from the Bloomington Stock Center (#15512). The deficiency, *Df(3L)Aprt32/TM6 Ubx e*, removes the 62B1-62E3 region and was obtained from the Bloomington Stock Center (#5411) (Wang et al., 1994).

The EPgy2 line was crossed to the deficiency line and the progeny were collected for egg laying experiments. Heterozygous sibling females were separated from mutant *EPgy2/Df* females, and egg laying was monitored over 6 to 12 hour intervals. Fertility was determined by allowing the females to lay eggs for 3 days and monitoring for larvae. SEM was performed on a Jeol JSM5600LV SEM in low vacuum mode with an acceleration voltage of 5kV and a spot size of 42. Images were collected using the shadow mode of the backscatter detector. Samples were prepared by adhering 0-12 hour embryos on double-stick carbon tabs.

ACKNOWLEDGMENTS

We thank Hugo Bellen for providing unpublished P element insertion lines, Ethan Lee and Marc Kirschner for the gift of Taq/Pfu polymerase, Stephen Bell for use of a real-time PCR machine, Tom Volkert, Dave MacAlpine, and Alex Ensminger for helpful advice, Nicki Watson for help with SEM, Raissa Fomina for help with stocks, Helena Kashevsky for assistance with stocks and the third microarray experiment, Ryan Casillo for performing the third microarray experiment, and Gerry Rubin and BDGP for the *Drosophila* Gene Collection. Experiments were done in the Corning Microarray Facility and the Keck Imaging Center at the WI. Dave Bartel, Steve Bell, Hannah Blitzblau, Stacie Bumgarner, Gerry Fink, Susan Lindquist, and Ilaria Rebay provided helpful comments on the manuscript. M.B. and D.F. were supported by funds from Corning, and G.B. by fellowships from the Damon-Runyon Cancer Fund and Merck. G.B. was a Margaret and Herman Sokol fellow of the Whitehead Institute. J.C. is a David Koch Graduate Fellow. This research was funded by the Stewart Trust, the Mathers Foundation, and NIH grant GM57451 to T.O-W.

REFERENCES

Albert, S., Bhattacharya, D., Klaudiny, J., Schmitzova, J., and Simuth, J. (1999). The family of major royal jelly proteins and its evolution. J. Mol. Evol. *49*, 290-297.

Austin, R. J., Orr-Weaver, T. L., and Bell, S. P. (1999). *Drosophila* ORC specifically binds to *ACE3*, an origin of DNA replication control element. Genes and Dev. *13*, 2639-2649.

Bellen, H., http://flypush.imgen.bcm.tmc.edu/pscreen/. (2003).

Bielinsky, A. K., Blitzblau, H., Beall, E. L., Ezrokhi, M., Smith, H. S., Botchan, M. R., and Gerbi, S. A. (2001). Origin recognition complex binding to a metazoan replication origin. Curr. Biol. *11*, 1427-1431.

Bosco, G., and Orr-Weaver, T. L. (2002). Regulation of the cell cycle during oogenesis and early embryogenesis in *Drosophila*. In Regulation of gene expression at the beginning of animal development, M. DePamphilis, ed. (Amsterdam: Elsevier), pp. 107-154.

Brown, D. D., and Dawid, I. B. (1968). Specific gene amplification in oocytes. Science *160*, 272-280.

Calvi, B. R., Lilly, M. A., and Spradling, A. C. (1998). Cell cycle control of chorion gene amplification. Genes and Dev. *12*, 734-44.

Carminati, J. L., Johnston, C. J., and Orr-Weaver, T. L. (1992). The *Drosophila ACE3* chorion element autonomously induces amplification. Mol. Cell. Biol. *12*, 2444-2453.

Claycomb, J. M., MacAlpine, D. M., Evans, J. G., Bell, S. P., and Orr-Weaver, T. L. (2002). Visualization of replication initiation and elongation in *Drosophila*. J. Cell Biol. *159*, 225-236.

de Cicco, D., and Spradling, A. (1984). Localization of a cis-acting element responsible for the developmentally regulated amplification of *Drosophila* chorion genes. Cell *38*, 45-54.

Degelmann, A., Hardy, P. A., and Mahowald, A. P. (1990). Genetic analysis of two female-sterile loci affecting eggshell integrity and embryonic pattern formation in *Drosophila melanogaster*. Genetics *126*, 427-34.

Delidakis, C., and Kafatos, F. C. (1989). Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*. EMBO J. *8*, 891-901.

Drapeau, M. D. (2001). The family of Yellow-related *Drosophila melanogaster* proteins. Biochem. Biophys. Res. Commun. *281*, 611-613.

Gall, J. G. (1968). Differential synthesis of the genes for ribosomal RNA during amphibian oogenesis. Proc. Natl. Acad. Sci. USA *60*, 553-560.

Glover, D. M., Zaha, A., Stocker, A. J., Santelli, R. V., Pueyo, M. T., De Toledo, S. M., and Lara, F. J. S. (1982). Gene amplification in *Rhynchosciara* salivary gland chromosomes. Proc. Natl. Acad. Sci. USA *79*, 2947-2951.

Gray, J. W., and Collins, C. (2000). Genome changes and gene expression in human solid tumors. Carcinogenesis *21*, 443-52.

Han, Q., Fang, J., Ding, H., Johnson, J. K., Christensen, B. M., and Li, J. (2002). Identification of *Drosophila melanogaster* yellow-f and yellow-f2 proteins as dopachrome-conversion enzymes. Biochem. J. *368*, 333-340.

Heck, M., and Spradling, A. (1990). Multiple replication origins are used during *Drosophila* chorion gene amplification. J. Cell Biol. *110*, 903-914.

Henderson, D. S., Wiegand, U. K., Norman, D. G., and Glover, D. M. (2000). Mutual correction of faulty PCNA subunits in temperature-sensitive lethal *mus209* mutants of *Drosophila melanogaster*. Genetics *154*, 1721-1733.

Higgins, M. J., Walker, V. K., Holden, J. A., and White, B. N. (1984). Isolation of two *Drosophila melanogaster* genes abundantly expressed in the ovary during vitelline membrane synthesis. Dev. Biol. *105*, 155-65.

Kawabata, S., Nagayama, R., Hirata, M., Shigenaga, T., Agarwala, K. L., Saito, T., Cho, J., Nakajima, H., Takagi, T., and Iwanaga, S. (1996). Tachycitin, a small granular component in horseshoe crab hemocytes, is an antimicrobial protein with chitin-binding activity. J. Biochem. (Tokyo) *120*, 1253-1260.

Komitopoulou, K., Gans, M., Margaritis, L. H., Kafatos, F. C., and Masson, M. (1983). Isolation and characterization of sex-linked female-sterile mutants in *Drosophila melanogaster* with special attention to eggshell mutants. Genetics *105*, 897-920.

Konrad, K. D., Wang, D., and Marsh, J. L. (1993). Vitelline membrane biogenesis in *Drosophila* requires the activity of the alpha-methyl dopa hypersensitive gene (I(2)amd) in both the germline and follicle cells. Insect Mol. Biol. *1*, 179-87.

Landis, G., Kelley, R., Spradling, A. C., and Tower, J. (1997). The *k43* gene, required for chorion gene amplification and diploid cell chromosome replication, encodes the *Drosophila* homolog of yeast origin recognition complex subunit 2. Proc. Natl. Acad. Sci. USA *94*, 3888-3892.

Landis, G., and Tower, J. (1999). The *Drosophila chiffon* gene is required for chorion gene amplification, and is related to the yeast dbf4 regulator of DNA replication and cell cycle. Development *126*, 4281-4293.

LeMosy, E. K., and Hashimoto, C. (2000). The nudel protease of *Drosophila* is required for eggshell biogenesis in addition to embryonic patterning. Dev. Biol. *217*, 352-361.

Li, J. (1994). Egg chorion tanning in *Aedes aegypti* mosquito. Comp. Biochem. Physiol. A Physiol. *109*, 835-843.

Lilly, M. A., and Spradling, A. C. (1996). The *Drosophila* endocycle is controlled by Cyclin E and lacks a checkpoint ensuring S-phase completion. Genes and Dev. *10*, 2514-2526.

Lu, L., Hongjun, Z., and Tower, J. (2001). Functionally distinct, sequence specific replicator and origin elements are required for *Drosophila* chorion gene amplification. Genes and Dev. *15*, 134-146.

Mahowald, A., and Kambysellis, M. (1980). Oogenesis. In The Genetics and Biology of *Drosophila*, M. Ashburner and T. Wright, eds. (New York: Academic Press), pp. 141-224.

Maleszka, R., and Kucharski, R. (2000). Analysis of *Drosophila* yellow-B cDNA reveals a new family of proteins related to the royal jelly proteins in the honeybee and to an orphan protein in an unusual bacterium *Deinococcus radiodurans*. Biochem. Biophys. Res. Commun. *270*, 773-776.

Orr, W., Komitopoulou, K., and Kafatos, F. C. (1984). Mutants supressing in *trans* chorion gene amplification in *Drosophila*. Proc. Natl. Acad. Sci. USA *81*, 3773-3777.

Orr-Weaver, T. L. (1991). *Drosophila* chorion genes: cracking the eggshell's secrets. BioEssays *13*, 97-105.

Parks, S., and Spradling, A. (1987). Spatially regulated expression of chorion genes during *Drosophila* oogenesis. Genes and Dev. *1*, 497-509.

Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat. Genet. *23*, 41-46.

Popodi, E., Minoo, P., Burke, T., and Waring, G. L. (1988). Organization and expression of a second chromosome follicle cell gene cluster in *Drosophila*. Dev. Biol. *127*, 248-56.

Prescott, D. M. (1994). The DNA of ciliated protozoa. Microbiol. Rev. *58*, 233-267.

Royzman, I., Austin, R. J., Bosco, G., Bell, S. P., and Orr-Weaver, T. L. (1999). ORC localization in *Drosophila* follicle cells and the effects of mutations in *dE2F* and *dDP*. Genes and Dev. *13*, 827-840.

Royzman, I., Hayashi-Hagihara, A., Dej, K. J., Bosco, G., Lee, J. Y., and Orr-Weaver, T. L. (2002). The E2F cell cycle regulator is required for *Drosophila* nurse cell DNA replication and apoptosis. Mech. Dev. *119*, 225-37.

Royzman, I., Whittaker, A. J., and Orr-Weaver, T. L. (1997). Mutations in *Drosophila DP* and *E2F* distinguish G1-S progression from an associated transcriptional program. Genes and Dev. *11*, 1999-2011.

Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M., and Harvey, D. A. (2000). A *Drosophila* complementary DNA resource. Science *287*, 2222-4.

Rudkin, G., and Corlette, S. L. (1957). Disproportionate synthesis of DNA in a polytene chromosome region. Proc. Natl. Acad. Sci. USA *43*, 964-968.

Savant, S. S., and Waring, G. L. (1989). Molecular analysis and rescue of a vitelline membrane mutant in *Drosophila*. Dev. Biol. *135*, 43-52.

Schwed, G., May, N., Pechersky, Y., and Calvi, B. R. (2002). *Drosophila* minichromosome maintenance 6 is required for chorion gene amplification and genomic replication. Mol. Biol. Cell *13*, 607-620.

Spradling, A., and Orr-Weaver, T. (1987). Regulation of DNA replication during *Drosophila* development. Ann. Rev. Genetics *21*, 373-403.

Spradling, A. C. (1993). Developmental genetics of oogenesis. In The development of *Drosophila melanogaster*, M. Bate and A. Martinez Arias, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 1-70.

Spradling, A. C. (1981). The organization and amplification of two clusters of *Drosophila* chorion genes. Cell *27*, 193-201.

Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., Misra, S., and Rubin, G. M. (1999). The Berkeley *Drosophila* Genome Project gene disruption project: Single P-element insertions mutating 25% of vital *Drosophila* genes. Genetics *153*, 135-177.

Sugumaran, M. (2002). Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. Pigment Cell Res. *15*, 2-9.

Underwood, E. M., Briot, A. S., Doll, K. Z., Ludwiczak, R. L., Otteson, D. C., Tower, J., Vessey, K. B., and Yu, K. (1990). Genetics of 51D-52A, a region containing several

maternal-effect genes and two maternal-specific transcripts in *Drosophila*. Genetics *126*, 639-650.

Wang, M., Champion, L. E., Biessmann, H., and Mason, J. M. (1994). Mapping a mutator, *mu2*, which increases the frequency of terminal deletions in *Drosophila melanogaster*. Mol. Gen. Genet. *245*, 598-607.

Waring, G. L. (2000). Morphogenesis of the eggshell in *Drosophila.* Int. Rev. Cytol. *198*, 67-108.

Whittaker, A. J., Royzman, I., and Orr-Weaver, T. L. (2000). *Drosophila* double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. Genes and Dev. *14*, 1765-1776.

Wu, N., Liang, C., DiBartolomeis, S. M., Smith, H. S., and Gerbi, S. A. (1993). Developmental progression of DNA puffs in *Sciara coprophila*: amplification and transcription. Dev. Biol. *160*, 73-84.

Conclusions/Afterword

CONCLUSIONS

The work presented in this thesis advances the field of *Drosophila* gene amplification in several ways. First, the studies have established that amplification is a more powerful model for the study of metazoan DNA replication than we may have previously appreciated. We have discovered by microscopy and real-time PCR that the initiation and elongation portions of replication occur during distinct phases of oogenesis for at least DAFC-66D and 7F, and likely for DAFC-30B, thus making the amplified regions an attractive model for the *in vivo* study of replication elongation. The newly isolated amplicon DAFC-62D displays an increase in copy number at a specific genomic position in stage 13 egg chambers, indicating that the origin fires at a time when ORC2 has dissociated and only elongation is occurring at other DAFCs. We should note, however, that although ORC2 may have dissociated from DAFC-62D by the stage 13 origin firing, the other components of the pre-replication complex, such as Dup/Cdt1 and MCM2-7, may have already been loaded and reside at the origin waiting for the proper signal to fire. It will be important to determine the purpose and regulation of this late firing in the future.

Second, this work has elaborated on the roles of proteins involved in gene amplification. The localization patterns of replication proteins previously not observed at amplification loci, PCNA and MCM2-7, indicate that they play a role in amplification, both as origins fire and at replication forks. The precise localization of Dup/Cdt1 and ORC2 in relation to each other and BrdU incorporation has been deteremined here and indicates that these proteins may play distinct roles in amplification. We have shown that

the replication initiation factor Dup/Cdt1 is present and hypothesize that it may play a role at elongating replication forks. By mutant analysis, we have begun to understand the functions of Dup/Cdt1 in gene amplification, in that the protein is necessary for the proper loading of MCM2-7 onto amplification origins and perhaps for the proper import/export of MCM2-7 to/from the nucleus during amplification. Preliminary evidence suggests that CDC6 is present, albeit transiently, at amplification foci, but thus far we cannot conclusively assign a role for CDC6 in gene amplification (Appendix 3).

Third, and perhaps the most important contribution of this work to the field of gene amplification is the proof of principle that the microarray assay is capable of detecting developmentally important gene amplification events, even when they occur at a low level. Real-time PCR, FISH, RNA *in situ* hybridization, and mutant studies indicate that genes in DAFC-30B and 62D are amplified and highly expressed after amplification, that amplification is necessary to achieve proper levels of expression, and that at least *yellow-g* is necessary for proper egg formation. The necessity of *yellow-g* in egg formation and the presence of mostly enzymes in these new amplicons implicates enzymes as targets of gene amplification, thus reinforcing the hypothesis that gene amplification is a more widespread phenomenon than we appreciate currently. Further mutant and expression analysis of the genes in DAFC-30B and 62D will elaborate on the reasons for amplifying these genes and provide insights into egg development.

Although this work has provided new insights into the mechanism and developmental implications for gene amplification, many questions still remain. The future directions for this work will be discussed in detail below.

Rescue of *dup* mutants

The mystery of which alleles solely affect *dup* remains unsolved (Appendix 2), so the rescue experiments could be attempted again, after making a rescue construct containing the *dup* genomic region and generating transgenic fly lines. This may be an unfruitful endeavor, as the regulatory region is not characterized, and may be up to tens of thousands of basepairs long, making cloning difficult. The $dup^{a2, a4, L5}$ alleles have not been molecularly characterized yet, so sequencing those alleles and finding the mutation could be useful for understanding why certain combinations of alleles are not lethal or sterile. Expression of Dup alone may not be enough to rescue replication defects, so co-expressing CDC6, another protein necessary for forming the pre-replication complex and loading MCM2-7, with Dup may be necessary to observe rescue. Furthermore, we could attempt to rescue the questionable alleles with expression of the kinase gene, to approach the question from a different perspective.

The Role of Dup in Elongation

What is the role of Dup in elongation? The most definitive way to answer this question would be to isolate a conditional allele of *dup* that could be inactivated at specific points during gene amplification. Unfortunately, such alleles are quite difficult to isolate in *Drosophila*, so we are currently unable to address this question in this manner. Perhaps searching for additional alleles of *dup* that disrupt different regions of the protein and result in a separation of function (an initiation domain vs. an elongation domain of the Dup protein, for instance) in various mutant collections, could be useful,

but will likely not be very fruitful. Another way to attempt to dissect what Dup could be doing during elongation could be to create *dup* transgenes carrying different portions of the protein and determine which, if any portions of the protein could be capable of initiation but not elongation.

Regulation of DAFC-30B and DAFC-62D

Currently, we do not have a clear idea of where the origin for DAFC-30B resides, as we have observed no peak of copy number within the gradient. Thus, the developmental timing of amplification along the entire gradient for DAFC-30B should be determined by real-time PCR. Furthermore, the developmental timing of replication fork movement along the entire DAFC-62D region should also be determined. These are the first steps in determining the behavior and regulation of the new amplicons.
Because the level of amplification is lower, localization of replication proteins to these DAFCs is somewhat difficult to observe. Therefore, chromatin immunoprecipitation experiments could be done to determine when each of the characterized replication factors bind to the amplicons. Also of special curiosity are the transcription factors, E2F, Rb, and Myb. Considering the role these proteins play at DAFC-66D, it is of interest to know whether they may be regulating the origin firing of the new amplicons.
Preliminary bioinformatics searches suggest that there are E2F binding sites, perhaps as many as eight, surrounding the 4kb peak of DAFC-62D. More detailed bioinformatics studies will help to define E2F and Myb binding sites in the new amplicons. ChIP or footprinting experiments could then be used to verify that these proteins are in fact present *in vivo*. Furthermore, the variation of replication or transcription factor binding

throughout various developmental stages should be examined, as these factors could be involved in regulating the late firing of DAFC-62D.

Finally, more stringent bioinformatics studies could help to distinguish regulatory sequences for each of the new amplicons. These studies should be done using very small windows of sequence and looking for even slight variations in A/T content, as even slight increases could be significant. Certainly transgenic studies should be utilized to determine what sequences are important for origin firing at the new amplicons, and are now being examined by another graduate student, Fang Xie.

Other replication factors involved in gene amplification

Recently, a collection of P-element mutants has been donated to the *Drosophila* community by the company Exelixis (Artavanis-Tsakonas, 2004). Among this collection are homozygous viable insertions into genes such as *mcm7, mcm10, cdc45, cdc7,* and *cdc6*. None of these genes have previously been studied during gene amplification, although several have been implicated as playing a role in the process by immunofluorescence (for MCM7, CDC6, CDC45), or by mutant analysis of partners (Dbf4 is the partner of CDC7) (Landis and Tower, 1999; Loebel et al., 2000; Claycomb et al., 2002). Study of these mutants could yield valuable insights into the regulation of origin firing and replication fork movement, and could provide a basis for doing suppressor screens to find interactions between replication factors in gene amplification.

The role of *yellow-g* and *yellow-g2*

We have yet to determine if the *yellow-g* mutant specifically disrupts the vitelline membrane or if it is some other portion of the eggshell that has not properly formed. This could be determined by performing Transmission Electron Microscopy (TEM) studies on the eggshells of mutant mothers.

Why is it that *yellow-g2* is incapable of functioning in place of *yellow-g* in the mutant? Do they perform independent functions or are the levels of both proteins critical for forming a proper eggshell? We could test if the levels were the critical factor by introducing *UAS-yellow-g2* transgenes into the genome and expressing them via *hsp70-gal4* in the *yellow-g* mutant background. It would be possible to test a separate role for *yellow-g2* by generating a mutant in the gene. This could be done via local P-element hopping of the P-element inserted in *yellow-g*, which alone may generate a mutant allele of *yellow-g2*, or could subsequently be used for creating excisions in *yellow-g2*. As with *cdc6*, we should always check P-element screen databases for new insertions in *yellow-g2*. If a genetic allele is not generated, RNAi for *yellow-g2* using the UAS system should be attempted to knock down the *yellow-g2* levels.

Another way to approach the question of separable functions for each of the *yellow* genes is to perform a biochemical assay on purified proteins. One such assay that could be performed is described in (Han et al., 2002) and involves the conversion, through multiple steps, of dihidroxyindole to dopachrome. Most of the various intermediates in this conversion pathway can be detected by spectrometry.. The ability of Yellow-g or g2 to catalyze various steps in this process should be analyzed, and compared to each other and the other Yellow proteins. If Yellow-g was found to have an

enzymatic function in the assay, extracts from *yellow-g* mutant eggs should be prepared and also used in the assay alongside wild-type egg extracts, to show *in vivo* relevance.

Are *yellow-g* and *yellow-g2* only necessary during oogenesis? The mutant phenotype suggests so, but it would be useful to determine the expression profile of each of these genes throughout various stages of development, using RT-PCR or RNA *in situ* hybridization. It may be interesting to test whether other *yellow* family members that are normally expressed during different developmental stages than *yellow-g* and *g2* can function for *yellow-g* in oogenesis when it is mutant. It would also be useful to examine the protein localization of Yellow-g and g2 in the eggshell, by immuno-EM or possibly by regular immunofluorescence. This would require an antibody for each of the proteins to be generated.

Are the *yellow-g* mutant eggs fertilized? The eggs laid by *yellow-g* mutant mothers collapse only as they travel down the uterine tract. The collapsed eggs and bleach sensitivity indicate that egg activation did not occur in the mutant eggs, but we have not yet determined whether the eggs can be fertilized prior to activation. To do this, we should mate homozygous mutant mothers to males carrying a spermtail protein tagged with GFP, and look to see if there is any GFP signal inside the laid eggs. This experiment may be technically difficult because the majority mutant eggs burst in bleach while being dechorionated, but perhaps with sufficient numbers of eggs we could collect enough data to make a determination on the fertilization status.

Neutral red studies indicate that replication factor mutants have compromised vitelline membranes like the *yellow-g* mutant, but do not collapse (Appendix 4). It would

be valuable to perform TEM on the replication factor mutant eggs to determine why they do not collapse and what the overall structure of their eggshell is compared to wild-type.

Examination of other genes in DAFC-30B and 62D

There are a number of other genes of potential interest in the new amplicons whose role in oogenesis is yet to be determined. RNAi studies and P-element screens for mutants in these interesting genes may give us clues into their necessity and function. It would be most advantageous to begin examining genes that were expressed highly in the late stages of oogenesis, such as *CG13113*. More RNA *in situ* hybridizations should be done to determine the expression pattern of all genes in each amplicon, and these results should be used as a basis for selecting which genes to mutate or knock down.

Some of the first genes that should be examined are those that encode the chitin-binding domain proteins, as they are found in both new amplicons. Furthermore, we should attempt to address the question, is chitin a component of the eggshell? This could be done by staining the eggshell with calcofluor or some other chitin-binding dye, but it should be noted that with the autofluorescence of the eggshell and the lack of a positive or negative control for this experiment, the results may be inconclusive.

Transporters and membrane-bound or trans-membrane protein-encoding genes may be of interest, in that they may be parts of signaling pathways necessary for setting up egg polarity and proper eggshell formation or for influencing the timing of amplification and gene expression.

Another group of genes to be examined are those that encode small proteins that appear to have homology to structural elements in other tissues. These proteins may be part of the eggshell or vitelline membrane.

Summary

In summary, there are a number of avenues to further explore in regards to this work. Studies of replication factors, including CDC6 and DUP/Cdt1 and their roles in amplification will help us to better understand the mechanism of gene amplification and of DNA replication in metazoa. Analyzing the sequences of DAFC-30B and 62D in and around peak amplified loci may help us to identify common motifs or elements necessary for regulating gene amplification. Obtaining additional mutants in the DAFC-30B and 62D amplicons will help us to understand the developmental importance of these genes. Further assessing the functions of amplified genes known to be important will aid in understanding the events that occur late in oogenesis and contribute to the formation of a proper eggshell.

References:

Artavanis-Tsakonas (2004). Accessing the Exelixis collection. Nat Genet. *36*, 207.

Bell, S. P., and Dutta, A. (2002). DNA replication in eukaryotic cells. Annu Rev Biochem *71*, 333-374.

Claycomb, J. M., MacAlpine, D. M., Evans, J. G., Bell, S. P., and Orr-Weaver, T. L. (2002). Visualization of replication initiation and elongation in *Drosophila*. J. Cell Biol. *159*, 225-236.

Han, Q., Fang, J., Ding, H., Johnson, J. K., Christensen, B. M., and Li, J. (2002). Identification of *Drosophila melanogaster* yellow-f and yellow-f2 proteins as dopachrome-conversion enzymes. Biochem J *368,* 333-340.

Landis, G., and J. Tower. (1999). The *Drosophila chiffon* gene is required for chorion gene amplification, and is related to the yeast dbf4 regulator of DNA replication and cell cycle. Development *126,* 4281-4293.

Loebel, D., H. Huikeshoven, and S. Cotterill. (2000). Localisation of the DmCdc45 DNA replication factor in the mitotic cycle and during chorion gene amplification. Nucleic Acids Res *28,* 3897-3903.

Appendix One

Generation and Characterization of Mutants in DAFC-30B

After obtaining a mutant in the *yellow-g* gene of DAFC-62D and demonstrating that the amplified genes in this new amplicon were essential for egg viability, we wanted to do the same for DAFC-30B. A number of P-elements were inserted in the region (Figure 1). Thus, we collected as many of these lines as possible and tested them for fertility and overall egg morphology in the homozygous state (if possible; some of the insertions were lethal), and in trans to a deficiency for the region, *Df(2L)N22-3*. The results of this small-scale screen are summarized in Table 1.

As a second approach to generating mutants in DAFC-30B, we mobilized P elements in the region and isolated the resultant excisions or re-insertions of the transposon. We focused these efforts on a gene annotated as *CG33298* (formerly genes *CG13112* and *CG18419*). We chose this gene because it was shown by RNA *in situ* hybridization (see chapter three for details) to be expressed from stages 11-13 in a subset of follicle cells in the anterior-dorsal portion of the egg chamber. Furthermore, this gene is predicted to encode a product with homology to Ca^{2+} transporting ATPases, and if we could demonstrate its necessity for oogenesis, we would again demonstrate that not just structural proteins, but also enzymes, are amplified for proper development.

We chose two P elements inserted near the 3' end of the second exon of *CG33298*, *EP(2)890* and *EP(2)2080* (see Figure 1), because we thought that excisions in the 3' of the gene may generate a truncated protein by removing a portion of the coding region or decrease the stability of the transcript by deleting part of the 3' UTR. These two P elements in particular were chosen because they are inserted on opposite strands, thus potentially eliminating any problems we may have faced by only one P being able to mobilize. The crossing scheme for generating the excision mutants is shown in Figure 2.

Thus far approximately 30 excision lines and 10 re-insertion lines have been generated and will be characterized further.

Figure 1

A number of P-element transposons are inserted in DAFC-30B. This schematic diagram of DAFC-30B indicates the approximate positions of P-elements (red arrowheads) in the region and on which strand they are inserted. Genes in green were present on the microarrays and genes in blue are all other predicted and/or proven genes.

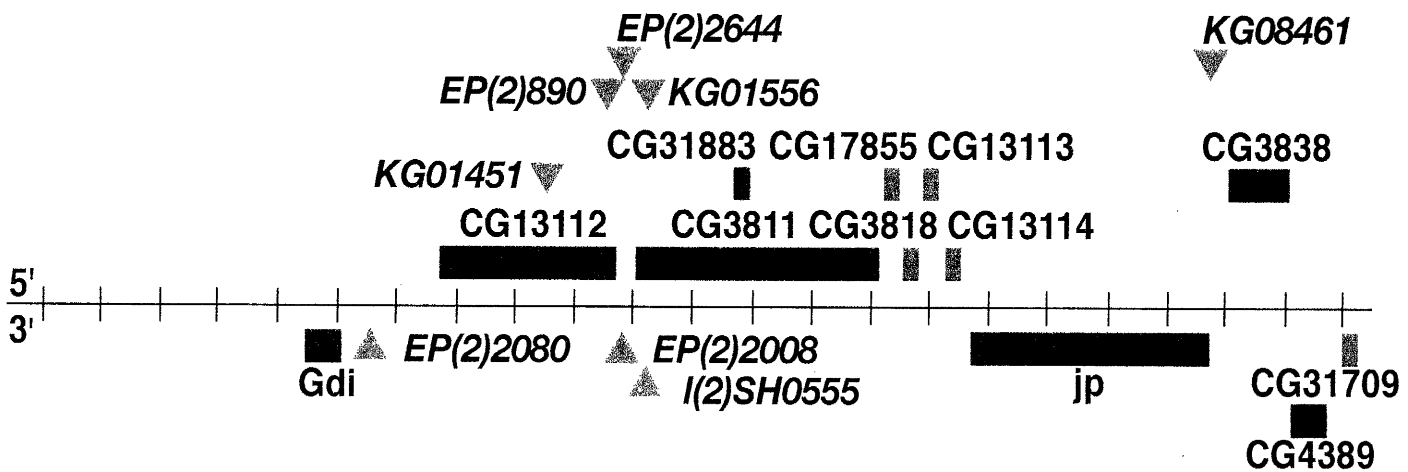

EP(2)2644
KG08461
EP(2)890
KG01556
CG31883 CG17855 CG13113
CG3838
KG01451
CG13112
CG3811 CG3818 CG13114
5'
3'
Gdi
EP(2)2080
EP(2)2008
l(2)SH0555
jp
CG31709
CG4389

Table 1. P-element mutants were examined for their viability and sterility in the homozygous state and/or in trans to a deficiency for the region. Eggs laid by homozygous or transheterozygous females were examined for overall morphology and screened for the uptake of neutral red dye. ND, not determined.

Table 1. Phenotype of P-element insertion mutants in DAFC-30B

P-element line	Homozygous viable?	Female Sterile in trans to Df?	Egg chamber defects?	Neutral Red uptake?
EP(2)890	NO	LETHAL	NONE	ND
EP(2)2008	YES	NO	NONE	ND
EP(2)2080	YES	NO	NONE	ND
EP(2)2644	YES	NO	NONE	ND
l(2)SH055	NO	LETHAL	NONE	ND
KG01451	YES	NO	NONE	NO
KG01556	YES	NO	NONE	NO
KG08461	SEMI	NO	NONE	NO

Crosses to Generate Excisions in DAFC-30B

In mass

w-; P[w+] / CyO × w-; Sp / CyO; Δ2,3 Sb / Tm6 Ubx

In mass

w-; P[w+] / CyO; Δ2,3 Sb / + × w-; Sco / CyO

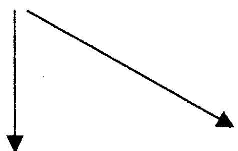
Single males

excision w-; [ ] / CyO or Sco

or

hop w-; P[w+], P[w+] / CyO or Sco

× w-; Sco / CyO



Mate siblings to establish lines

Recover single males for PCR assay

Appendix Two

Studies of the Role of DUP/Cdt1 in *Drosophila* Replication

- **Expression of a *dup* Transgene Under UAS-Gal4 Control**
- **Chromatin IP of DUP at *ACE3***
- **Co-IPs of DUP with other Replication Factors**

Studies of the *dup* Transgene

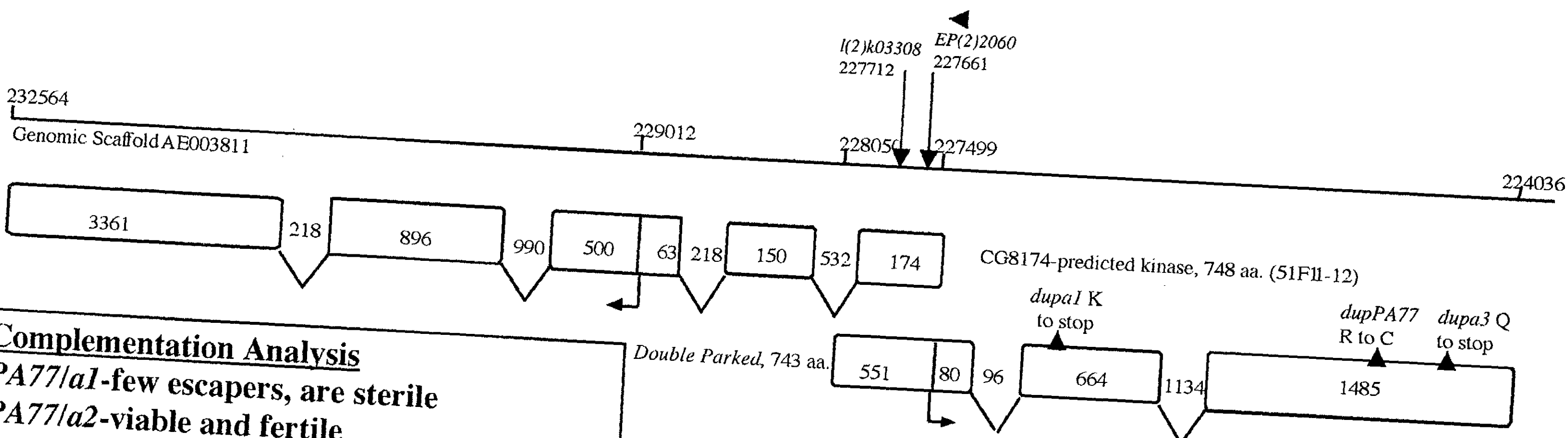
The 5' UTR of *dup* overlaps with the 5'UTR of a putative serine/threonine protein kinase gene, *CG8174,* encoded on the opposite DNA strand (Figure 1A). This is worrisome, as complementation tests of the various *dup* alleles indicated that not all allelic combinations give the *dup* phenotype (Figure 1B). The embryonic lethal alleles, $dup^{a1\text{-}a4}$ and $dup^{l(2)51ec}$, are all lethal in combination with each other (Whittaker et al., 2000). The dup^{pa77} allele is sterile in homozygous form, results in only a few sterile escapers in trans to dup^{a1} and dup^{a3}, and gives rise to no escapers in trans to dup^{a4}. However, when crossed to the lethal alleles, $dup^{l(2)51ec}$ or dup^{a2}, viable and fertile transheterozygous offspring result. This led us to question whether the *dupl(2)51ec* and dup^{a2} alleles, neither of which have been molecularly characterized, could be affecting both *CG8174* and *dup*. To assess which alleles affect *dup* and/or *CG8174*, we set out to rescue the mutant phenotype with a *UAS-dup* cDNA transgene under the control of Gal4.

I generated the pUASp-*dup* construct (named pJC1, Figure 2) and verified that there were no point mutations in the insert by sequencing the plasmid before injecting it into embryos to produce transgenic lines. I obtained four transgenic P[*w+, dup*] fly lines, and determined that three of the insertions were on the *third* chromosome (lines A20, A52, and A57) and one on the *second* chromsome (A62). The P[*w+, dup*] (from each line A20, A52, A57) was crossed into the *dup* background. The driver lines, *nos-Gal4* or *act5c-Gal4*, were also crossed into the *dup* background and then the two resultant lines were crossed to generate flies that had a single copy of the driver, a single copy of the transgene and were mutant for *dup* (transheterozygous combinations of alleles were

Figure 1. The 5' UTR of *dup* overlaps with the 5' UTR of a putative serine-threonine protein kinase gene, *CG8174*. In (A), the genomic region surrounding *dup* on the genomic scaffold AE003811 is depicted. Exons are shown as blocks with sizes in base pairs and introns are intervening lines. The locations of the molecularly-characterized *dup* alleles are denoted, as well (the molecular lesions for the other alleles have not yet been solved). Large arrows show the translational start sights. In (B), the results of complementation tests between the various *dup* alleles are described.

Genomic Region Surrounding *dup*

A.

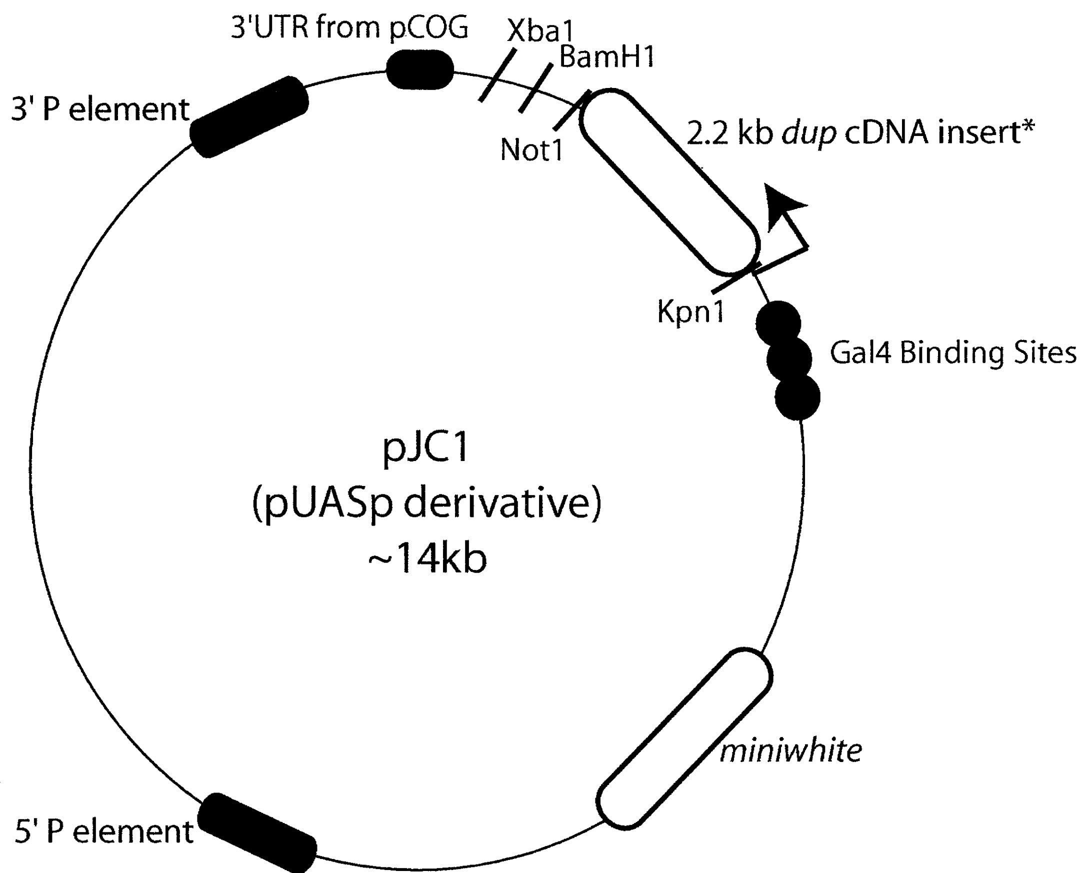


B.

<u>Complementation Analysis</u>

***PA77/a1*-few escapers, are sterile**
***PA77/a2*-viable and fertile**
PA77/a3*-fewer escapers than *PA77/a1
***PA77/a4*-no escapers**
***L5/PA77*-viable and fertile**
***a1* through *a4/a1* through *a4*-lethal**
***EP(2)2060/a1* though *a4, L5, l(2)k03308*-lethal**
***EP(2)2060/PA77*-sterile females, fertile males**
***l(2)k03308/a1* through *a4, L5*-lethal**

Figure 2. Structure of the pJC1 pUASp derivative. The 2.2 kb *dup* coding region was PCR amplified from cDNA clone, LD35784, using primers that added a Kpn1 site to the 5' end and a Not1 site to the 3' end for cloning into the pUASp vector. After cloning the vector was purified via cesium chloride gradient and sequenced before injection into *yw* embryos.



*Used *dup* PCR fragment from clone LD35784 for the insert.

Adapted from P. Rorth, Mechanisms of Development, 1998. 78(1-2): 113-118.

used). The transgene was unable to rescue the sterility or the lethality of any of the *dup* mutants, even those that had been sequenced and characterized. From those alleles (dup^{a1} and dup^{a3}) that have been sequenced and characterized, and that we believe specifically affect *dup*, we would expect to see rescue from the *dup*-associated lethality. The fact that none of the allelic combinations showed rescue led us to conclude that either *dup* expression is tightly regulated and the mutants could only be rescued by a genomic rescue construct but not by overexpression via the UAS-Gal4 system, or that there was a problem with the transgene itself.

To verify that the transgene had not been rearranged upon insertion into the genome, I performed PCR for the cDNA transgene on genomic DNA isolated from single transgenic flies. PCR verified that the transgene was the appropriate size of 2.2kb, and did not appear to be rearranged by gross observation (Figure 3). In addition to testing the transgene itself, I also tested whether mRNA and protein were produced from the transgene by driving P[*w+, dup*] expression in alternating segments of the embryo using the *paired-Gal4* driver, then performing either RNA *in situ* hybridization (Figure 4) or antibody immunofluorescence for Dup protein (Figure 5). These assays demonstrated that both mRNA and protein are produced from the *dup* transgene. In sum, the PCR, RNA *in situ* hybridization, and protein localization studies indicate that the transgene is fully functional and suggest that the complex regulation of *dup* expression may be the culprit for a lack of rescue. However, it remains formally possible that the lack of rescue for some, but not all, *dup* alleles resulted from the mutations affecting the kinase gene and not *dup*.

Figure 3. PCR of the *dup* transgene indicates that it is not rearranged. Single fly genomic DNA preps were made as described in (Ballinger and Benzer, 1989) and PCR was performed on the genomic templates with the primers used in the original cloning of the transgene. (PCR reactions were each 1ml genomic DNA, 2.5ml Extaq buffer, 2ml dNTPs supplied by Extaq, 0.25ml Extaq polymerase, 0.25pM each primer and dH20 to 25ml final volume. Thermocycling was done at 94°C for 5 minutes, then 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, and finally 72°C for 10 minutes.) The genotype of each genomic template is as indicated, where *A52*, *A57*, and *A20* are DNA from fly lines carrying one copy of the *P[w+, dup]* transgene. The *dup* cDNA PCR product is 2.2kb in size (arrow).

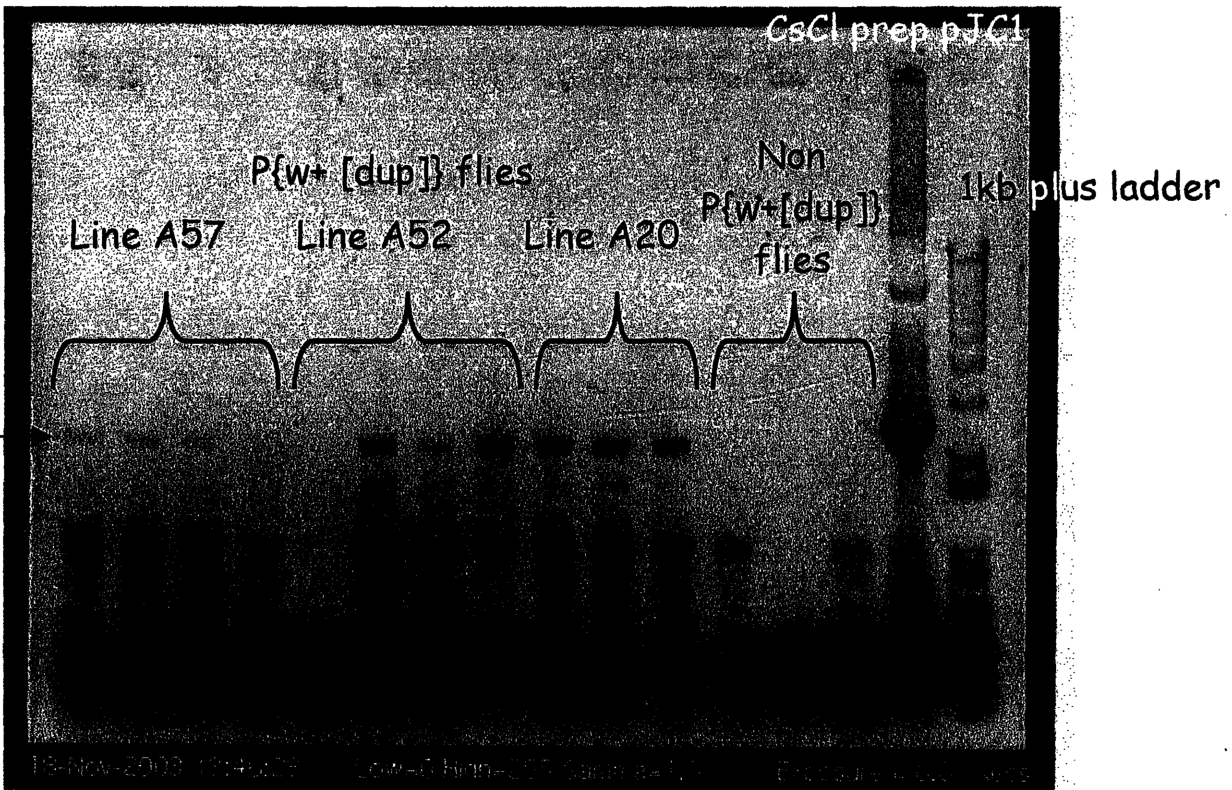
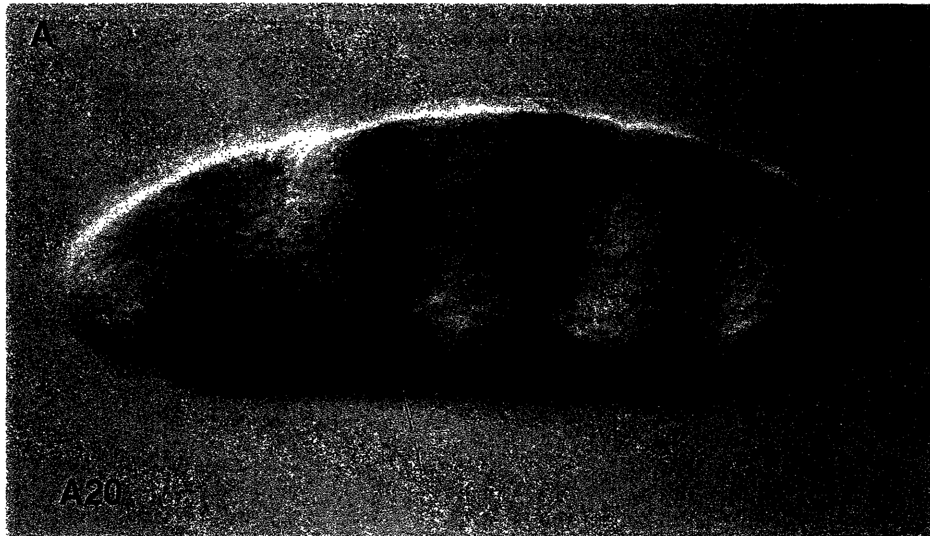
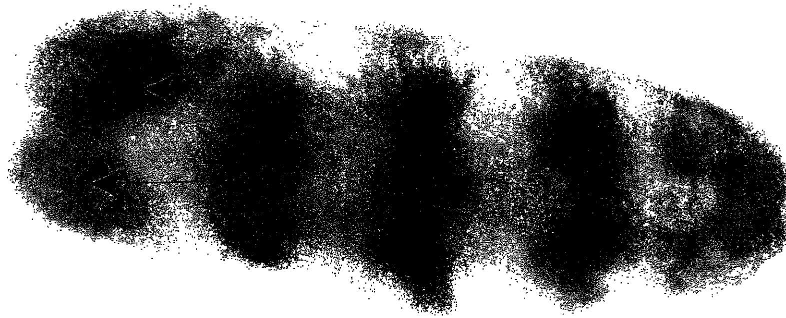

CsCl prep pJC1
P{w+ [dup]} flies
Line A57
Line A52
Line A20
Non P{w+[dup]} flies
1kb plus ladder
2.2 kb dup cDNA
Random other genomic products

Figure 4. Transcripts from the *dup* transgene are detected by RNA *in situ* hybridization when driven by *paired-Gal4*. Embryos carrying the *dup* transgene (Line *A20*, A; *A52*, B; and *A57*, C) and the *paired-Gal4* driver show *dup* mRNA in the pattern of *paired* expression (purple staining), demonstrating that the transgene can be transcribed. Note that in some of the embryos, the normal pattern of *dup* mRNA can also be observed in the developing CNS and PNS (B, arrows show brain staining). (This experiment was done as described in (Royzman et al., 1997) with the *dup* probe being generated from an approximately 680bp fragment of the *dup* second exon. This fragment was generated by PCR and cloned into pBluescript using EcoRI and XhoI sites, as described for other genes in Chapter 3.)



B



A52

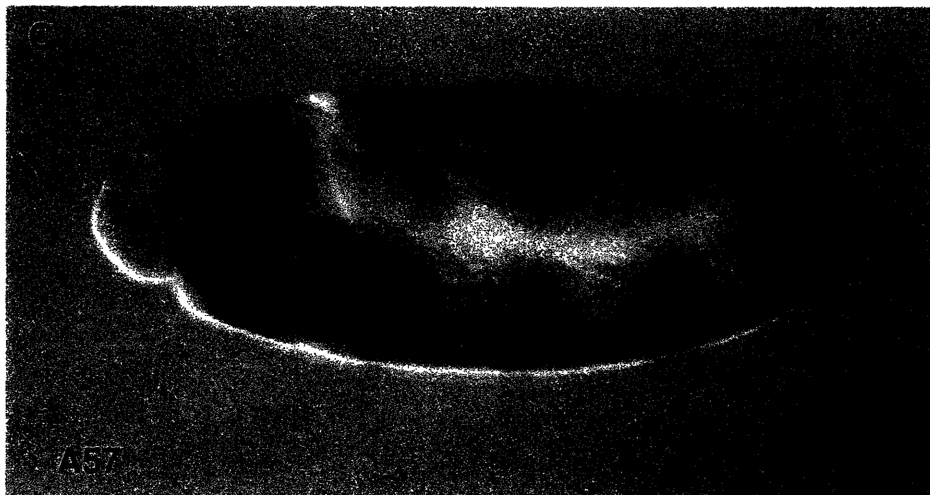
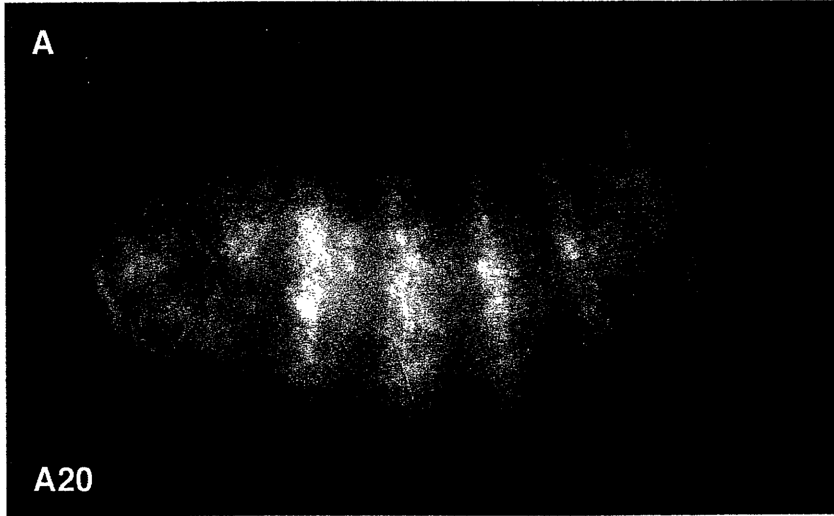
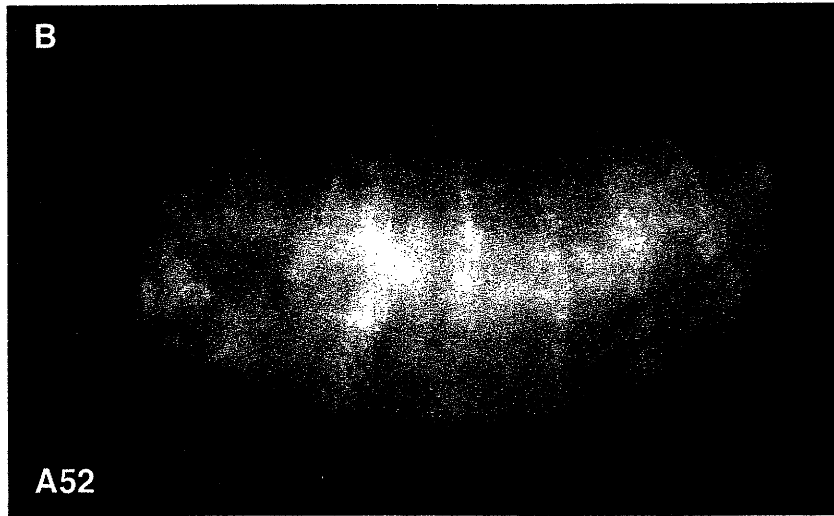
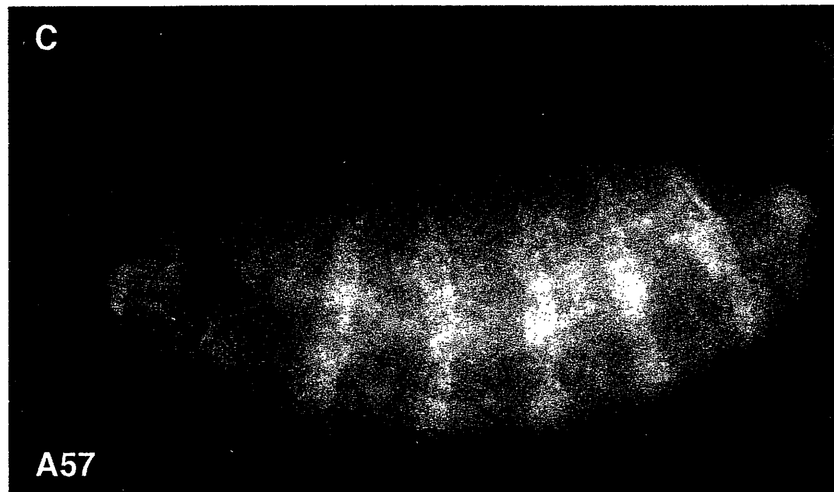


Figure 5. The DUP protein is detected in the *paired* pattern when the *dup* transgene is driven by *paired-Gal4*. Embryos carrying both the *dup* transgene and *paired-Gal4* driver (Lines *A20*, A; *A52*, B; *A57*, C) display DUP staining in alternating segments, indicating that the DUP protein is produced from the transgene. (This experiment was performed by fixing the embryos in formaldehyde and using a 1:500 dilution of the DUP antisera at 4°C overnight.)


A
A20


B
A52


C
A57

ChIP of DUP at *ACE3*

In order to determine if DUP/Cdt1 was directly bound to replication sequences at DAFC-66D, I attempted Chromatin IP experiments using stage 10B egg chambers and the DUP antibody, as has been done for ORC2 (Austin et al., 1999; Bosco et al., 2001). Unfortunately, I was never able to conclusively ChIP DUP/Cdt1 at *ACE3*, although I could ChIP ORC2 at *ACE3* effectively (Figure 6). This may have been due to performing the experiment under sub-optimal ChIP conditions, or simply because DUP/Cdt1 does not directly contact the DNA at *ACE3*.

Co-IPs of DUP and Replication Factors from Ovaries

To determine whether DUP/Cdt1was acting in ovaries as it does in other organisms, functioning in conjunction with CDC6 to load MCM2-7 onto chromatin, Allyson Whittaker initiated Co-IP experiments on ovary extracts, which I followed up on (performed as described in (Bosco et al., 2001). The DUP/Cdt1 antisera effectively precipitate the DUP/Cdt1 protein (Figure 7A, B) (Quinn et al., 2001), but neither Allyson nor I was ever able to see a convincing interaction between DUP/Cdt1 and any of the replication factors tested, including ORC2, CDC6, MCM2, 3, 5 (See Figure 7A for MCM2, Figure 7D for CDC6). However, there may be a weak interaction between DUP/Cdt1 and CDC6 (Figure 7D, arrow). Incidentally, there seems to be conflicting data about whether the anti-CDC6 (antisera #2144, see Appendix 3) can IP the CDC6 protein (Figure 7C shows a positive IP, Figure 7D shows a negative result). Some of the reverse Co-IPs were also attempted, again showing no conclusive interactions. It is entirely possible that the conditions were not optimized to detect the interactions between the

proteins, or the interactions may be transient enough not to be detectable without crosslinking to enrich for complexes.

References:

Austin, R. J., Orr-Weaver, T. L., and Bell, S. P. (1999). *Drosophila* ORC specifically binds to *ACE3*, an origin of DNA replication control element. Genes and Dev *13*, 2639-2649.

Ballinger, D. G., and Benzer, S. (1989). Targeted gene mutations in *Drosophila.* Proc Natl Acad Sci USA *86*, 9402-9406.

Bosco, G., Du, W., and Orr-Weaver, T. L. (2001). DNA replication control through interaction of E2F-RB and the origin recognition complex. Nature Cell Biol *3*, 289-295.

Quinn, L. M., Herr, A., McGarry, T. J., and Richardson, H. (2001). The *Drosophila* Geminin homolog: roles for Geminin in limiting DNA replication, in anaphase and in neurogenesis. Genes and Dev *15*, 2741-2754.

Royzman, I., Whittaker, A. J., and Orr-Weaver, T. L. (1997). Mutations in *Drosophila DP* and *E2F* distinguish G1-S progression from an associated transcriptional program. Genes Dev *11*, 1999-2011.

Whittaker, A. J., Royzman, I., and Orr-Weaver, T. L. (2000). *Drosophila* double parked: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. Genes Dev *14*, 1765-1776.

Figure 6. Chromatin Immunoprecipitations of ORC2 and DUP/Cdt1 at *ACE3*, from stage 10B egg chambers. ORC2 can be ChIPed effectively at *ACE3* (lanes 1 and 2), as compared to input serial dilutions. DUP/Cdt1, on the other hand, was unable to be ChIPed at *ACE3* in multiple experiments, and the results of two such experiments are shown here (DUP ChIP A and B).

ORC2 and DUP ChIP Experiments

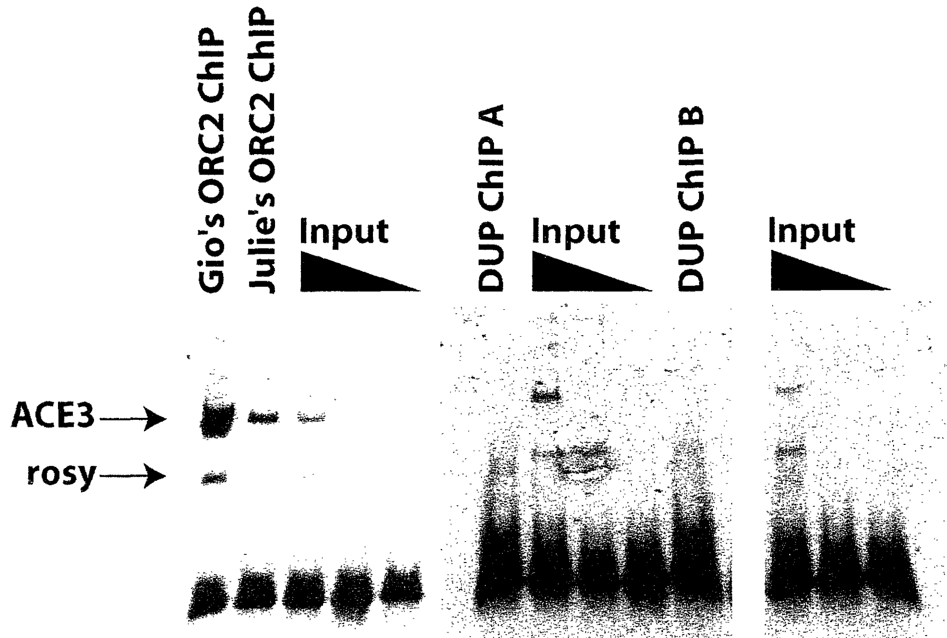
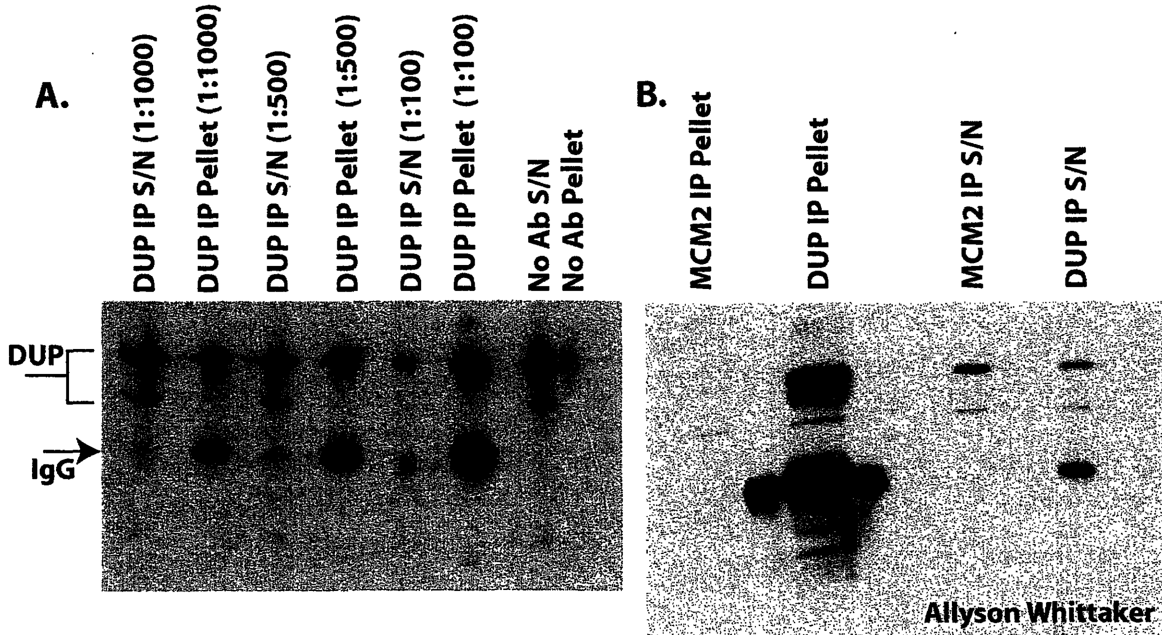
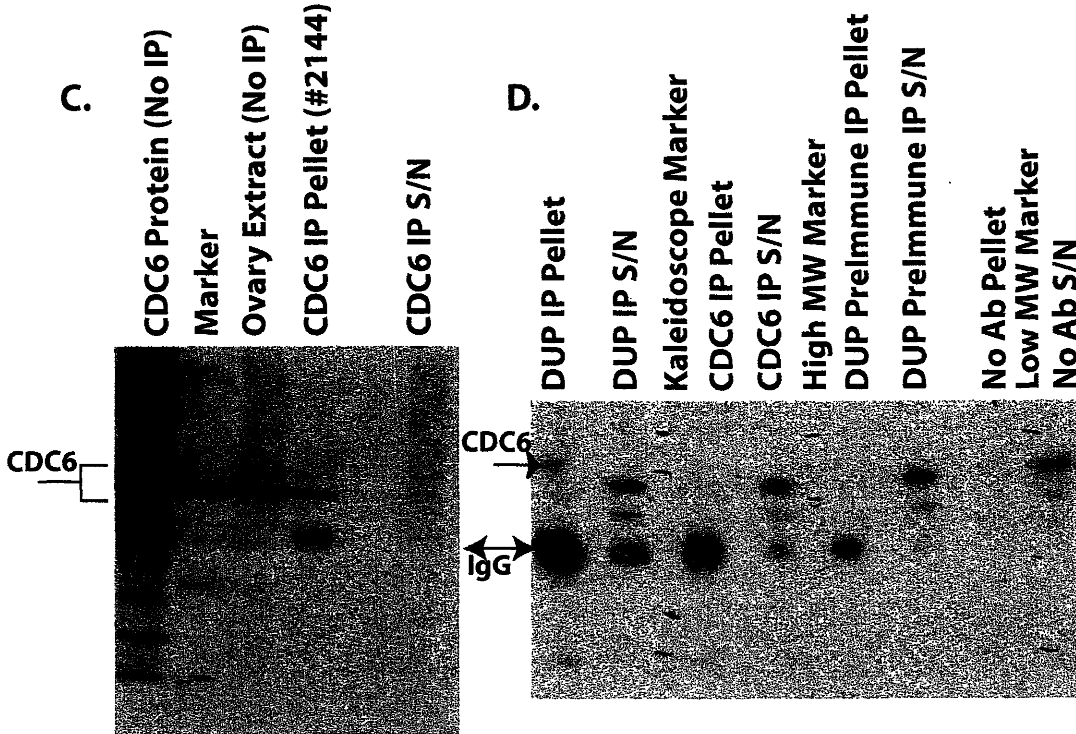


Figure 7. Immunoprecipitations and Co-Immunoprecipitations of DUP/Cdt1 and other replication factors. A. Various concentrations of anti-DUP (antibody #1574) were used to IP the protein from ovary extracts. B. An IP experiment done by Allyson Whittaker (Western is probed with anti-DUP) shows similar results to A, and demonstrates that the anti-MCM2 cannot effectively Co-IP the DUP/Cdt1 protein. Note the various migrating forms of DUP/Cdt1 protein. C. IP of CDC6 with the 2144 antibody (Western is probed with anti-CDC6) at a 1:400 dilution shows that the antibody can IP the CDC6 protein. (Note the spillover of ovary extract into the Marker lane. Purified CDC6 was the baculovirus expressed protein used as the antigen.) D. An IP experiment (Western is probed with anti-CDC6, 1:400 antibody dilutions) using anti-DUP shows that there may be some Co-IP of CDC6 with DUP/Cdt1 (arrow). This experiment gave results contrary to C in that it did not demonstrate that the anti-CDC6 can IP the CDC6 protein.

DUP IPs



CDC6 IPs



Appendix Three

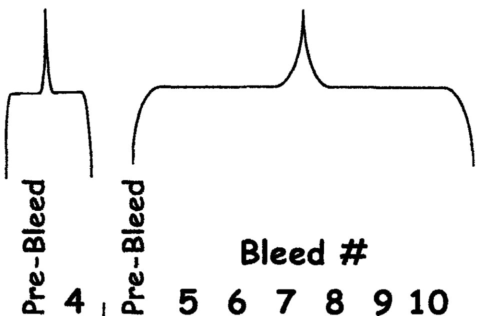
Preliminary Studies of the Role of CDC6 in *Drosophila* Gene Amplification

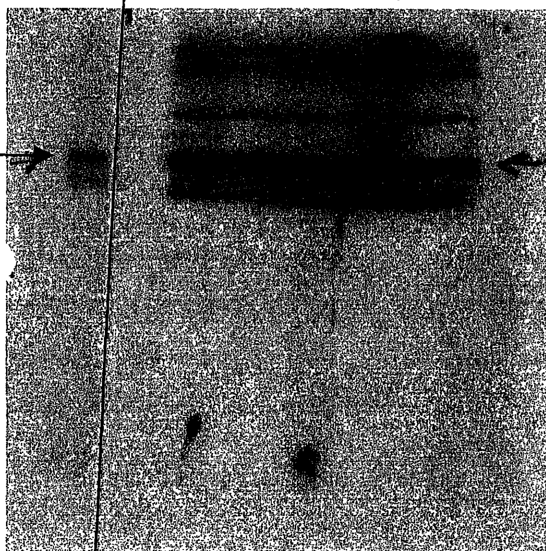
To date, no mutants in the gene *cdc6* have been described and no studies of the function of CDC6 in *Drosophila* have been initiated. A former postdoc in the Bell lab, Rick Austin, began the process of generating CDC6 antibodies by creating baculoviruses to express the protein. Giovanni Bosco, a former postdoc in the Orr-Weaver lab then expressed the protein and sent the purified protein for injection into guinea pigs. Two forms of CDC6 protein were injected into two animals: a native conformation protein preparation and a denatured protein preparation (denatured by treatment with SDS). I tested the sera generated by these guinea pigs on Western blots of *Drosophila* ovary protein extracts, and showed that the antibodies generated to the native CDC6 (#2144) recognized only two bands, one at approximately 86kDa and one at a slightly lower molecular weight. The antibodies generated to the denatured CDC6 protein (#2143) recognized a number of proteins of various molecular weights, with the predominant band matching the 86kDa band of the 2144 antibody (Figure 1). As CDC6, with 643 amino acids, is predicted to have a molecular weight of approximately 86kDa, it is likely that the antibodies are recognizing CDC6, although without a null or truncation mutant in CDC6 this will be difficult to determine.

With the 2144 antibody in hand, I wanted to determine whether CDC6 localized to sites of gene amplification in the follicle cells by immunofluorescence. The protein was localized to the cytoplasm of both follicle cells and nurse cells until stage 10A, although some nuclei (both follicle cell and nurse cell) displayed slight nuclear localization, as well (Figure 2). By stage 10B, the CDC6 protein relocated from the follicle cell cytoplasm to the follicle cell nuclei, and some follicle cell nuclei displayed

Figure 1. The CDC6 antisera recognize a predominant band of approximately 86kDa, the predicted size of the CDC6 protein. The 2144 antibody was generated to native CDC6 and recognizes only 2 bands in ovary extracts, while the 2143 antibody was generated to denatured CDC6 and recognizes multiple bands in the ovary extracts. Both antibodies, however, recognize the same ~86kDa band, which is likely to be CDC6. Pre-immune sera from each of the guinea pigs used to generate the antibodies are also shown, and the bleed numbers from which the sera were taken are indicated above the lanes. The antibodies were used at a 1:7500 concentration. (See Bosco, Du et al. 2001 for details on the Western blot.)

GP Anti-CDC6 Recognizes a band of CDC6's Approximate Size



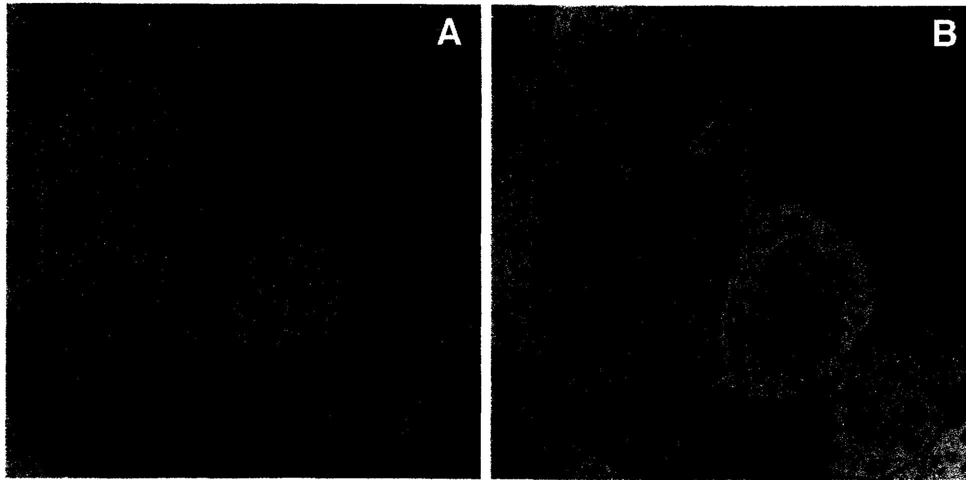


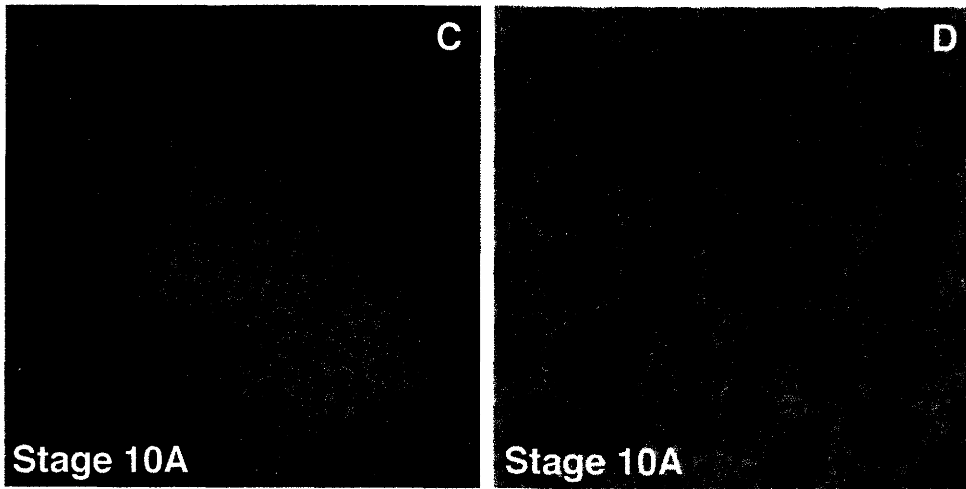
Antibodies:
#2144= to native CDC6
#2143=to denat. CDC6

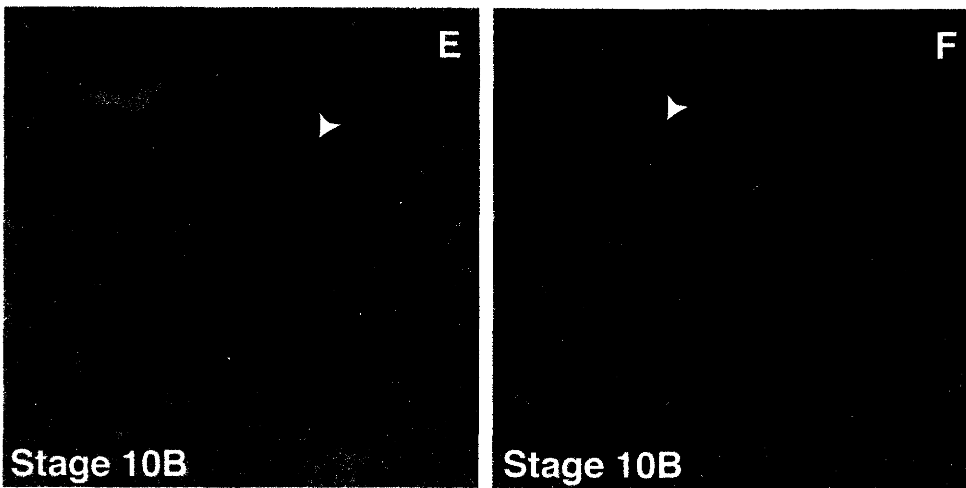
1:7500 Ab dilution
Ovary extract
20" exposure

Antibody generated by Rick Austin and Giovanni Bosco

Figure 2. The 2144 antibody shows a dynamic localization pattern in ovaries. The antibody shows largely cytoplasmic, but perhaps some nuclear staining, in both follicle cells (A, C, D) and nurse cells (B) until stage 10B. In stage 10B (E, F), the protein recognized by the 2144 antibody shifts its localization from cytoplasmic to nuclear in the follicle cells, and in a subset of eggchambers, subnuclear foci are observed (E, F arrows). It is possible that these foci correspond to one of the amplified regions (most likely the third chromosome chorion locus), and the transient nature of localization to these foci is consistent with the activity of CDC6 in other organisms. The caveats to this experiment are that we have no *cdc6* mutant to test for the specificity of the antibody, nor was the pre-immune sera from the guinea pig in which the 2144 antibody was generated tested on ovaries to show specificity.


A
B


C
D
Stage 10A
Stage 10A


E
F
Stage 10B
Stage 10B

subnuclear foci (Figure 3). The subnuclear foci appeared uniform per an egg chamber, however very few egg chambers displayed the subnuclear foci. This may indicate that CDC6 localizes only temporarily to the amplified regions and its localization is a difficult event to detect, similar to the case in yeast (for review, see Bell and Dutta 2002). It seems that the high salt/high detergent buffer used to wash non-chromatin bound MCM2-7 from the nucleus, allowing us to see MCM2-7 at amplified loci (see Chapter 2), will not be a viable strategy for attempting to visualize CDC6 at the amplified loci, because CDC6 is stripped from the chromatin, even in relatively low salt conditions. Thus, we must utilize several other strategies to determine if CDC6 is present and necessary for gene amplification.

A first step to assessing the function of CDC6 is to perform immunoprecipitations with the CDC6 antibody and determine if the protein can be pulled out of extracts from amplification-stage egg chambers (or any replicating tissue). I have tried preliminary IP experiments (see Appendix 4, Figure 7), but do not have conclusive data about the antibody. If the antibody can IP the CDC6 protein, co-IPs should be attempted to determine if CDC6 is interacting with any of its usual replication factor partners, such as ORC2, DUP/Cdt1, and MCM2-7, for which we have antibodies, and the reverse co-IPs should also be done. Again, I have attempted preliminary experiments of this sort and have failed to see an interaction between CDC6 and DUP/Cdt1 in ovaries, but have not tried to vary the conditions for the IP, something that may be necessary to observe the interaction.

If the antibodies are capable of immunoprecipitating CDC6, they could be used to attempt chromatin immunoprecipitation of CDC6 at the amplicons. This may be a difficult experiment, as CDC6 is not easily ChIPed in other systems. Perhaps alternate fixation methods giving higher efficiency of cross-linking should be utilized, and/or the number of egg chambers collected for the experiment may need to be increased, if loss of chromatin throughout the procedure is a problem. The ChIP should first be tested at DAFC-66D, because this amplicon possesses the highest copy number and would be most likely to give a positive result. Stages 10A and 10B should be tested both separately and combined, because CDC6 may be present at the amplified loci prior to the loading of other replication factors, and may only be present for a short developmental window.

As no mutants were available in *cdc6*, we turned to overexpression studies in an attempt to implicate CDC6 in gene amplification. I generated pUASP-CDC6 constructs, in a manner similar to the *dup* transgene (see Appendix 2), and verified that there were no point mutations in the construct by sequencing. With the help of Helena Kashevsky, the pUASP-CDC6 construct was injected into embryos and approximately 20 transgenic fly lines were obtained. After isolating a P[*w+, cdc6*] line with the transgene on the *third* chromosome, I crossed the transgene to both an *hsp70-Gal4* driver line and the *323a-Gal4* driver line, in which Gal4 is expressed in the follicle cells from stage 10B onward (as well as in other tissues; Manseau, Baradaran et al. 1997). When *cdc6* expression was driven by either of the Gal4 driver lines, no appreciable differences in BrdU incorporation during gene amplification stages were observed. Elevated levels of CDC6 protein were detected in the nuclei of follicle cells when expression was driven by *323a-*

gal4, but no striking differences in CDC6 levels were observed when expression was driven by *hsp70-gal4*. These experiments should be repeated to thoroughly assess the role of CDC6 in amplification.

The CDC6 expression studies should be pushed further, trying other transgenic lines, as the previously described experiments were done using only a single transgenic line, and with only one copy of the transgene. Recently, we have obtained a number of *pUASp-cdc6* transgenic lines from Dr. Maki Asano, a new collaborator on this project. Perhaps different lines or additional transgene copies would display different effects from the overexpression, and this could be monitored by BrdU incorporation or real-time PCR. Additionally, it is important to determine what effect overexpressing CDC6 has on the levels of MCM2-7. Thus the MCM2-7 staining should be performed in the background where CDC6 is overexpressed to test for increased MCM2-7 present on the chromatin. Finally, the CDC6 produced by overexpression could be inactive without its partners, so additional replication factors could be coexpressed with CDC6, such as Dup/Cdt1.

As stated above, until a *cdc6* mutant is obtained, we will not know for certain the necessity of CDC6 in gene amplification, or in any DNA replication, in *Drosophila*. We should continue to search P-element insertion databases to be aware if any insertions in or nearby *cdc6* are generated. These alleles may have a phenotype of their own, or could be used to perform a small-scale P-element excision or hopping screen, looking for deletions of *cdc6* or new insertions into the gene that affect transcript and protein levels. A potential allele of *cdc6* will deposited in the near future by the company Exelixis, so it will be important to obtain this line as soon as possible.

It seems that *cdc6* may be a difficult gene in which to generate a genetic mutation, as no mutants have yet arisen. Thus, an alternative to obtaining a mutant in *cdc6* would be to perform RNA Interference studies. *cdc6* RNAi could be administered during amplification stages using the *UAS*-gal4 system, similar to CDC6 overexpression experiments, with *hsp70*-gal4 or *323a*-gal4 drivers. As controls for the effectiveness of the RNAi, the levels of *cdc6* mRNA could be examined by RNA *in situ* hybridization, and the level of CDC6 protein could be monitored by antibody staining. BrdU and possibly real-time PCR could be used to determine the extent of gene amplification when the RNAi was administered.

CDC6 localizes to endocycling follicle and nurse cells, so the RNAi studies could be expanded to examine the necessity of CDC6 in follicle and nurse cell endocycles, with the *hsp70*-gal4 or *nos*-gal4 (for nurse cells only) drivers. Furthermore, the localization pattern of CDC6 seems to indicate a movement of the protein between the nurse and possibly follicle cell nuclei and cytoplasm, so it should be determined by CDC6 and BrdU co-labeling how closely the presence of CDC6 in the nucleus correlates with the onset of S-phase.

How is the activity of CDC6 regulated during amplification? Is CDC6 a target of Cyclin E, and can overexpressing Cyclin E drive CDC6 onto the chromatin or into the nucleus? Alternatively, does Cyclin E add an inhibitory phosphorylation on CDC6 to send it out of the nucleus during amplification stages? We have a *UAS-cyclinE* line in the lab (Richardson et al., 1995) that could be used to drive expression in the follicle cells, and then CDC6 staining can be done to establish whether the pattern of CDC6 localization changes in response to increases in Cyclin E levels. Western blots could also

be done to determine if there is any change in mobility for the CDC6 protein. This would require a purified population of follicle cells from which protein extracts would be generated (Bryant et al., 1999). If possible, nuclear and cytoplasmic protein preparations could also be made to look at the condition of CDC6 in these two compartments and see if it varies.

References:

Bell, S. P. and A. Dutta (2002). "DNA replication in eukaryotic cells." Annu. Rev. Biochem. **71**: 333-74.

Bosco, G., W. Du, et al. (2001). "DNA replication control through interaction of E2F-RB and the origin recognition complex." Nature Cell Biol. **3**: 289-295.

Bryant, Z., Subrahmanyan, L., Tworoger, M., LaTray, L., Liu, C. R., Li, M. J., van den Engh, G., and Ruohola-Baker, H. (1999). "Characterization of differentially expressed genes in purified *Drosophila* follicle cells: toward a general strategy for cell type-specific developmental analysis." Proc Natl Acad Sci USA **96**: 5559-5564.

Manseau, L., A. Baradaran, et al. (1997). "GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of *Drosophila*." Dev. Dynamics **209**: 310-322.

Richardson, H., O'Keefe, L. V., Marty, T., and Saint, R. (1995). Ectopic cyclin E expression induces premature entry into S phase and disrupts pattern formation in the *Drosophila* eye imaginal disc. Development **121**: 3371-3379.

Appendix Four

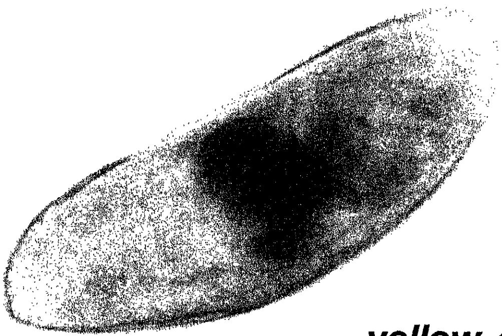
Analysis of Vitelline Membrane Integrity in *yellow-g* and Replication Factor Mutants Using the Neutral Red Assay

The *yellow-g* collapsed egg phenotype is reminiscent of defects in the formation of the vitelline membrane, and we have hypothesized that *yellow-g* encodes an enzyme necessary for the crosslinking of the vitelline membrane and/or chorion layers. Previously, several other groups have used the uptake of neutral red dye by de-chorionated embryos as an assay for the integrity of the vitelline membrane (Komitopoulou, Gans et al. 1983; Degelmann, Hardy et al. 1990; Konrad, Wang et al. 1993; LeMosy and Hashimoto 2000). Embryos with intact vitelline membranes are unable to absorb the dye, while embryos possessing a compromised vitelline membrane take up the dye become red to varying degrees. When the embryos laid by *yellow-g* mothers were subjected to the assay, the vast majority burst upon exposure to bleach in the de-chorionation step, another indication that the vitelline membrane has not been properly crosslinked upon egg activation (Limbourg and Zalokar 1973). Of those embryos that survived the de-chorionation, nearly 100% absorbed the neutral red dye, as compared to the embryos laid by heterozygous sibling controls in which virtually none of the embryos absorbed the dye (Figure 1). This indicates that the vitelline membranes of the embryos laid by *yellow-g* mutant mothers are compromised.

These data left us with the conundrum: mutants with compromised vitelline membranes display collapsed eggs, yet the replication factor female sterile mutants *mcm6*, *chiffon/dbf4-like*, and *dup* do not show the same collapsed egg phenotype even though they amplify DAFC-62D (the amplicon in which *yellow-g* resides) to no significant degree and display significantly reduced *yellow-g* transcript levels. This

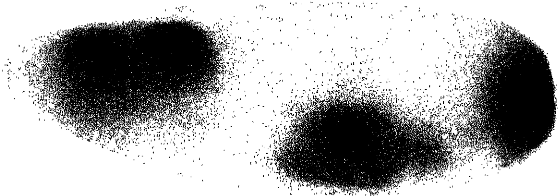
Figure 1. The embryos laid by *yellow-g* mutant mothers absorb neutral red dye. In (A), a control embryo laid by a heterozygous mother shows no neutral red uptake. (B-D) show increasing degrees of dye uptake in embryos laid by *yellow-g* mutant mothers. The staining was done as described in (LeMosy and Hashimoto 2000), using *EY01493/EY01493* or *EY01493/TM3* females. Anterior is left, and images were collected as described in Chapter 3 for RNA *in situ* hybridizations.

A



yellow-g $^{+/-}$

B



yellow-g $^{-/-}$

C

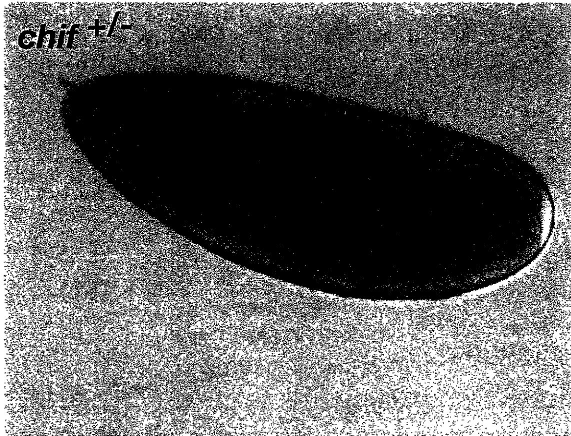
yellow-g $^{-/-}$

D

yellow-g $^{-/-}$

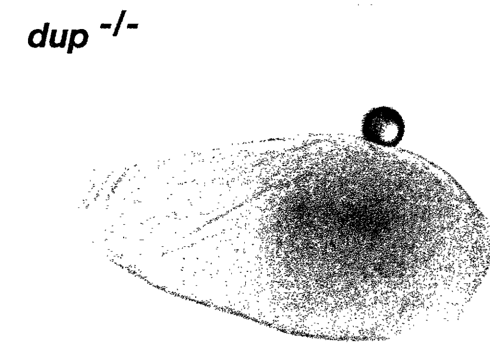
observation caused us to question whether amplification of *yellow-g* was truly necessary to produce a proper vitelline membrane. Thus, we employed the neutral red assay on embryos laid by the replication factor mutant mothers. We observed that the majority of embryos laid by mutant mothers were positive for neutral red uptake, and up to approximately one-quarter of these embryos actually burst upon bleach exposure prior to the staining (Figure 2). These data are consistent with the notion that the amplification of *yellow-g*, and perhaps other amplified genes, is necessary for the proper formation of the vitelline membrane. We hypothesize that the reason the embryos laid by replication factor mutant mothers do not collapse when laid is that, in addition to having the disrupted vitelline membrane, they do not have the full force of an intact chorion weighing down upon the compromised membrane, as the chorion amplicons have not been sufficiently amplified or expressed in these mutants either.

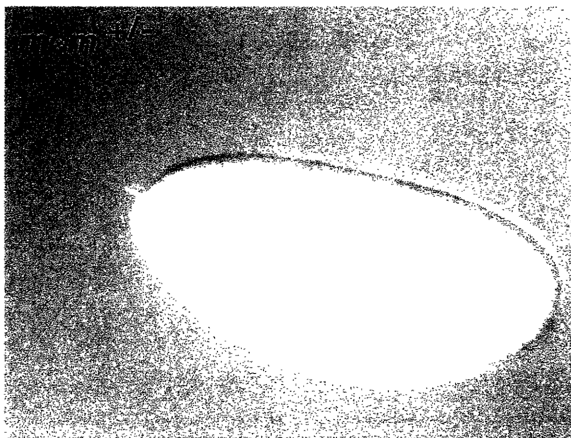
Figure 2. Embryos laid by replication factor mutant mothers absorb neutral red dye. Embryos laid by heterozygous sibling mothers are shown in the left column, those laid by mutant mothers are in the right column. The allelic combinations used are as described in Chapter 3 (Figure 5). Anterior is left.

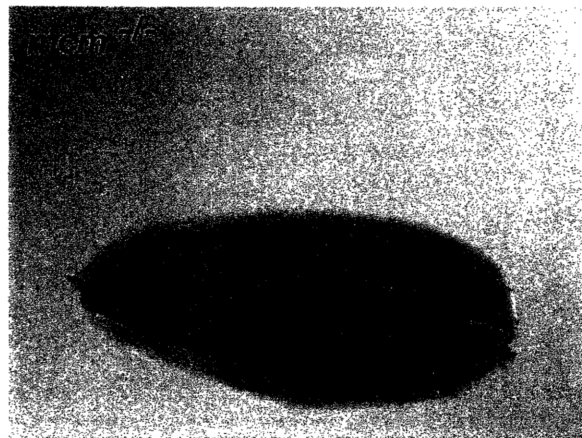


dup+/-

dup -/-







References:

Degelmann, A., P. A. Hardy, et al. (1990). "Genetic analysis of two female-sterile loci affecting eggshell integrity and embryonic pattern formation in *Drosophila melanogaster*." Genetics **126**(2): 427-34.

Komitopoulou, K., M. Gans, et al. (1983). "Isolation and characterization of sex-linked female-sterile mutants in *Drosophila melanogaster* with special attention to eggshell mutants." Genetics **105**: 897-920.

Konrad, K. D., D. Wang, et al. (1993). "Vitelline membrane biogenesis in *Drosophila* requires the activity of the alpha-methyl dopa hypersensitive gene (I(2)amd) in both the germline and follicle cells." Insect Mol. Biol. **1**(4): 179-87.

LeMosy, E. K. and C. Hashimoto (2000). "The nudel protease of *Drosophila* is required for eggshell biogenesis in addition to embryonic patterning." Dev. Biol. **217**: 352-361.

Limbourg, B. and M. Zalokar (1973). "Permeabilization of *Drosophila* eggs." Dev. Biol. **35**: 382-387.

Appendix Five:

Real-time PCR Determination of *ACE1* (DAFC-7F) and DAFC-30B Developmental Timing of Amplification

Having determined the profile of amplification for DAFC-66D by real-time PCR during different developmental stages, we thought it important to do the same for DAFC-7F. By performing the real-time PCR experiments in 10kb intervals on this amplicon, we were able to determine that the timing of origin firing was the same for DAFC-7F as it was for DAFC-66D, in that, origin firing occurred during stages 10B and 11 of egg chamber development, and in stages 12 and 13 only the existing replication forks progressed outward (Figure 1). The DAFC-7F amplicon reaches a peak copy number of approximately 14-fold amplification in this assay, as compared to the 18 to 20-fold amplification observed by quantitative Southern blotting (Spradling 1981; Delidakis and Kafatos 1989).

In addition to determining the developmental timing of amplification in DAFC-7F, we wanted to know what the developmental profile was for the DAFC-30B amplicon. DAFC-30B displays a plateau of peak copy number, over approximately 75kb in stage 13. This could indicate that origin firings ended earlier than for DAFC-7F and 66D, and that in subsequent stages, the replication forks had more time to proceed bidirectionally, and thus replicated more of the flanking sequences than in the original two amplicons. To test this, we chose three primer sets in the center of the gradient and performed the real-time PCR in different developmental stages at these loci. We chose these primer sets assuming that they were at or near the replication origin, because they were in the center of the copy number gradient, where the origins for DAFC-7F and 66D reside. However, if the origin is in any way offset from the center, these primer sets would not be the appropriate ones to use in the assay for deterimining the timing of origin firing. With that caveat in mind, we observed that all copy number increases had already occurred by

stage 10B at each of these loci (Figure 2), and we interpreted this to mean that all origin firing at DAFC-30B occurs earlier in stage 10B than we could observe in our heterogeneous population of egg chamber DNA (stage 10B is the longest stage of amplification, over 6 hours). Other possibilities include that origin firing may have occurred in stage 10A or, as stated above, that we chose the incorrect loci to test, thus it would be useful to perform the real-time PCR in various developmental stages across all of DAFC-30B.

References:

Delidakis, C. and F. C. Kafatos (1989). "Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*." EMBO J. **8**: 891-901.

Spradling, A. C. (1981). "The organization and amplification of two clusters of *Drosophila* chorion genes." Cell **27**: 193-201.

Figure 1. Amplification at DAFC-7F displays a similar profile to that at DAFC-66D. Quantitative real-time PCR experiments were performed on the DAFC-7F amplicon in 10kb intervals along the 100kb amplified region as described in Chapter 2. Origin firing occurs in stages 10B and 11, while in stages 12 and 13, only elongation occurs. Peak copy number at the *ACE1* region is 14-fold amplification.

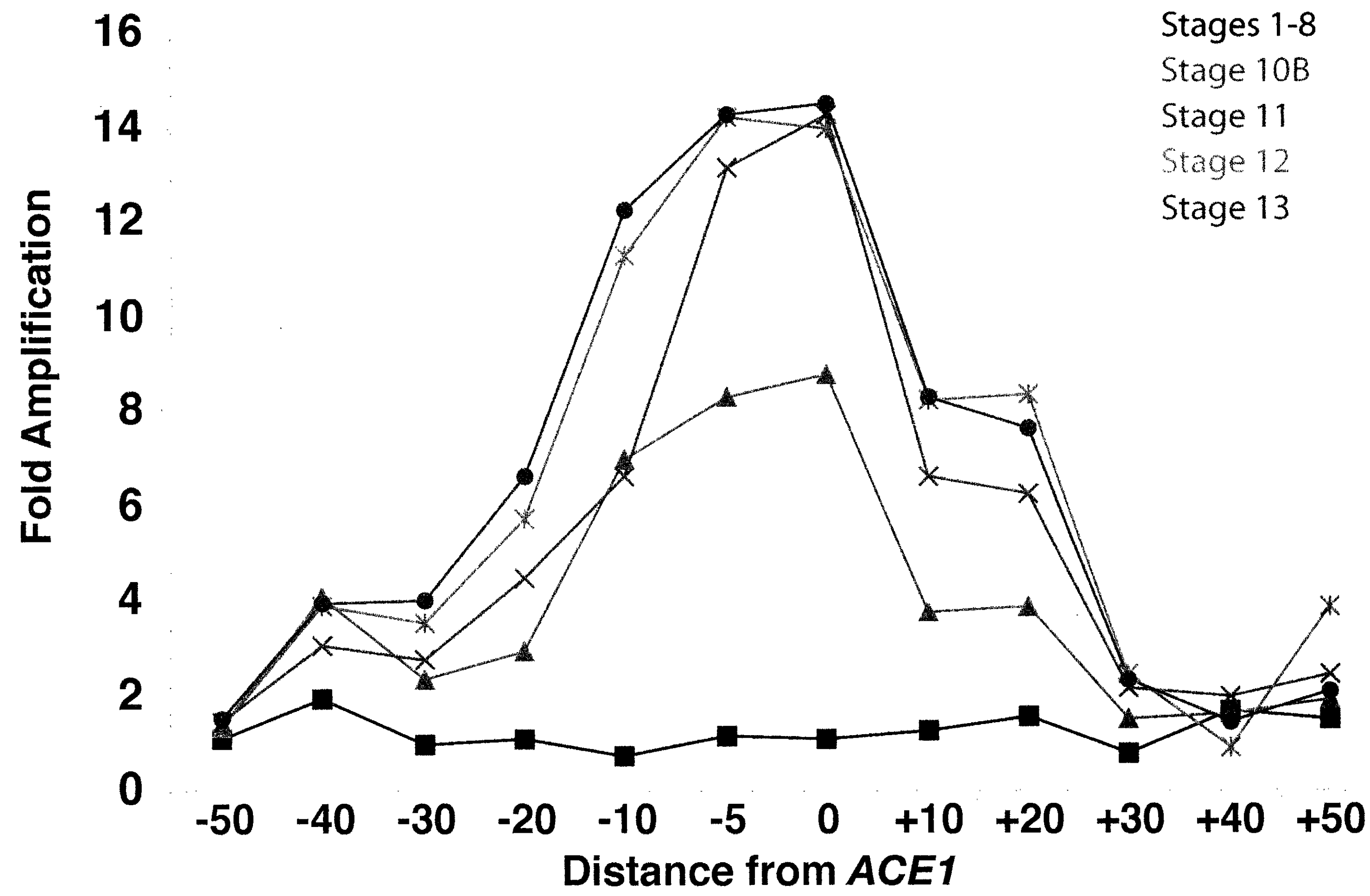
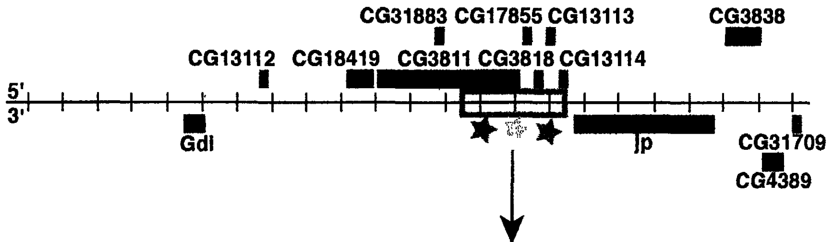
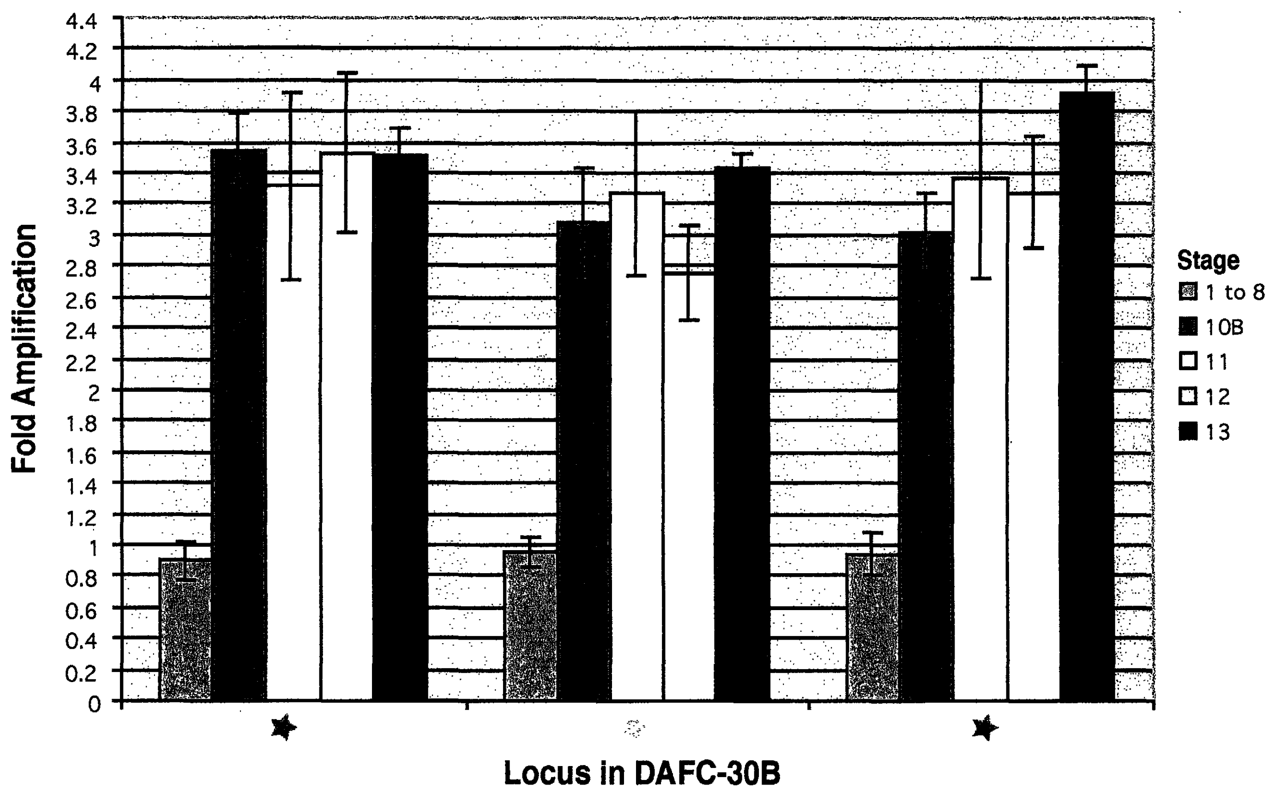

Fold Amplification
16
14
12
10
8
6
4
2
0
-50
-40
-30
-20
-10
-5
0
+10
+20
+30
+40
+50
Distance from ACE1
Stages 1-8
Stage 10B
Stage 11
Stage 12
Stage 13

Figure 2. Origin firing appears to have ended in DAFC-30B by stage 10B. Real-time PCR was performed (as in Chapters 2 and 3) at three loci (stars) along DAFC-30B in each stage of gene amplification. These results are depicted in bar graph format, with the loci chosen as the x-axis, the fold amplification as the y-axis, and each developmental stage shown in a different color (see legend).


CG31883 CG17855 CG13113
CG3838
CG13112 CG18419 CG3811 CG3818 CG13114
5'
3'
Gdl
jp
CG31709
CG4389


Fold Amplification
4.4
4.2
4
3.8
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Stage
1 to 8
10B
11
12
13
Locus in DAFC-30B

~The End~