OCCURRENCE AND IMPLICATIONS OF BIOLOGICAL NETWORK EVOLUTION FOLLOWING POLYPLOIDY

A Thesis

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

DUSTIN MAYFIELD-JONES

Dr. J. Chris Pires

MAY 2014

The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

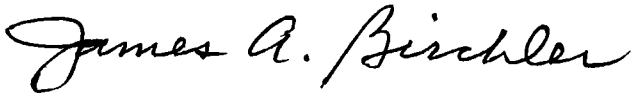
OCCURRENCE AND IMPLICATIONS OF BIOLOGICAL NETWORK EVOLUTION FOLLOWING POLYPLOIDY

presented by Dustin Mayfield-Jones,

a candidate for the degree of master of biology,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor J. Chris Pires



Professor James A. Birchler

Professor Gavin C. Conant

ACKNOWLEDGEMENTS

First and foremost, I thank my advisor, Dr. Chris Pires. He has been an inspiration to me both professionally and personally. His consistent demonstration of an incredibly high degree of curiosity and enthusiasm for his work is most admirable. Without his support and expertise, this work would not have been possible.

I thank my committee members, Drs. Gavin Conant and Jim Birchler, and Dr. Paula McSteen, for their time and thoughtful conversations about my work and life.

Also, I thank my current and past graduate students and post-docs in the Pires Lab: Dr. Patrick Edger, Jacob Washburn, Sarah Unruh, Andrea Ravelo, Dr. Kate Hertweck, Dr. Erica Wheeler, Dr. Tatiana Arias-Garzon, Dr. Pamela (Roxi) Kellar, and Dr. Zhiong Xiong. In addition to science, they each have taught me a great deal about other aspects of life for which I am ever grateful.

I thank the wonderful undergraduate students in the Pires Lab: Michelle Tang, Woodson Smelser, Kevin Bird, Matthew Wiegman, Wade Dismukes and Madison Knapp. Each student has touched me in his or her unique way, and together they have taught me about my role as mentor and friend.

I thank other students, scientists, and staff that were essential to my growth and sanity during my work: Mitch Tucker, Caitlin Vore, Barb Sonderman, Debbie Allen, Dr. Mark Hannink, Dr. Kathy Newton, Dr. Abbie Ferierri, Clayton Coffman, Johanna Marrow, Dr. Diana Roberts, Dr. Danny Stark, Dane Lund, Kara Riggs, Dan Leuchtman, Preston Owen, Jonas Fernbach, additional Liscum Lab members Anthony, Scott, and Kyle, and Dr. Emmanuel (Mannie) Liscum.

I thank my close friends who have helped me through difficult times: Huan Truong, Avinash Karn, Chris Dove, Kyle Meinke, Ashley Winkler, Dr. Allision Flammang, Rachel Andreas, Erika Halsey, and Sarah Huber.

Thanks to the authors that marked this period of my life by demonstrating their wisdom, courage, and freedom: Don Dellilo, Jonathan Franzen, David Foster Wallace, Paula Fox, Thomas Pynchon, and Raymond Carver. I thank Italo Calvino for uncompromising creativity.

Lastly, I thank my parents, Dennis Mayfield and Debora Lintner. my stepmother, Anamika Anand, and spouse, Karen Mayfield-Jones, for their love and support, now and in the future.

TABLE OF CONTENTS

LIST OF FIGURES

OCCURRENCE AND IMPLICATIONS OF BIOLOGICAL NETWORK EVOLUTION FOLLOWING POLYPLOIDY

Dustin Mayfield-Jones

Dr. J. Chris Pires, Dissertation Supervisor

ABSTRACT

Recent advances in technology have allowed for unprecedented mapping of biological networks. Recovering and measuring these networks is essential for understanding the forces of evolution that shape and act on phenotypes. Here, we synthesize information regarding polyploidy, or whole genome duplication, and its effects on the rewiring of said networks. These changes may result in the phenotypic evolution thought to lead to the diversity and survival of some lineages of life. Specifically, we examine the consequences of polyploidy on an adaptive trait, flowering time, critical to the establishment and maintenance of plant speciation. Future studies of plant polyploids and of the changes of their complex biological networks will allow for the development of improved models of disease and biological processes useful for producing better crops for food, fuel, fiber, and pharmaceuticals. Discovery and characterization of such networks will also providing knowledge about the largely unknown constraints on the design space of life.

Introduction

Polyploid cells and organisms are classically defined as containing more than two homologous sets of chromosomes. Unlike diploids, that receive one set of homologous chromosomes from each parent, polyploid organisms inherit two or more sets. While this phenomenon occurs in many diverse eukaryotic organisms, it is particularly frequent in plants [1] and thought to be responsible for the success of some lineages. The duplication in genetic information and increase in combinatorial complexity is thought to allow for evolutionary change [2]. The redundancy in duplicated biological networks allows for novel changes in pathways while not drastically perturbing essential functions [3].

The success of polyploids over diploid progenitors could be due to their ability to mask deleterious mutations [4], stabilize heritable heterosis [5], and/or diversify function of duplicated genes via neofunctionalization or subfunctionalization [6].

Lineages can experience multiple rounds of polyploidy, each followed by the loss of many of the resulting duplicates. As reviewed in chapter one, following a polyploidy event, the genome often undergoes changes that return chromosomal pairing to a diploid-like state. However, the effects of polyploidy on genome architecture can be seen long after the return to diploid pairing during meiosis [7]. With the advent of genome sequencing of higher plants, it is now appreciated how all flowering plants share an ancient polyploidy event, and genes duplicated at this time were important in the development of the flower [8, 9]. While the possibility for novel traits existing as a consequence of polyploidy has long been appreciated,

biologists are now observing clues as to the fate of which genes may have contributed to the expression of novel traits.

The biological consequences of polyploidy can be observed at instantaneous, recent, and ancient time scales. Thus, these events are classified from recent to older events by the terms neopolyploidy, mesopolyploidy, and paleopolyploidy. In chapter one, we review mechanisms of formation of both polyploid cells and entirely polyploid individuals, the influence of polyploidy on gene loss and network rewiring, and how these changes affect long-lasting complexity and evolution of the organism. In chapter two, we then focus on the effect of polyploidy on a specific adaptive trait, flowering time.

Specifically, chapter two examines the role of polyploidy and the regulation of flowering time via epigenetic changes. The study of flowering time has resulted in a well characterized set of pathways in the model plant, Arabidopsis [10]. Comparisons with related species have elucidated interesting facts about the evolution of biological networks [11]. In the case of rice and Arabidopsis, aspects of the autonomous and photoperiod pathways are well conserved. However, a gene, *CONSTANS* (*CO*), that promotes flowering during long days in Arabidopsis does quite the opposite in rice, promoting flowering in short days and repressing flowering in long days [12]. Like in this example of *CO*, the analogy of nature reusing old tools from its toolbox in new ways may be very appropriate for thinking about both nature and the future engineering of biological systems. From both experimental observations and theory, the authors argue for new avenues of research including the study of both natural and resynthesized polyploids.

In addition to engineering life and creating exact models of cells to ecosystems, the goals of understanding biological systems involves discovering underlying design principles. Once identified, these principles would in turn reveal the answer to a major remaining question in systems biology: what are the general design constraints of life? In addition to constraints, remaining questions regarding the evolution of biological networks involve the rate and degree of network change [13]. To address this, measures of complex biological networks, such as robustness and connectivity, have recently been applied to biological systems to begin to measure these changes and have been reviewed elsewhere [14, 15]. The burgeoning field of systems biology and the study of biological network evolution will hopefully reveal answers to these deeper questions about the possibilities of life in addition to helping us engineer better fuel, food, fiber, and pharmaceuticals.

Together, these chapters synthesize our current understanding about the importance of polyploidy and possible roles of subsequent biological network evolution.

Figures

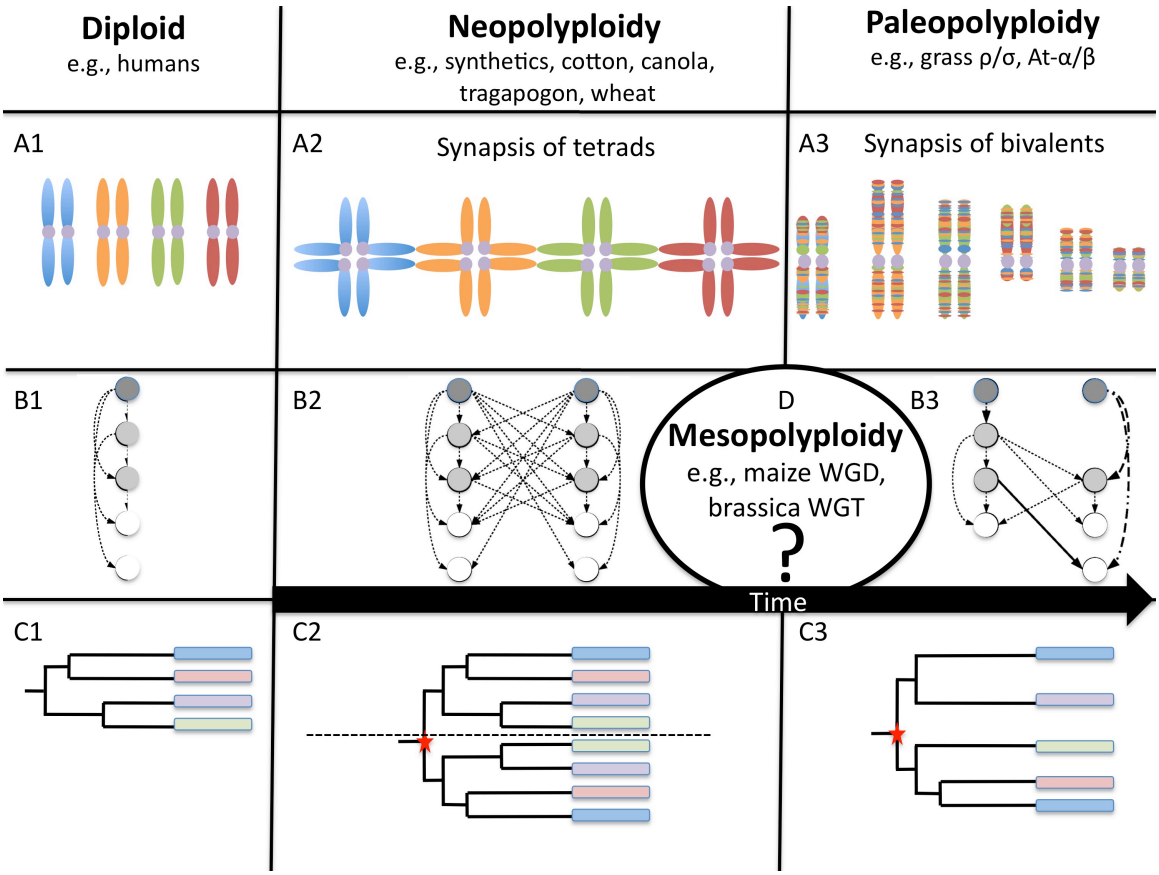


Figure 1. Chromosomal, network, and phyletic consequences of polyploidy through time with relevant organismal examples. (A) Chromosome pairs for diploid and polyploid organisms during meiosis. (A1) A hypothetical diploid, 2x with 8 chromosomes. (A2) A neopolyploid, 4x with 16 chromosomes. (A3) A paleopolyploid, 2x with 12 chromosomes. (B) Network evolution over time following a whole genome duplication event. (B1) Hypothetical network with general genetic elements represented by shaded circles and edges representing molecular interactions. (B2) A fully redundant network duplicated by whole genome duplication. (B3) Over time, the network undergoes loss of interactions and elements resulting in a rewired state. (C) Phyletic consequences of whole genome

duplication, represented by a red star. Each leaf represents a gene from a taxon. A different colored box represents each taxon. (C1) Lineage divergence within a group of diploid organisms following a typical bifurcating pattern. (C2) Speciation following a recent whole genome duplication, resulting in duplicates redundantly mirroring the species tree. (C3) Over time, gene loss contributes to the erosion of the mirrored species tree present immediately following whole genome duplication and initial speciation. (D) Mesopolyploidy marks the period of major network changes. While many of the mechanisms of these changes are hypothesized, the timing and lineage specific differences or similarities are largely unknown.

Literature Cited

1. Mable, B.K., *'Why polyploidy is rarer in animals than in plants': myths and mechanisms.* Biological Journal of the Linnean Society, 2004. **82**(4): p. 453-466.

2. Madlung, A. and J.F. Wendel, *Genetic and epigenetic aspects of polyploid evolution in plants.* Cytogenet Genome Res, 2013. **140**(2-4): p. 270-85.

3. De Smet, R. and Y. Van de Peer, *Redundancy and rewiring of genetic networks following genome-wide duplication events.* Curr Opin Plant Biol, 2012. **15**(2): p. 168-76.

4. Gu, Z., et al., *Role of duplicate genes in genetic robustness against null mutations.* Nature, 2003. **421**(6918): p. 63-6.

5. Birchler, J.A., et al., *Heterosis.* Plant Cell, 2010. **22**(7): p. 2105-12.

6. Moore, R.C. and M.D. Purugganan, *The evolutionary dynamics of plant duplicate genes.* Curr Opin Plant Biol, 2005. **8**(2): p. 122-8.

7. Simillion, C., et al., *The hidden duplication past of Arabidopsis thaliana.* Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13627-32.

8. Jiao, Y., et al., *Ancestral polyploidy in seed plants and angiosperms.* Nature, 2011. **473**(7345): p. 97-100.

9. Project, A.G., *The Amborella Genome and the Evolution of Flowering Plants.* Science, 2013. **342**(6165).

10. Pose, D., L. Yant, and M. Schmid, *The end of innocence: flowering networks explode in complexity.* Curr Opin Plant Biol, 2012. **15**(1): p. 45-50.

11. Higgins, J.A., P.C. Bailey, and D.A. Laurie, *Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses.* PLoS One, 2010. **5**(4): p. e10065.

12. Turck, F., F. Fornara, and G. Coupland, *Regulation and identity of florigen: FLOWERING LOCUS T moves center stage.* Annu Rev Plant Biol, 2008. **59**: p. 573-94.

13. Kitano, H., *Towards a theory of biological robustness.* Mol Syst Biol, 2007. **3**: p. 137.

14. Chae, L., et al., *Towards understanding how molecular networks evolve in plants.* Curr Opin Plant Biol, 2012. **15**(2): p. 177-84.

15. Barzel, B. and A.L. Barabasi, *Universality in network dynamics.* Nat Phys, 2013. **9**.

CHAPTER 1. Watching the grin fade: Tracing the effects of polyploidy on different evolutionary timescales

Authors: Dustin Mayfield-Jones, Jacob D. Washburn, Tatiana Arias, Patrick Edger, J. Chris Pires, and Gavin C. Conant

Published: Seminars in Cell and Developmental Biology, Vol. 24: 320-331, April 2013

Abstract

Polyploidy, or whole-genome duplication (WGD), is a recurrent mutation both in cell lineages and over evolutionary time. By globally changing the relationship between gene copy number and other cellular entities, it can induce dramatic changes at the cellular and phenotypic level. Perhaps surprisingly, then, the insights that these events can bring to understanding other cellular features are not as well appreciated as they could be. In this review, we draw on examples of polyploidy from animals, plants and yeast to explore how investigations of polyploid cells have improved our understanding of the cell cycle, biological network complexity, metabolic phenotypes and tumor biology. We argue that the study of polyploidy across organisms, cell types, and time scales serves not only as a window into basic

cell biology, but also as a basis for a predictive biology with applications ranging from crop improvement to treating cancer.

Introduction

Polyploidy and cellular biology

For a biologist interested in the mechanisms of cell function, it is tempting to think of polyploidy as something rare and probably occurring in someone else's system. Almost all model organisms have chromosomes that behave as diploids. If one studies the cell biology of humans, mouse, zebrafish, Drosophila, *Caenorhabditis elegans*, yeast, Paramecium, or Arabidopsis, one can easily convince oneself that whole genome duplications are unimportant as most polyploidy events are likely to be "evolutionary dead ends." However, like the Cheshire Cat's grin, the faint traces of whole genome duplications are to be found in most of these organisms. As more genomes are sequenced, we have found that not only is polyploidy rampant, but also that the apparent distinction between diploids and polyploids is eroding: phylogenomic results have now confirmed that many important lineages (e.g., all vertebrates and flowering plants) contain one or more rounds of whole genome duplications (WGDs) in their history. Over evolutionary time, the lingering effects of these WGDs have altered nearly every aspect of cell biology: from the immediate impacts of polyploidy (neopolyploidy) on chromosome pairing and epigenetic control of gene regulation (Section 3 of this review), to the process of diploidization (mesopolyploidy) and the fractionation of genomes (Section 4), to ancient whole genome duplications (paleopolyploidy) and the rewiring of metabolic and

regulatory networks (Section 5).

The evolutionary story of WGD, while compelling, is complex. But WGD can also be thought of as simply a mutational mechanism. Using natural or artificial polyploids as experimental systems has altered our understanding of and ability to investigate basic cellular functions, such as reproduction, metabolism, regulation and the relationship between cell volume and gene copy number. In this review, we outline several ways in which WGD has either illuminated particular processes or has served as a useful experimental system for probing such processes. Our examples will include how polyploidy illustrates the mechanisms of meiosis, how it may drive some of the complexity of regulatory networks seen in eukaryotes, how it may have produced some apparently paradoxical phenotypes (such as the propensity of bakers' yeast to ferment glucose) and speculative but tantalizing links between polyploidy, glucose fermentation and cancer. As we will show, some of the most interesting polyploids are those where only the grin remains: our aim is to bring the full WGD cat back into view.

Neo, meso or paleo? A timeline of WGD

A polyploidy event, strictly speaking, is simply a mutation in a single individual, but its results are often discussed at various timescales, sometimes without a clear distinction as to which one is at issue. Roughly speaking, we will discuss events occurring within a few generations of the mutation (neopolyploids), at an intermediate timescale, where speciation events have occurred since the WGD (mesopolyploids) and at long timescales, where larger-scale evolutionary

transitions may have occurred (paleopolyploids).

Who can be polyploid?

Variation in ploidy is not only common in plant and (some) animal lineages but also in certain cell lineages within an organism. Polyploid cells that aid in metabolic functions are an essential part of the developmental program in many organisms [1]. However, polyploidy is also associated with pathological conditions such as cancer and gall formation in plants [1], [2] and [3]. Here we explain what kinds of cells and taxonomic groups are likely to become polyploid and through which processes. We also discuss some of the physiological, morphological, and evolutionary implications of polyploid cell lineages.

WGD and cell lineages

Cells with roles in growth, metabolism, development and reproduction in mammals, insects and in plants are known to undergo polyploidy. Some of the most famous examples of polyploid cell lineages come from arthropods: e.g., salivary gland, gut, trachea, liver and kidney cells show ploidy changes throughout development [4] and [5]. In mammals, megakaryocytes (blood cells specialized for platelets production) may become polyploid, with a DNA content of up to 128N [6]. Later, these cells will fragment into circulating platelets [4]. Mammalian cells such as hepatocytes, arterial smooth muscle cells, and cardiac myocytes may also become

polyploid [6]. In plants, potentially polyploid cells include those in trichomes, galls, the leaf epidermis, roots, female gametophyte central cells [1N–14N [7]], callus tissue [8], endosperm (3N–15N), pollen, and hypocotyls [9] and [10]. Some of these lineages are obligate polyploids while others become polyploid due to external stimuli [6]. In the first case, polyploidy is a programmed part of development: e.g., megakaryocyte endoreduplication [4] or Drosophila polytene chromosomes [5]. In the second, one of the key stimuli that seems to induce polyploidy is stress, which induces WGD in uterine muscle during pregnancy, the thyroid during hyperthyroidism, and seminal vesicles with aging [4], [6] and [11]. Similarly, extreme temperature changes can induce polyploidy in plant cells [12] as can disease and telomere damage in animals [11] and [13] or a number of other stressors in the lab [14].

The reasons for these cell-lineage WGD events are varied, but some trends can be identified. In particular, cell volume and relative metabolic rate are clearly involved: topics we will return to in our discussion of the yeast WGD in Section 5. Thus, polyploidy tends to increase cell size and to be associated with high metabolic rates [4]. In plants, polyploid endosperm cells are known to increase metabolism and rates of gene transcription [4], [7] and [15]. Endoreduplication in animal embryonic cells can channel energy to proteins and cellular components needed during initial phases of development that would otherwise be used for cellular division. In addition to surface-area to volume relationships, polyploidy also alters the nuclear vs. organelle genome ratio, changing tissue development and morphology [4], [15] and [16]. Understanding these effects may be critical to

understanding cellular metabolism.

WGD in plant and animal species

Although polyploid organisms are known in both animals and plants, the general perception is often that they are too rare to have been a significant factor in animal evolution [17] and [18]. Nevertheless, recent polyploidy is observed across a vast range of animals, and is particularly common in fish and amphibians [19]. But it is with the blossoming of comparative genomics that ancient polyploidy events are now widely recognized to have played a significant role in the structure, content, and evolution of most eukaryotic genomes [18], [20], [21], [22] and [23]. As it happens, these WGD events have been best studied in plants, and the repeated discovery of ancient plant WGDs over the past 20 years has changed the question from how many plants are polyploid to how often a given lineage has experienced WGD [24] and [25]. From an evolutionary perspective, these events raise questions such as whether WGD provides the fodder for evolutionary novelty and whether it leads to changes in diversification and extinction rates [26], [27] and [28]. Such changes are expected, as WGD generates new genomic interactions starting with an initial "genomic shock" that must be resolved in the new lineage [see especially references in [29]]. Recent analyses have found that newly formed polyploid species have higher extinction rates than their diploid relatives. It is believed that this difference may be due to meiotic abnormalities and other negative fitness consequences of the inability to resolve said genomic shock. These data suggest a

model whereby most WGDs are evolutionary dead-ends, with only rare polyploids surviving over the long term. However, those rare survivors have left a substantial legacy in plant genomes, becoming some of the fundamental drivers of phenotypic novelty, resulting in genomic changes that may ultimately promote adaptive speciation and diversification [27], [28] and [30].

Neopolyploidy: the cell cycle, WGD and chromosome behavior

Mechanisms of polyploidy formation

Polyploidy may arise in several ways, generally through disruptions in either mitosis or meiosis. Polyploid cell lineages are formed most commonly via mitotic disruptions. Among these lineages, the most straight-forward route of formation is endoreduplication, where mitosis is bypassed under conditions of low cyclin-dependent kinase (CDK)/cyclin complexes [4] and [31]. Some lineages undergo several cycles of endoreduplication without entering mitosis [4], [9], [10] and [32]. A second mechanism of polyploid formation due to a disruption of mitosis involves the repression of telomerase. This repression leads to telomere shortening that results in end-to-end fusion of chromosomes [13]. Yet another type of polyploidy is seen in binucleated cells, which result from a failure of cytokinesis [2]. Cytokinesis failure [33] and/or failure to complete anaphase prior to reentering the cell cycle may cause polyploidy in mammalian megakaryocytes [6]. Finally, polyploid cell lineages can occur through the merging of cells or the engulfing of one cell by another [2] and [11]: for instance, transplanted bone marrow cells can

spontaneously adopt the phenotype of recipient cells through cell fusion [34].

While polyploid cell lineages have important physiological roles, researchers have historically been more interested in mutations that give rise to an entirely polyploid individual. Such mutations are generally the result of defects in meiosis. However, parthenogenesis and apomixes, where development occurs without fertilization, can also be considered a form of polyploidy [35], [36], [37] and [38]. In general, the meiotic process involves the production of unreduced gametes that are formed when the parent cell completes the first meiotic division but fails to enter meiosis II. The result is a dyad of two diploid cells rather than a tetrad of four haploid cells. Defects in cytokinesis can also lead to unreduced gametes [31]. The patterns of formation of these gametes and their subsequent evolution have been extensively reviewed [39] and [40]. The prevailing mechanisms of plant polyploid formation are somatic doubling in the meristem and 2n gamete formation [39].

The immediate impacts of polyploidy on a cell can be diverse. One common, though not universal, occurrence is an increase in cell size in concert with DNA content [22], [41], [42] and [43]. This occurrence sometimes referred to as the "nucleotypic effect" also causes volume and surface area scaling effects on the cell [41]. Some have posited that the smaller surface-to-volume ratios of polyploid cells may cause lower growth rates, but this occurrence is not universal [22]. Changes in cell volume associated with WGD may also cause changes in protein concentrations, thus affecting cellular kinetics [41]. Other observations include those of genetic, epigenetic, gene expression, and phenotypic changes (such as flowering time) in

recent allopolyploids [22], [44], [45], [46] and [47]. However, the fact that these changes are not seen in autopolyploids [48] may indicate that they are a function of hybridization more than of WGD itself [22].

Insights into the cell cycle from polyploid cells

Studies of various mitotic abnormalities that result from polyploidy have improved our understanding of mitosis and meiosis. In Drosophila, longer anaphase stages, chromosome bridges, delayed mitotic transits, and lagging chromosomes have been observed in polyploid cells [49]. Extra centrosomes also seem to result from polyploidy [50] and [51]. In this case, tetraploid animal cells generate four centrosomes which, if not properly positioned, can result in multipolar spindle formation and aneuploid daughter cells [52] and [53]. Although underutilized, these polyploid phenotypes may hold keys to further understanding of mitotic mechanisms.

Restoring functional meiosis after WGD: examples from wheat

Meiosis is problematic and occasionally fatal for polyploids because it requires the homologous pairing of chromosomes. When three or more very similar chromosomes exist within the cell, they tend to form multivalents, often resulting in incorrect segregation and unbalanced gametes. Because unbalanced gametes decrease reproductive fitness, polyploids have developed mechanisms for deterring

their formation. These mechanisms provide invaluable insights into diploid meiosis but would have been difficult to discover directly in those systems.

One example is from domestic wheat, *Triticum aestivum* L. Although this species is a hexaploid with three genomes, its chromosomes pair disomically, allowing it to function genetically as a diploid. Surprisingly, however, this differential pairing is not driven by differences between the three genomes, which are not diverged enough to give proper chromosomal pairing on their own [54] and [55]. Instead, several genes have been discovered which contribute to this process. The most important is named *Ph1* (pairing homoeologous): mutations in *Ph1* allow homoeologous chromosome pairing (Fig. 1) [56]. Only parts of the mechanism by which *Ph1* facilitates disomic pairing are known. For example, in the presence of *Ph1*, homoeologous chromosomes come together prior to meiosis for chromatin remodeling [57] and [58]. When *Ph1* is lacking, chromatin remodeling occurs asynchronously and non-homologous pairing occurs more frequently [57] and [58]. Given these results, some researchers proposed that *Ph1* suppresses crossing over and as a result prevents potential pairing between non-homologous chromosomes [59]. They introgressed, or transferred, chromosome 5B of wheat (containing *Ph1*) into autotetraploid rye (*Secale cereale* L.). This resulted in fewer non-homologous chromosome pairings when compared to normal rye. The same researchers also looked at chromosome pairing in rye when 2 copies of wheat *Ph1* were present. In this case, pairing (and crossing over) was restricted to the point that univalents and bivalents were formed more often than in controls [59] indicating that *Ph1* can even suppress homologous chromosome pairing. Logically,

one can assume that *Ph1*-like genes may be present in diploid organisms and be responsible for the degree of pairing and crossing over that occurs in these species. A similar recombination (and perhaps pairing) suppression gene exists in allotetraploid *Brassica napus* (AACC; 2n = 38) [60]. It has been named *PrBn* (Pairing regulator in *B. napus*) and its mutants display some similar phenotypes to *Ph1* mutants in wheat [61]. *PrBn* has been mapped to an area of ~10 cm but appears to function epistatically with other genes [62], making the dissection of its function difficult. The discovery of this second genetic system for meiotic pairing regulation presents an interesting opportunity for comparisons between the two systems. Due to the close phylogenetic relationship between *B. napus* and the model species Arabidopsis, many molecular tools are available for dissecting the *PrBn* system. Polyploids thus provide interesting windows into meiosis, a process contributing to both evolutionary change and various diseases [e.g., Down syndrome [63]]. The fact that a protein homolog of 5B2, the enzyme encoded by the *Ph1* locus, exists in humans (CDK2) [64], [65] and [66] makes the possible applications of polyploid meiosis research intriguing.

Mesopolyploidy: gene loss and gene networks

Once the immediate challenges of polyploidy have been resolved (e.g., restoring pairing, establishment of a breeding population), a longer-term process of genomic change begins. Over the last decade, plant researchers have gained new insights into both these processes through studies investigating either recent natural polyploids

(e.g. *Tragopogon mirus* and *Tragopogon miscellus*) or domesticated and resynthesized polyploids (e.g. *B. napus* – canola, *T. aestivum* – wheat, *Gossypium hirsutum* – cotton, *Nicotiana tabacum* – tobacco, *Arabidopsis suecica*). These studies have identified epigenetic and genomic changes that occur in the early generations following WGD, including alterations in DNA methylation and gene expression. These changes set the stage for arguably the main event of post-WGD evolution: the fading of the WGD grin through duplicate gene loss and chromosomal rearrangements [45], [47], [67] and [68].

Diploidization

Collectively, this entire "stabilizing" process has been termed diploidization (i.e. transitioning from the polyploid back to the more stable diploid state). An organism that has been "caught in the act" of this process is autotetraploid rye (*S. cereale* L.). Although it still forms multivalents during meiosis, it displays only ~60% of those expected of a randomly pairing autotetraploid [59]. It thus illustrates the importance of understanding the various mutations and underlying mechanisms that contribute to diploidization [69]. Again, the two key processes of duplicate loss and genome rearrangement are at play: working together, they reduce the similarity of homoeologous chromosomes, allowing the restoration of meiotic stability. We can visualize the process of genome rearrangement with comparative cytogenetic data. These changes are due to various large-scale mechanisms of karyotype evolution, especially chromosome fusion and fission events [70]. For example, *Arabidopsis*

thaliana has five haploid chromosomes while its ancestor is inferred to have had eight [70]. Based on both cytogenetic approaches and comparing genetic maps, it is clear that several of the ancestral chromosomes fused in order to reach the current Arabidopsis karyotype.

Fractionation

Of equal importance for diploidization is the loss of the duplicate genes produced by the WGD, a process termed fractionation. From each pair of duplicates, it is generally possible to lose one copy through genetic drift, since the second copy retains the required function (although see below). Recent evidence suggests that these losses occur via a series of short deletions [71]. The most naïve expectation is that this process would randomly eliminate duplicates (unbiased fractionation), leaving essentially half of each of two copies of the pre-WGD genome (see Sections 4.4 and 4.6). One of the more striking results of two decades of polyploidy research is the discovery that the trajectory of fractionation is often nonrandom, sometimes in two distinct ways: biased fractionation across the genome and the over-retention of dosage sensitive genes.

Biased fractionation

If the polyploid was formed by the combination of two distinct genomes (e.g., an allopolyploid rather than an autopolyploid), losses are often more common in one of the two contributing genomes than in the other. This pattern of biased fractionation has been reported in both monocot and eudicot genomes following independent

polyploid events [72], [73], [74] and [75]. Similarly, the noncoding regulatory sequences are also fractionated over time across homoeologous regions in plant genomes [Fig. 2A; [74] and [76]]. For example, *Brassica rapa* (Chinese cabbage) underwent a whole genome triplication event (i.e. a process involving a first WGD, a hybridization with a diploid, and then a second WGD) that is not shared with its relative, the model plant *A. thaliana*. An analysis of the *B. rapa* genome revealed a dominant subgenome that has undergone the least fractionation (i.e., has the highest duplicate gene retention rate). The other two contributing subgenomes have undergone more fractionation and are termed, in order of the number of retained duplicates, the highly- and most-fractionated subgenomes, respectively (Fig. 2A). Thus, gene retention has been biased toward one subgenome (i.e. one progenitor parental genome or 'dominant' subgenome), with duplicates having been largely lost from the two other subgenomes. The mechanism(s) underlying this genome-wide pattern is unknown, although researchers have hypothesized that they involve epigenetic phenomena and repetitive elements [76]. Hence, it was recently proposed that studying the fractionation of known cis-acting sequences in plant genomes would provide an opportunity to determine the function of individual conserved noncoding sequences [termed "Nature's Promoter Basher [76]].

Unbiased fractionation and reciprocal gene loss

Although biased fractionation has been the more frequent outcome of allopolyploidy [75], [77] and [78], it is important to note that preferential retention of one sub-

genome is not the invariable result of WGD: teleost and yeast lineages generally show more random and equal loss patterns [79] and [80]. One potential reason for this difference would be if the later two events had been true autopolyploidies, with identical genomes coming together. Unfortunately, current methods are not yet sensitive to clearly distinguish an ancient autopolyploidy from an ancient allopolyploidy.

Gene Balance Hypothesis: retention of dosage sensitive genes

After biased fractionation, the second type of bias seen in duplicate gene retention and loss patterns is more general and can occur even in cases of unbiased fractionation such as in yeast (described below). Even in those cases, gene retention following WGDs is often biased toward dosage-sensitive genes, which includes highly connected transcription factors, kinases and ribosomal proteins. These same dosage sensitive genes, which are over-retained in duplicate after WGD, are significantly under-represented among small-scale duplications [21]. Reciprocally, dosage-insensitive genes, which are most frequently duplicated via smaller-scale duplications, tend to be lost following WGDs [21]. These gene retention and loss biases are most easily explained by events occurring at the protein level and are likely due to a need to maintain proper stoichiometric balance in networks and macromolecular complexes [81] and [82]. For example, previous studies have demonstrated that the over-expression of a single dosage-sensitive subunit will impact the assembly kinetics of a macromolecular complex, which will result in a

decrease in the number of fully assembled complexes and an increase in intermediate products and unassembled subunits, which are often toxic [73] and [74]. The reduction in fully assembled complexes may also result in a network imbalance (i.e. imbalance with direct interacting proteins). For example, it is vital to maintain proper balance between an activator and repressor acting on a common target following WGD (e.g. a kinase and a phosphatase).

Polyploid events simultaneously increase the dosage of all genes; thus proper balance is retained in the absence of gene loss. As the process of gene loss sets in, duplicates of interacting dosage-sensitive genes must be co-retained to maintain stoichiometric balance in pathways (signaling and regulatory) and in large macromolecular complexes. This balance can be critical for normal function and development. On the other hand, smaller scale duplications of certain dosage-sensitive genes (i.e. highly connected signaling proteins) would result in a stoichiometric imbalance similar to gene losses following a polyploid event. Thus, these same dosage-sensitive genes are under-represented among smaller-scale duplications and seldom exhibit copy number variation.

These data and the associated model are the basis of the Gene Balance Hypothesis [83], which argues that groups of genes that function together are dosage sensitive because a change in expression of one of them has the potential to break the subtle balance required for them to properly function together. This hypothesis is supported by analyses of various eukaryotic genomes, including yeast, Drosophila, Paramecium, and flowering plants [84], [85] and [86]. More recently,

our analysis of the entire Arabidopsis primary metabolic network [87] and certain secondary metabolic pathways [88] have also supported specific predictions of the hypothesis: in particular we observed over-retention of duplicates of high connectivity from the metabolic network. In addition, evidence revealed that a small subset of these duplicates were also retained due to selection for increased gene product [87].

Fractionation in yeast

As an example of the overall process of fractionation, the polyploidy yeasts (including the model eukaryote *Saccharomyces cerevisiae*) are exceptional due to their small, well-studied genomes and the large quantities of genomic data available. That a WGD had occurred in the lineage leading to *S. cerevisiae* was argued shortly after the genome sequence was released [89]. This hypothesis was confirmed with the sequencing of several yeast species that split from the bakers' yeast lineage prior to the WGD [90], [91] and [92]. The pattern of post-WGD evolution in yeast is particularly rewarding to study given the existence of a well-curated genome browser that lays out the patterns of paralog loss post-WGD [Fig. 2B, the Yeast Gene Order Browser [93]]. What is most striking when comparing these genomes is the pattern of interleaved blocks of genes from pairs of *S. cerevisiae* chromosomes relative to their homologous single copy genes in species such as *Lachancea waltii* or *Eremothecium gossypii* (Fig. 2B). These doubly conserved synteny blocks cover 90% of the genome in *L. waltii* [91] and 96% of that in *E. gossypii* [90], giving clear,

visual, evidence of WGD followed by fractionation.

The Gene Balance Hypothesis and yeast

Examining the pattern of losses in the post-WGD yeasts, it is clear that several of the predictions of the balance hypothesis are borne out. Thus several classes of genes were over-retained after the WGD, including ribosomal proteins, protein kinases and transcription factors [94] and [95]. Similarly, genes that tend to have been fixed by WGD are less likely to have undergone single-gene duplications in other yeast species [96]. However, WGD-duplicates produced by genome duplication have more protein interactions [97] and [98], more phosphorylation sites [99] and tend to be highly expressed [94]. In interpreting these results, a certain caution is in order because network patterns (such as interactions and phosphorylation sites) are measured in the same organism that underwent the WGD. As a result, it is possible that, rather than reflecting retention patterns, these associations are indications of post-WGD neofunctionalization among the preserved duplicates. In other words, WGD may have allowed the accumulation of new interactions in the duplicates it produced. Such interactions would then be implicated in new functions resulting from WGD. While such functional innovation has certainly occurred (see below), we find it an implausible explanation for the general trend for two reasons. First, it would require a very high proportion of WGD-produced duplicates to have undergone the rare process of neofunctionalization, Second, while early work on network evolution found that protein interaction and regulatory networks rapidly

lose duplicate interactions [100] and [101], more recent analyses have found that the rate of appearance of novel interactions is much lower [102]. Thus, a more parsimonious explanation is that high-resolution functional data from yeast illustrate many of the Gene Balance Hypothesis' predictions.

The implications of the Gene Balance Hypothesis are many. In lineages with multiple WGDs in their history, it is dangerous to speculate as to divergence in function for large duplicated families of transcription factors, given that such families are predicted to accumulate simply through duplicate retention to preserve balance. However, we emphasize that maintaining dosage balance is a general trend, not a universal requirement. Similarly, one might understand how maintaining dosage balance is important for conserving pathway functionality and multi-component machineries and still wonder how this conservation achieves the phenotypic variation and functional innovation thought to be the result of WGD.

We argue that the key to this apparent paradox lies in the environment in which the new polyploidy finds itself. In an unchanging external environment, expression levels among genes are probably tuned to functional needs and balance requirements [103] and [104]. Under these circumstances, duplicate retention may be favored among interacting proteins. However, as we discuss below in the case of the yeast WGD, if the selective conditions have changed, altering dosage balance may actually be beneficial. Similarly, retention for balance is likely a transitional state after WGD: in both the *A. thaliana* metabolic network and among the yeast ribosomal proteins, we have found that selection to maintain dosage balance

eventually relaxes post-WGD, allowing functional divergence of the associated duplicates [87] and [105].

More generally, fractionation can also be misleading in the search for orthologous genes, since apparently single copy genes in two genomes sharing a WGD can easily be paralogs [106]. The reason is that reciprocal loss of paralogs long after the WGD gives rise to situations such as that seen in Fig. 2B, where genes such as *NUS1* in *S. cerevisiae* and *2000.17* in *Kluyueromyces polysporus* are single copy and homologous in the two genomes, but share a common ancestor at the WGD (e.g., are paralogs) and not at the more recent speciation of *S. cerevisiae* and *K. polysporus* (in which case they would have been orthologs). Whether this fact is of functional, rather than merely evolutionary, importance is still being investigated.

Paleopolyploidy: evolution, complexity and WGD

Polyploidy also has the potential to induce long-lasting changes at levels of complexity above that of the linear genome. For instance, there is evidence that WGD dramatically altered the developmental regulatory network of vertebrates [107] and [108]. Because WGD duplicates many genes at the same time, it creates the potential for correlated, multi-gene, alterations in cellular networks [109]. One of the most carefully worked out of these examples concerns a well-known, and yet paradoxical, characteristic of the laboratory yeast *S. cerevisiae*, namely the Crabtree effect. This effect describes this yeast's unusual metabolic preference for partially oxidizing glucose into ethanol even when oxygen is available. Doing so at least

initially forgoes much of the energy that would be realized if the glucose was instead fully converted into CO_2 and water through the TCA cycle [110] and [111]. Intriguingly, there is an apparent association between a yeast species' possession of the ancient WGD and the Crabtree effect [112]. The case for this association is strengthened by the functional differentiation of a group of duplicated genes from the WGD, all involved in glucose metabolism (Fig. 3A). These gene pairs, two glucose sensors (*SNF3* and *RGT2*), two glucose transporters (*HXT6/HXT1*) and two duplicate enzymes that catalyze the initial step of glycolysis, have specialized such that one member of the pair acts at low glucose concentrations and the other at high ones [113].

Given these observations, we and others have proposed that the yeast WGD initiated a change in the patterns of glucose metabolism in yeast. This idea was based on a proposed sequence of linked events. First, the increase in gene copy-number from the WGD gave rise (after some gene losses in other parts of the genome) to an increased flux through glycolysis [112], [114] and [115]. Second, because the rate of pyruvate use by the mitochondrion is constrained by oxygen concentrations and surface area to volume rules, the WGD-possessing cells redirected some of this increased glycolytic flux to the (previously anaerobic) fermentative pathways [115]. [Artificial yeast polyploids are larger than diploid cells, and as a result differ in their surface area to volume relationships [116].] Thus, the end result was to set the new polyploidy yeast on a path toward increased Crabtree effect.

Because several post-WGD yeast genomes have been sequenced, it is possible to follow these events in some detail. We can, at each node in the phylogeny at the bottom of Fig. 3B, infer the status of every WGD-created duplicate pair [e.g., whether it remains duplicated or has been converted to single copy [117]]. As expected, the fraction of duplicated genes declines over time (blue line, right axis), with approximately 10% of the genes surviving in duplicate in the extant S. cerevisiae genome. We can then use the complete yeast metabolic network to predict the flux through each enzyme under a variety of nutrient conditions [118]. For each time point (blue dots in the phylogeny of Fig. 3B), we can calculate the average flux both of the enzymes coded for by genes retained in duplicate (right axis, dark red line) and that of those returned to single copy (pink). Over time, the number of surviving duplicates falls, but those duplicates are increasingly enriched for enzymes of high flux reactions. In contrast, the remainder of the enzymes is generally of lower flux and is returned to single copy. Thus, the pattern of duplicate retention across the genome and through time supports the notion of a metabolic adaptation driving the resolution of the yeast WGD. This premise was also supported by an independent analysis by van Hoek and Hogeweg [119], who were able to show computationally that similar WGD events modeled in modern *S. cerevisiae* could also be expected to result in over-retention of glycolytic enzymes and increased glycolytic flux.

These ideas do not address the apparent difficulty of why evolution would deliberately "choose" to make metabolism less efficient. The reason they do not is surprisingly simple: increased metabolic efficiency does not always increase fitness. Theoretical work on resource competition among organisms inhabiting a large but

ephemeral environmental resource shows that such competition among cells can actually favor "wasteful" lineages if their more efficient competitors are also slower-growing [120], [121] and [122]. The phenomenon is referred to as the "tragedy of the commons" [123] and occurs when the efficient cells are able to convert more glucose into energy but pay for this efficiency in lower growth rates, meaning that the fast, wasteful, cells come to dominate the local environment.

If the WGD in fact was a trigger that allowed *S. cerevisiae* to move toward an increasing Crabtree effect, it most likely would have been followed by later, reinforcing changes. And at least two such changes are known. First, in WGD yeasts, there was a loss of cis-regulatory elements among the genes for the mitochondrial ribosomal proteins that has effectively de-coupled the expression of the cytosolic and mitochondrial ribosomal proteins [124]. This decoupling allows *S. cerevisiae* to up-regulate production of cytosolic ribosomes independently of mitochondrial ones, increasing fermentative efficiency by avoiding unnecessary ribosome synthesis in the quiescent mitochondria. The second example is a post-WGD single gene duplication of an alcohol dehydrogenase. The result of this event was specialized ADH loci, one for ethanol synthesis and a second one responsible for the back-conversion of ethanol to pyruvate [once glucose is exhausted Crabtree yeasts can re-import and respire the ethanol they previously produced [125]]. Such later evolutionary refinements provide further evidence for a WGD-produced shift in metabolism and remind us again how an evolutionary perspective can greatly enlighten our understanding of complex biological processes.

Polyploidy and cancer

Surprisingly, evolutionary histories of WGD can also shed light on more immediate events: for instance cancer progression. That tetraploidy and particularly aneuploidy occurring in tumor lineages has long been known. However, they were often thought to be only incidental phenomena in the tumors' evolution [126]. This minor role for WGD in cancer was challenged nearly 25 years ago with a conceptual model for the role of polyploidy in cancers that suggested that tetraploidy represents an intermediate stage during carcinogenesis of many human solid tumors [127]. More recent evidence supports a pattern of polyploidy-to-aneuploidy transition in at least some cancers. Generally, similar chromosomal observations regarding mitotic pairing and genomic instability have been made of cells following a polyploidy event in both normal and pathological states [11]. Large-scale chromosomal amplifications, deletions, inversions, and translocations have been detected in cancers [128] and [129]; similar chromosomal behavior has been observed in both recent natural and resynthesized polyploid plants [Fig. 1; [130] and [131]]. In addition to similar early effects on chromosome behavior, most aneuploid tumors are indeed found to have genomes in the triploid to tetraploid range [132]. Most convincingly, actual tetraploid intermediates have been found in both murine and human carcinogenesis (e.g., human cervical carcinoma and Barrett's esophagus) [133], [134], [135], [136] and [137]. Others have encouraged research on manipulating ploidy in animal models of cancer to directly test whether

tetraploidy can lead to aneuploidy similar to those observed in many cancers [11]. While evidence contradicts the requirement for a polyploid-to-aneuploid transition in some cancers [138], many of the same biological implications regarding cellular adaptation apply to both conditions where multi-gene duplications potentially result in the genetic novelty that is the fodder for selection [139].

Because of the complexity of cancer and the difficulty of distinguishing causal from merely associative events, it is still an open question the degree to which polyploidy causes cancer phenotypes. This difficulty is compounded by the fact that "causal" is used both in an immediate molecular sense (absence of protein X causes uncontrolled division) and in a more remote process sense (selection is favoring cell lineages with faster growth properties). For the purposes of understanding polyploidy and cancer, we strongly believe that this later view of causality is more illuminating [140]. This view is complicated by several factors, unfortunately: mutation patterns and rates in cancer cells are not well understood [141], most cancers are either removed or treated immediately following detection, and cancers are not homogeneous but rather the product of local adaptation to microenvironments [133], [142], [143] and [144]. The inability to assess the evolution of these cancers prohibits study of the effects of ploidy changes on the development and maintenance of cancer cells in vivo. However, some slow developing premalignant neoplasms have become models for understanding the temporal events of neoplastic progression in solid tumors including the Barrett's esophagus just mentioned. With the advent of single cell sequencing, studying the landscape of affected cells will hopefully provide insight into how cancers evolve

and adapt across heterogeneous environments [145].

Polyploidy and cancer: insights from the Warburg and Crabtree effects

In the meantime, there are intriguing suggestions of a link between polyploidy and a well known but poorly understood cancer phenotype: the Warburg effect. This effect refers to the prevalence of glucose fermentation over glucose respiration in cancer cells and is often explained by reference to the hypoxic environment that many tumors experience. In such circumstances, energy production by the citric acid cycle is infeasible. However, there is some evidence that the Warburg effect is observed in circumstances where lack of oxygen is not a factor [146]. Instead, one part of the Warburg effect seems to stem from a similar type of glucose repression of oxidative phosphorylation as is seen in the yeast Crabtree effect [147] and [148]. Thus, while the Warburg effect is no doubt complex and due to a variety of factors, it is possible that the same type of selection for fast, inefficient growth is acting to favor its appearance [120] and [121]. Polyploidy or aneuploidy followed by gene losses could alter relative enzyme dosages so as to favor glycolysis, as we have suggested occurred in yeasts. Our new ability to easily resequence tumor genomes [149] provides an attractive method for testing this hypothesis: the difficulty with this approach lies in accounting for the changes in gene expression that could similarly drive a Warburg phenotype. If the hypothesis is in fact confirmed, it would be striking that the same type of mutations (WGD followed by duplicate loss) has been employed in both systems in order to achieve the same end: rapid cell division

in environments where resources are effectively unlimited.

Conclusions

Polyploidy has revealed several interesting facets of cell biology. Some of them involve direct responses to polyploidy in certain cells (e.g., the yeast WGD and ethanol fermentation). More often, polyploidy incidentally illuminates other aspects of biology, such as the nature of chromosome pairing in meiosis or the structure of duplicated genes in cellular networks. On this latter point, the Gene Balance Hypothesis allows for a relaxation of epistatic constraints between genes through the process of post-WGD subfunctionalization [150]: such processes may be another slow and mysterious part of the influence of WGD on extant organisms. But there is also an even more general principle that polyploidy illustrates: the necessity of seeking an integrative understanding of the machinery of the cell. The global changes that result from polyploidy remind us of the interconnected nature of biology, a connectedness that even includes cats from Cheshire [107].

Figures

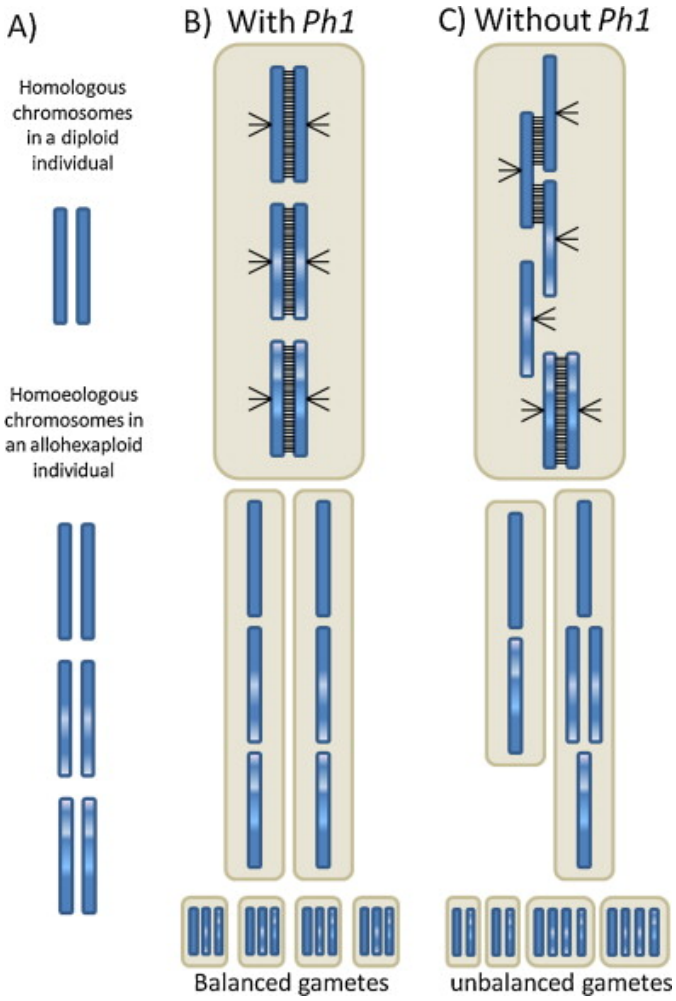


Figure 1. Simplified illustration of meiosis I in allohexaploid wheat with and without the *Ph1* gene. The figure shows only one chromosome set (three chromosome pairs) and does not show crossing over or meiosis II. (A) An example of homologous chromosomes in a diploid individual and those same chromosomes in a hexaploid individual after divergence. (B) In the presence of the *Ph1* gene, homologous chromosomes pair to the exclusion of their homoeologs in diploid-like fashion and eventually form balanced gametes. (C) In the absence of *Ph1*, diploid-like pairing breaks down and many pairing combinations may result (only one

hypothetical combination is shown here) possibly leading to unbalanced gametes and infertility.

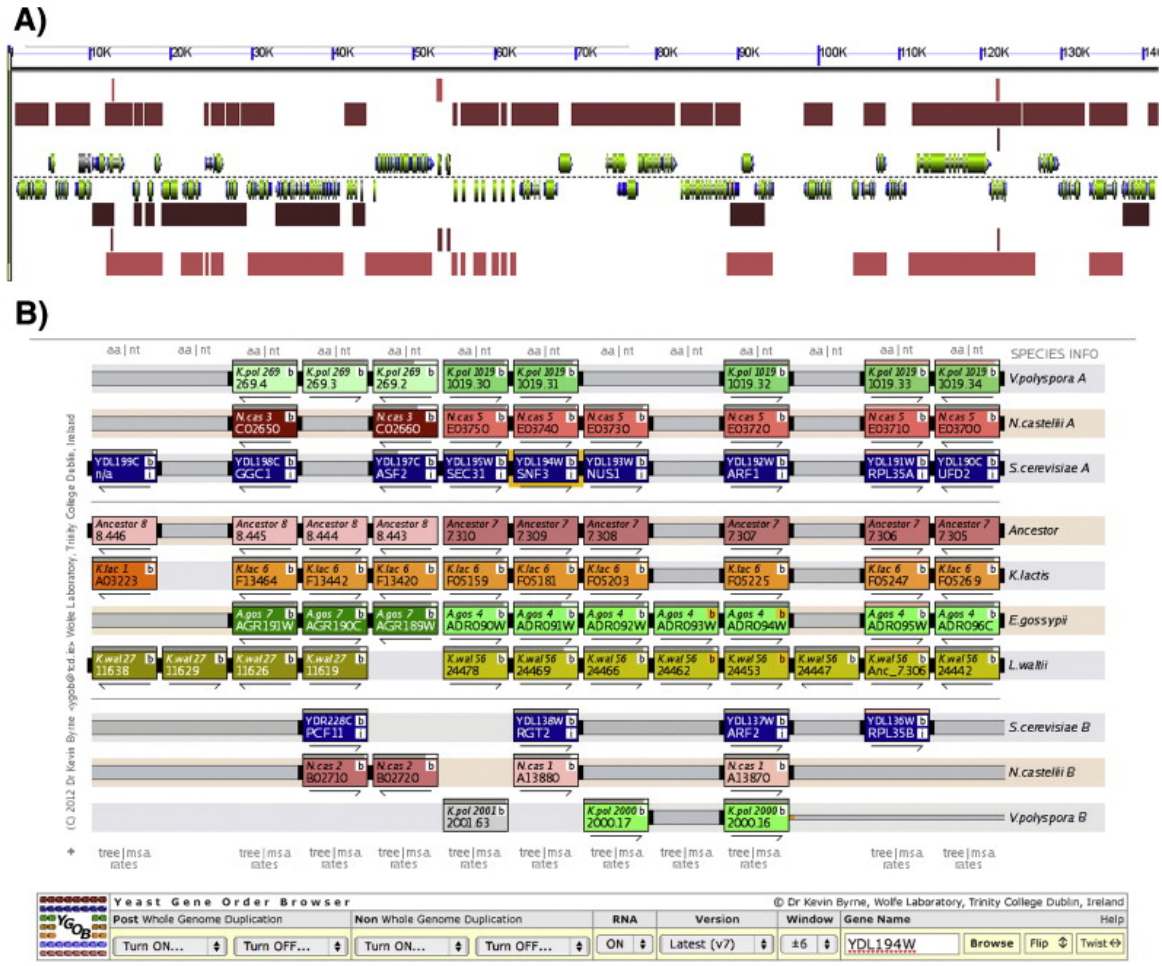


Figure 2. Whole genome duplication and fractionation patterns in Brassica and yeast. (A) The Brassica whole genome triplication event and fractionation patterns are clearly observed by comparing orthologous regions between *Arabidopsis thaliana* and *Brassica rapa*. This figure was constructed using CoGe (genomevolution.org). A 140 Kb genomic region centered on locus *AT3G01390* is shown in the middle, with gene models (green are exons and blue are introns) on the +/− strands. The three orthologous regions in *Brassica rapa* are shown above (light brown – 'dominant sub-genome') and below (dark brown – 'most fractionated sub-genome'; pink – 'highly fractionated sub-genome') the Arabidopsis region. The dominant sub-genome is consistently the least fractionated homoeologous region in

Brassica rapa (i.e. showing the highest rate of gene retention). Sequence conservation of each Brassica region against the orthologous Arabidopsis region is depicted by the presence of a colored block. (B) Little evidence of biased fractionation in yeast. Shown is a screen shot from the Yeast Genome Order Browser [YGOB [93]]: in the center are gene orders for three yeasts lacking the WGD (*K. latics, E. gossypii, and L. waltii*) and the presumed ancestor of the WGD yeasts [151]. Above and below are the corresponding pairs of chromosomes from three yeasts with the WGD (from inner to outer *S. cerevisiae, N. castellii and K. polysporus*). The interleaved pattern among the post-WGD yeasts relative to the others is clear. The region of the genome shown corresponds to the *SNF3/RTG2* gene pair discussed in the text.

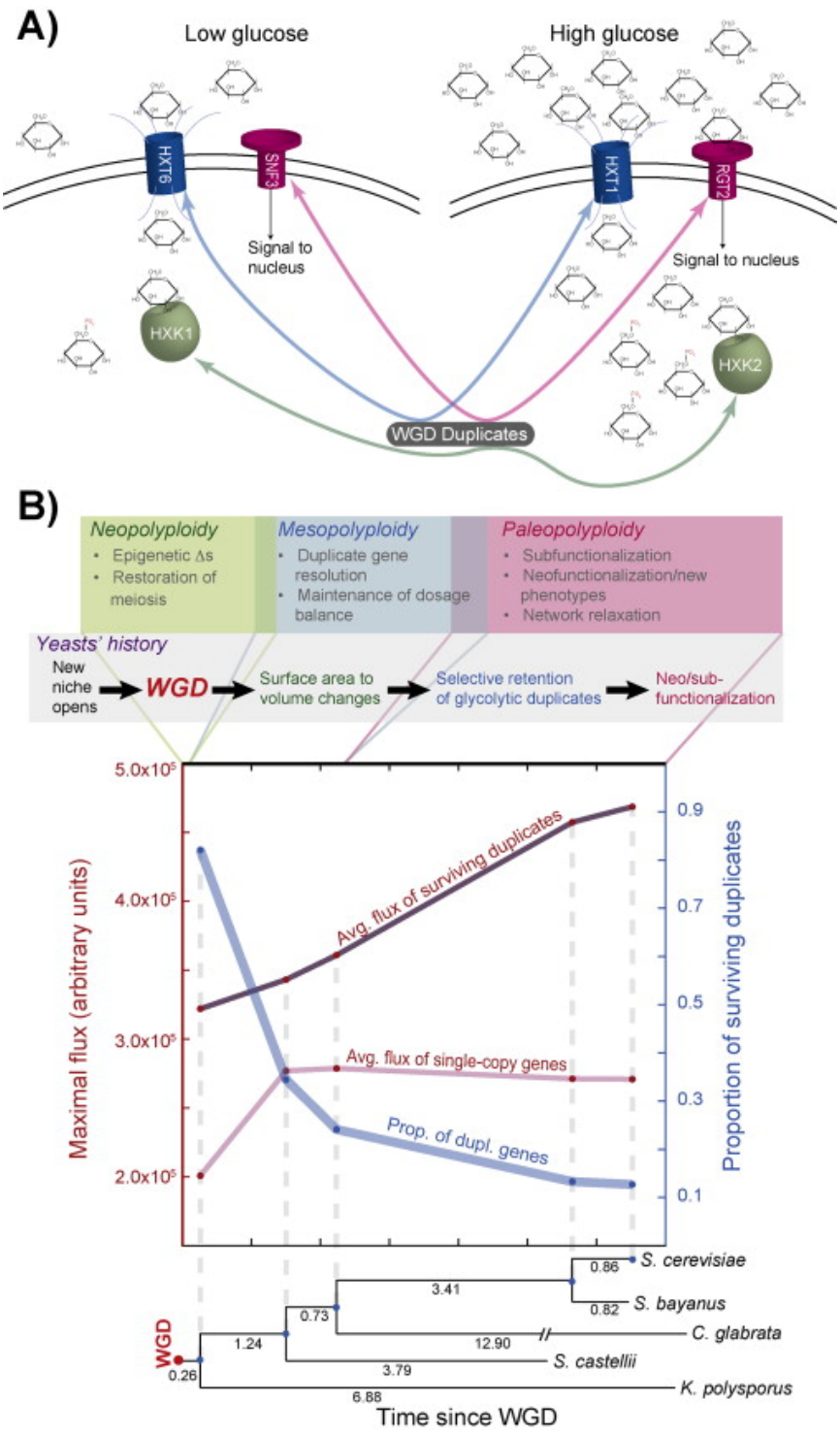


Figure 3. New evolutionary trajectories opened by the yeast WGD. (A) Subspecialization of genes for glucose metabolism in yeast after the WGD. Shown are a pair of duplicated glucose sensors (*SNF3/RGT2*), a pair of duplicated glucose

transporters (*HXT6/HXT1*) and a pair of duplicated enzymes for the first step of glycolysis (*HXK1/HXK2*). In all three cases, the first member of the pair is active at low glucose concentrations and the second at higher ones (see main text). (B) Timing of events after WGD. In the boxes are examples of the types of event occurring at each stage of WGD resolution. (Obviously, there are no fixed boundaries between the timeframes and we have illustrated this fact with overlapping boxes). We show a speculative sequence of events leading from the yeast WGD to the modern bakers' yeast as an example of these phenomena. What is clear is that metabolism has played a key role in the evolution of the yeast WGD. To illustrate this point, we have inferred whether each yeast WGD-produced duplicate gene pair was still duplicated at five points in the phylogeny of the post-WGD yeasts (blue dots; bottom). To do so, we used our existing likelihood model of WGD resolution [117]: branch lengths indicate the relative rate of gene loss and duplication fixation along that branch. The result is the curve of the proportion of surviving duplicates at each point (blue line, left axis). We then computed the maximal flux through each yeast metabolic reaction under a variety of conditions [152]. We then took the average of the flux of the genes likely to have been present in duplicate and of those likely to have been in single copy at the five points mentioned. As we had seen previously [153], even immediately after WGD, there was a bias between the enzymes carrying low flux, which are relatively likely to have been returned to single-copy and those of higher flux, which tended to remain duplicated. Here we extend this result and now see that, as duplicates are lost over time, there is an increasing distinction between the retained duplicates (right axis, dark red) which

are over-represented among the high flux enzymes, and the remaining single-copy genes (pink).

Boxes

Allopolyploid – having multiple chromosome sets derived from different species as a result of WGD.

Autopolyploid – having multiple chromosome sets derived from a single genome or single species as a result of WGD.

(We note that the meanings of auto- and allo-polyploid have changed with advances in technology. Early in the study of polyploidy, the terms referred to the cytological modes of inheritance, where auto- and allo-polyploids exhibited polysomic and disomic inheritance, respectively. With the advent of genome sequencing and improved cytological methods, auto- and allo-polyploids have come to refer to the nature of the two progenitor genomes.)

Crabtree effect – the observation that some species of yeast use the less energetically efficient partial oxidization of glucose into ethanol in the presence of oxygen rather than its full conversion to CO_2 and water through the TCA cycle.

Diploidization – the process of transition from a polysomic state (with incomplete chromosome pairing) back to a more stable diploid-like state with homologous chromosome pairing in meiosis.

Disomic – chromosomes faithfully pair with their homolog forming bivalents during meiosis.

Fractionation – the process of post-WGD gene loss through nonfunctionalizing mutations and deletions. If one of the two parental genomes that served to form the

polyploidy population experiences more losses, the processes is said to be biased. Otherwise, it is termed unbiased fractionation.

Genome Balance Hypothesis – this hypothesis argues for the existence of reasonably strong purifying selection acting on changes in relative gene dosage between genes whose products interact physically or in regulatory cascades. A key mechanism thought to drive this selection is the formation of non-functional or even toxic complexes when interaction balance is lost. The hypothesis thus predicts that dosage changes, like single gene duplications, will be disfavored if they possess such interactions. Complementarily, such interacting genes should be preserved in duplicate after polyploidy, because the loss of one member will also introduce dosage imbalances.

Homoeolog – homologous genes, genomic regions, or entire chromosomes in a polyploidy genome that derive from the two diploid parental genomes.

Homolog – two or more genes, genomic regions, or entire chromosomes that share descent from a common ancestral DNA sequence.

Introgression – the transfer of genomic elements from one species to another as a result of their hybridization and repeated backcrossing.

Polyploidy event/whole genome duplication (WGD) – an event that results in the formation of a new organism with an extra complete copy of its genome. Section 2 gives an overview of the mechanisms leading to WGD formation.

Polysomic – chromosomes that are completely homologous and segregate to form

multivalents or random bivalents in meiosis.

Reciprocal gene loss – WGD followed by speciation and fractionation can lead to a situation where species sharing the WGD have lost alternative copies of certain genes through independent fractionation. It has been argued that reciprocal gene loss may have contributed to reproductive isolation among post-WGD yeasts because hybrids would produce a high proportion of non-viable offspring due to double-null loci [79].

Segmental allopolyploids – refers to polyploids that exhibit both polysomic and disomic modes of inheritance.

Smaller-scale duplication – the duplication of a small part of a genome, often that of a single gene. Small-scale duplications are often contrasted to larger scale events such as duplications of whole chromosome arms or even genomes (WGD).

Subgenome – one of the contributing parental genomes in an allopolyploid.

Warburg effect – the observation that, unlike most cells that prefer to use glucose respiratorily when oxygen is present, cancer cells often use fermentation to metabolize that glucose (see Crabtree effect).

Box 1: Definitions

Acknowledgments

We would like to thank C. Hudson, and E. Puckett for helpful discussions. DM, TA, PPE and JCP are supported by NSF grants NPGI 1202793, DBI 1110443, DEB 1146603, and DEB 1209137. GCC is supported by the Reproductive Biology Group of the Food for the 21st Century program at the University of Missouri.

Literature Cited

[1] Mosieniak G, Sikora E. Polyploidy: the link between senescence and cancer. Current Pharmaceutical Design 2010;16:734–40.

[2] Krajcovic M, Overholtzer M. Mechanisms of ploidy increase in human cancers: a new role for cell cannibalism. Cancer Research 2012;72:1596–601.

[3] Duensing A, Duensing S. Centrosomes, polyploidy and cancer. Advances in Experimental Medicine and Biology 2010;676:93–103.

[4] Edgar BA, Orr-Weaver TL. Endoreplication cell cycles: more for less. Cell 2001;105:297–306.

[5] Smith AV, Orr-Weaver TL. The regulation of the cell cycle dur- ing Drosophila embryogenesis: the transition to polyteny. Development 1991;112:997–1008.

[6] Zimmet J, Ravid K. Polyploidy: occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. Experimental Hematology 2000;28:3–16.

[7] Friedman WE, Madrid EN, Williams JH. Origin of the fittest and survival of the fittest: relating female gametophyte development to endosperm genetics. International Journal of Plant Sciences 2008;169:79–92.

[8] Fras A, Juchimiuk J, Siwinska D, Maluszynska J. Cytological events in explants of Arabidopsis thaliana during early callogenesis. Plant Cell Reports 2007;26:1933–9.

[9] Melaragno JE, Mehrotra B, Coleman AW. Relationship between endopoly- ploidy and cell size in epidermal tissue of Arabidopsis. Plant Cell 1993;5:1661–8.

[10] Galbraith DW, Harkins KR, Knapp S. Systemic endopolyploidy in Arabidopsis thaliana. Plant Physiology 1991;96:985–9.

[11] Storchova Z, Pellman D. From polyploidy to aneuploidy, genome instability and cancer. Nature Reviews Molecular Cell Biology 2004;5:45–54.

[12] Pécrix Y, Rallo G, Folzer H, Cigna M, Gudin S, Le Bris M. Polyploidization mechanisms: temperature environment can induce diploid gamete formation in Rosa sp. Journal of Experimental Botany 2011;62:3587–97.

[13] Davoli T, Denchi EL, de Lange T. Persistent telomere damage induces bypass of mitosis and tetraploidy. Cell 2010;141:81–93.

[14] Caperta AD, Delgado M, Ressurreic¸ ão F, Meister A, Jones RN, Viegas W, et al. Colchicine-induced polyploidization depends on tubulin polymerization in c-metaphase cells. Protoplasma 2006;227:147–53.

[15] Kondorosi E, Roudier F, Gendreau E. Plant cell-size control: growing by ploidy? Current Opinion in Plant Biology 2000;3:488–92.

[16] Lavania UC, Srivastava S, Lavania S, Basu S, Misra NK, Mukai Y. Autopolyploidy differentially influences body size in plants, but facilitates enhanced accumulation of secondary metabolites, causing increased cytosine methylation. Plant Journal 2012;71:539–49.

[17] Mable BK. Why polyploidy is rarer in animals than in plants': myths and mechanisms. Biological Journal of the Linnean Society 2004;82.

[18] Ohno S. Evolution by gene duplication. New York: Springer; 1970.

[19] Mable BK, Alexandrou MA, Taylor MI. Genome duplication in amphibians and fish: an extended synthesis. Journal of Zoology 2011;284:151–82.

[20] Soltis PS, Soltis D. Polyploidy and genome evolution. Berlin, Heidelberg: Springer-Verlag; 2013.

[21] Edger PP, Pires JC. Gene and genome duplications: the impact of dosage sensitivity on the fate of nuclear genes. Chromosome Research 2009;17:699–717.

[22] Otto SP. The evolutionary consequences of polyploidy. Cell 2007;131:452–62.

[23] Ohno S. Gene duplication and the uniqueness of vertebrate genomes circa 1970–1999. Seminars in Cell and Developmental Biology 1999;10:517–22.

[24] Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. Ancestral polyploidy in seed plants and angiosperms. Nature 2011;473:97–100.

[25] Masterson J. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 1994;264:421–4.

[26] Schranz ME, Mohammadin S, Edger PP. Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. Current Opinion in Plant Biology 2012;15:147–53.

[27] Arrigo N, Barker MS. Rarely successful polyploids and their legacy in plant genomes. Current Opinion in Plant Biology 2012;15:140–6.

[28] Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, et al. Recently formed polyploid plants diversify at lower rates. Science 2011;333:1257.

[29] Mayfield D, Chen ZJ, Pires JC. Epigenetic regulation of flowering time in polyploids. Current Opinion in Plant Biology 2011;14:174–8.

[30] Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. The frequency of polyploid speciation in vascular plants. Proceedings of the National Academy of Sciences of the United States of America 2009;106:13875–9.

[31] Brownfield L, Köhler C. Unreduced gamete formation in plants: mechanisms and prospects. Journal of Experimental Botany 2011;62:1659–68.

[32] Uryvaeva IV, Brodsky VYU. Genome multiplication, growth and development: biology of polyploid and polytene cells. Great Britain; 1985.

[33] Ravid K, Lu J, Zimmet JM, Jones MR. Roads to polyploidy: the megakaryocyte example. Journal of Cellular Physiology 2002;190:7–20.

[34] Terada N, Hamazaki T, Oka M, Hoki M, Mastalerz DM, Nakano Y, et al. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. Nature 2002;416:542–5.

[35] Hörandl E, Hojsgaard D. The evolution of apomixis in angiosperms: a reappraisal. Plant Biosystems 2012;146:681–93.

[36] Ravi M, Marimuthu MPA, Siddiqi I. Gamete formation without meiosis in Arabidopsis. Nature 2008;451:1121–4.

[37] Yamauchi A, Hosokawa A, Nagata H, Shimoda M. Triploid bridge and role of parthenogenesis in the evolution of autopolyploidy. American Naturalist 2004;164:101–12.

[38] Köhler C, Mittelsten Scheid O, Erilova A. The impact of the triploid block on the origin and evolution of polyploid plants. Trends in Genetics 2010;26:142–8.

[39] Ramsey J, Schemske DW. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 1998;29:467–501.

[40] Ramsey J, Schemske DW. Neopolylploidy in flowering plants. Annual Review of Ecology and Systematics 2002;33:589–639.

[41] Veitia RA, Bottani S. Whole genome duplications and a 'function' for junk DNA? Facts and hypotheses. PLoS One 2009;4:e8201.

[42] Cavalier-Smith T. Skeletal DNA and the evolution of genome size. Annual Review of Biophysics and Bioengineering 1982;11:273–302.

[43] Mirsky AE, Ris H. The desoxyribonucleic acid content of animal cells and its evolutionary significance. Journal of General Physiology 1951;34:451–62.

[44] Liu B, Vega JM, Feldman M. Rapid genomic changes in newly synthesized amphiploids of Triticum and Aegilops. II. Changes in low-copy coding DNA sequences. Genome 1998;41:535–42.

[45] Chen ZJ, Ni Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays 2006;28:240–52.

[46] Adams KL, Wendel JF. Novel patterns of gene expression in polyploid plants. Trends in Genetics 2005;21:539–43.

[47] Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC. Genomic changes in resynthesized Brassica napus and their effect on gene expression and phenotype. Plant Cell 2007;19:3403–17.

[48] Guo M, Davis D, Birchler JA. Dosage effects on gene expression in a maize ploidy series. Genetics 1996;142:1349–55.

[49] Fox DT, Gall JG, Spradling AC. Error-prone polyploid mitosis during normal Drosophila development. Genes and Development 2010;24:2294–302.

[50] Ganem NJ, Storchova Z, Pellman D. Tetraploidy, aneuploidy and cancer. Current Opinion in Genetics and Development 2007;17:157–62.

[51] Ganem NJ, Godinho SA, Pellman D. A mechanism linking extra centrosomes to chromosomal instability. Nature 2009;460:278–82.

[52] Borel F, Lohez OD, Lacroix FB, Margolis RL. Multiple centrosomes arise from tetraploidy checkpoint failure and mitotic centrosome clusters in p53 and RB pocket protein-compromised cells. Proceedings of the National Academy of Sciences of the United States of America 2002;99:9819–24.

[53] Comai L. The advantages and disadvantages of being polyploid. Nature Reviews Genetics 2005;6:836–46.

[54] Riley R, Chapman V. Genetic control of the cytologically diplod behaviour of hexaploid wheat. Nature 1958;182:713–5.

[55] Sears ER. Genetic control of chromosome pairing in wheat. Annual Review of Genetics 1976;10:31–51.

[56] Wall AM, Riley R, Gale MD. The position of a locus on chromosome 5B of Triticum aestivum affecting homoeologous meiotic pairing. Genetics Research 1971;18:329–39.

[57] Prieto P, Shaw P, Moore G. Homologue recognition during meiosis is associated with a change in chromatin conformation. Nature Cell Biology 2004;6:906–8.

[58] Colas I, Shaw P, Prieto P, Wanous M, Spielmeyer W, Mago R, et al. Effective chromosome pairing requires chromatin remodeling at the onset of meio- sis. Proceedings of the National Academy of Sciences of the United States of America 2008;105:6075–80.

[59] Lukaszewski AJ. The Ph1 locus from wheat controls meiotic chromosome pairing in Autotetraploid rye (Secale cereale L.). Cytogenetic and Genome Research 2010;129:117–23.

[60] Nicolas SD. Genetic regulation of meiotic cross-overs between related genomes in Brassica napus haploids and hybrids. Plant Cell 2009;21:373–85.

[61] Jenczewski E, Eber F, Grimaud A, Huet S, Lucas MO, Monod H, et al. PrBn, a major gene controlling homeologous pairing in oilseed rape (Brassica napus) haploids. Genetics 2003;164:645–53.

[62] Liu Z, Adamczyk K, Manzanares-Dauleux M, Eber F, Lucas MO, Delourme R, et al. Mapping PrBn and other quantitative trait loci responsible for the control of homeologous chromosome pairing in oilseed rape (Brassica napus L.) haploids. Genetics 2006;174:1583–96.

[63] Wiseman FK, Alford KA, Tybulewicz VL, Fisher EM. Down syndrome – recent progress and future prospects. Human Molecular Genetics 2009;18:R75–83.

[64] Yousafzai FK, Al-Kaff N, Moore G. The molecular features of chromosome pairing at meiosis: the polyploid challenge using wheat as a reference. Functional and Integrative Genomics 2010;10:147–56.

[65] Yousafzai FK, Al-Kaff N, Moore G. Structural and functional relationship between the Ph1 locus protein 5B2 in wheat and CDK2 in mammals. Functional and Integrative Genomics 2010;10:157–66.

[66] Al-Kaff N, Knight E, Bertin I, Foote T, Hart N, Griffiths S, et al. Detailed dissection of the chromosomal region containing the Ph1 locus in wheat Triticum aestivum: with deletion mutants and expression profiling. Annals of Botany 2008;101:863–72.

[67] Adams KL, Wendel JF. Polyploidy and genome evolution in plants. Current Opinion in Plant Biology 2005;8:135–41.

[68] Osborn TC, Chris Pires J, Birchler JA, Auger DL, Chen ZJ, Lee HS, et al. Understanding mechanisms of novel gene expression in polyploids. Trends in Genetics 2003;19:141–7.

[69] Schubert I, Lysak MA. Interpretation of karyotype evolution should consider chromosome structural constraints. Trends in Genetics 2011;27:207–16.

[70] Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I. Mechanisms of chromosome number reduction in Arabidopsis thaliana and related Brassicaceae species. Proceedings of the National Academy of Sciences of the United States of America 2006;103:5224–9.

[71] Woodhouse MR, Schnable JC, Pedersen BS, Lyons E, Lisch D, Subramaniam S, et al. Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homeologs. PLoS Biology 2010:8.

[72] Cheng F, Wu J, Fang L, Sun S, Liu B, Lin K, et al. Biased gene fractionation and dominant gene expression among the subgenomes of Brassica rapa. PLoS One 2012:7.

[73] Sankoff D, Zheng C, Wang B. A model for biased fractionation after whole genome duplication. BMC Genomics 2012;13(Suppl. 1).

[74] Schnable JC, Springer NM, Freeling M. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proceedings of the National Academy of Sciences of the United States of America 2011;108:4069–74.

[75] Thomas BC, Pedersen B, Freeling M. Following tetraploidy in an Arabidopsis a ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Research 2006;16:934–46.

[76] Freeling M, Woodhouse MR, Subramaniam S, Turco G, Lisch D, Schnable JC. Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants. Current Opinion in Plant Biology 2012;15:131–9.

[77] Sankoff D, Zheng C, Zhu Q. The collapse of gene complement following whole genome duplication. BMC genomics 2010;11:313.

[78] Langham RJ, Walsh J, Dunn M, Ko C, Goff SA, Freeling M. Genomic duplication, fractionation and the origin of regulatory novelty. Genetics 2004;166:935–45.

[79] Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 2006;440:341–5.

[80] Sémon M, Wolfe KH. Consequences of genome duplication. Current Opinion in Genetics and Development 2007;17:505–12.

[81] Birchler JA, Bhadra U, Bhadra MP, Auger DL. Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. Developmental Biology 2001;234:275–88.

[82] Freeling M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. Annual Review of P Plant Biology 2009;60:433–53.

[83] Birchler JA, Riddle NC, Auger DL, Veitia RA. Dosage balance in gene regulation: biological implications. Trends in Genetics 2005;21:219–26.

[84] Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, et al. Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature 2006;444:171–8.

[85] Papp B, Pal C, Hurst LD. Dosage sensitivity and the evolution of gene families in yeast. Nature 2003;424:194–7.

[86] Blomme T, Vandepoele K, De Bodt S, Simillion C, Maere S, Van de Peer Y. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biology 2006:7.

[87] Bekaert M, Edger PP, Pires JC, Conant GC. Two-phase resolution of polyploidy in the Arabidopsis metabolic network gives rise to relative followed by absolute dosage constraints. The Plant Cell 2011;23:1719–28.

[88] Bekaert M, Edger PP, Hudson CM, Pires JC, Conant GC. Metabolic and evolutionary costs of herbivory defense: systems biology of glucosinolate synthesis. New Phytologist 2012;196:596–605.

[89] Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 1997;387:708–13.

[90] Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, et al. The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome. Science 2004;304:304–7.

[91] Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 2004;428:617–24.

[92] Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, et al. Genome evolution in yeasts. Nature 2004;430:35–44.

[93] Byrne KP, Wolfe KH. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Research 2005;15:1456–61.

[94] Seoighe C, Wolfe KH. Yeast genome evolution in the post-genome era. Current Opinion in Microbiology 1999;2:548–54.

[95] Conant GC, Wolfe KH. Turning a hobby into a job: how duplicated genes find new functions. Nature Reviews Genetics 2008;9:938–50.

[96] Wapinski I, Pfeffer A, Friedman N, Regev A. Natural history and evolutionary principles of gene duplication in fungi. Nature 2007;449:54–61.

[97] Guan Y, Dunham MJ, Troyanskaya OG. Functional analysis of gene duplications in Saccharomyces cerevisiae. Genetics 2007;175:933–43.

[98] Hakes L, Pinney JW, Lovell SC, Oliver SG, Robertson DL. All duplicates are not equal: the difference between small-scale and genome duplication. Genome Biology 2007;8:R209.

[99] Amoutzias GD, He Y, Gordon J, Mossialos D, Oliver SG, Van de Peer Y. Post-translational regulation impacts the fate of duplicated genes. Proceedings of the National Academy of Sciences, USA 2010;107:2967–71.

[100] Teichmann S, Babu MM. Gene regulatory network growth by duplication. Nature Genetics 2004;36:492–6.

[101] Wagner A. The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. Molecular Biology and Evolution 2001;18:1283–92.

[102] Qian W, He X, Chan E, Xu H, Zhang J. Measuring the evolutionary rate of protein-protein interaction. Proceedings of the National Academy of Sciences of the United States of America 2011;108:8725–30.

[103] Drummond DA, Raval A, Wilke CO. A single determinant dominates the rate of yeast protein evolution. Molecular Biology and Evolution 2006;23:327–37.

[104] Wagner A. Energy constraints on the evolution of gene expression. Molecular Biology and Evolution 2005;22:1365–74.

[105] Evangelisti AM, Conant GC. Nonrandom survival of gene conversions among yeast ribosomal proteins duplicated through genome doubling. Genome Biology and Evolution 2010;2:826–34.

[106] Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. Proceedings of the National Academy of Sciences of the United States of America 2007;104:8397–402.

[107] Huminiecki L, Heldin CH. 2R and remodeling of vertebrate signal transduction engine. BMC Biology 2010;8:146.

[108] Huminiecki L, Goldovsky L, Freilich S, Moustakas A, Ouzounis C, Heldin CH. Emergence, development and diversification of the TGF-beta signalling pathway within the animal kingdom. BMC Evolutionary Biology 2009;9:28.

[109] Conant GC, Wolfe KH. Functional partitioning of yeast co-expression networks after genome duplication. PLoS Biology 2006;4:e109.

[110] Johnston M, Kim J-H. Glucose as a hormone: receptor-mediated glucose sensing in the yeast Saccharomyces cerevisiae. Biochemical Society Transactions 2005;33:247–52.

[111] Geladé R, Van de Velde S, Van Dijck P, Thevelein JM. Multi-level response of the yeast genome to glucose. Genome Biology 2003;4:233.

[112] Merico A, Sulo P, Pisˇkur J, Compagno C. Fermentative lifestyle in yeasts belonging to the Saccharomyces complex. FEBS Journal 2007;274: 976–89.

[113] Herrero P, Galíndez J, Ruiz N, Martínez-Campa C, Moreno F. Transcriptional regulation of the Saccharomyces cerevisiae HXK1, HXK2 and GLK1 genes. Yeast 1995;11:137–44.

[114] Blank LM, Lehmbeck F, Sauer U. Metabolic-flux and network analysis of fourteen hemiascomycetous yeasts. FEMS Yeast Research 2005;5:545–58.

[115] Conant GC, Wolfe KH. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. Molecular Systems Biology 2007;3:129.

[116] Galitski T, Saldanha AJ, Styles CA, Lander ES, Fink GR. Ploidy regulation of gene expression. Science 1999;285:251–4.

[117] Conant GC, Wolfe KH. Probabilistic cross-species inference of orthologous genomic regions created by whole-genome duplication in yeast. Genetics 2008;179:1681–92.

[118] Duarte NC, Herrgård MJ, Palsson BØ. Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. Genome Research 2004;14:1298–309.

[119] van Hoek MJ, Hogeweg P. Metabolic adaptation after whole genome duplication. Molecular Biology and Evolution 2009;26:2441–53.

[120] Pfeiffer T, Schuster S, Bonhoeffer S. Cooperation and competition in the evolution of ATP-producing pathways. Science 2001;292:504–7.

[121] Pfeiffer T, Schuster S. Game-theoretical approaches to studying the evolution of biochemical systems. Trends in Biochemical Sciences 2005;30:20–5.

[122] MacLean RC, Gudelj I. Resource competition and social conflict in experimental populations of yeast. Nature 2006;441:498–501.

[123] Hardin G. The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. Science 1968;162:1243–8.

[124] Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, et al. Rewiring of the yeast transcriptional network through the evolution of motif usage. Science 2005;309:938–40.

[125] Thomson JM, Gaucher EA, Burgan MF, De Kee DW, Li T, Aris JP, et al. Resurrecting ancestral alcohol dehydrogenases from yeast. Nature Genetics 2005;37:630–5.

[126] Boveri T. Concerning the origin of malignant tumours by Theodor Boveri. Journal of Cell Science 2008;121:1–84 Translated and annotated by Henry Harris.

[127] Shackney SE, Smith CA, Miller BW, Burholt DR, Murtha K, Giles HR, et al. Model for the genetic evolution of human solid tumors. Cancer Research 1989;49:3344–54.

[128] Frohling S, Dohner H. Chromosomal abnormalities in cancer. New England Journal of Medicine 2008;359:722–34.

[129] Mitelman F. Recurrent chromosome aberrations in cancer. Mutation Research 2000;462:247–53.

[130] Chester M, Gallagher JP, Symonds VV, Cruz da Silva AV, Mavrodiev EV, Leitch AR, et al. Extensive chromosomal variation in a recently formed nat- ural allopolyploid species, Tragopogon miscellus (Asteraceae). Proceedings of the National Academy of Sciences of the United States of America 2012;109:1176–81.

[131] Xiong Z, Gaeta RT, Pires JC. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid Brassica napus. Proceedings of the National Academy of Sciences of the United States of America 2011;108:7908–13.

[132] Raber MN, Barlogie B. DNA flow cytometry of human solid tumors. In: Melamed MR, Lindmo T MM, editors. Flow cytometry and sorting. New York: Wiley-Liss; 1990. p. 745–54.

[133] Merlo LM, Wang LS, Pepper JW, Rabinovitch PS, Maley CC. Polyploidy, aneuploidy and the evolution of cancer. Advances in Experimental Medicine and Biology 2010;676:1–13.

[134] Levine DS, Sanchez CA, Rabinovitch PS, Reid BJ. Formation of the tetraploid intermediate is associated with the development of cells with more than four centrioles in the elastase-simian virus 40 tumor antigen transgenic mouse model of pancreatic cancer. Proceedings of the National Academy of Sciences of the United States of America 1991;88:6427–31.

[135] Lv L, Zhang T, Yi Q, Huang Y, Wang Z, Hou H, et al. Tetraploid cells from cytokinesis failure induce aneuploidy and spontaneous transformation of mouse ovarian surface epithelial cells. Cell Cycle 2012;11:2864–75.

[136] Heselmeyer K, Schrock E, du Manoir S, Blegen H, Shah K, Steinbeck R, et al. Gain of chromosome 3q defines the transition from severe dysplasia to inva-

sive carcinoma of the uterine cervix. Proceedings of the National Academy of Sciences of the United States of America 1996;93:479–84.

[137] Ornitz DM, Hammer RE, Messing A, Palmiter RD, Brinster RL. Pancreatic neoplasia induced by SV40 T-antigen expression in acinar cells of transgenic mice. Science 1987;238:188–93.

[138] Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 1083 2011;333:1039–43. 1084

[139] Pavelka N, Rancati G, Li R. Dr Jekyll and Mr Hyde: role of aneuploidy in cellular adaptation and cancer. Current Opinion in Cell Biology 2010;22:809–15.

[140] Pepper JW, Scott Findlay C, Kassen R, Spencer SL, Maley CC. SYNTHESIS: cancer research meets evolutionary biology. Evolutionary Applications 2009;2:62–70.

[141] Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature 2009;458:719–24.

[142] Ye CJ, Stevens JB, Liu G, Bremer SW, Jaiswal AS, Ye KJ, et al. Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer. Journal of Cellular Physiology 2009;219:288–300.

[143] Friedl P, Alexander S. Cancer invasion and the microenvironment: plasticity and reciprocity. Cell 2011;147:992–1009.

[144] Parsons BL. Many different tumor types have polyclonal tumor origin: evidence and implications. Mutation Research 2008;659:232–47.

[145] Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, et al. Tumour evolution inferred by single-cell sequencing. Nature 2011;472:90–4.

[146] Kim JW, Dang CV. Cancer's molecular sweet tooth and the Warburg effect. Cancer Research 2006;66:8927–30.

[147] Dell' Antone P. Energy metabolism in cancer cells: how to explain the Warburg and Crabtree effects? Medical Hypotheses 2012;79:388–92.

[148] Diaz-Ruiz R, Rigoulet M, Devin A. The Warburg and Crabtree effects: on the origin of cancer cell energy metabolism and of yeast glucose repression. Biochimica et Biophysica Acta 2011;1807:568–76.

[149] Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 2012;486:405–9.

[150] Hittinger CT, Carroll SB. Gene duplication and the adaptive evolution of a classic genetic switch. Nature 2007;449:677–81.

[151] Gordon JL, Byrne KP, Wolfe KH. Additions, losses and rearrangements on the evolutionary route from a reconstructed ancestor to the modern Saccharomyces cerevisiae genome. PLoS Genetics 2009;5:e1000485.

[152] Pérez-Bercoff Å, McLysaght A, Conant GC. Patterns of indirect protein interactions suggest a spatial organization to metabolism. Molecular BioSystems 2011;7:3056–64.

[153] Huminiecki L, Conant GC. Polyploidy and the evolution of complex traits. International Journal of Evolutionary Biology 2012;2012:292068.

CHAPTER 2. Epigenetic regulation of flowering time in polyploids

Authors: Dustin Mayfield, Z. Jeffrey Chen, and J. Chris Pires

Published: Current Opinion in Plant Biology, Vol. 14: 174-178, April 2011

Abstract

Polyploidy plays a significant role in the evolution of flowering plants. Understanding the effects of polyploidy on the epigenetic regulation of adaptive traits may resolve questions about the success of polyploids. One such trait, flowering time, has been the subject of several gene expression studies because it has one of the best characterized genetic networks and because polyploidy has a significant impact on generating variation in flowering time. Future research on the epigenetic consequences of polyploidy on flowering time should begin to examine natural variation in an ecological context, while continuing to make use of resynthesized polyploids.

Introduction

Polyploidy is an important cause of plant speciation and diversification [1, 2•, 3•, 4, 5•, 6•, 7•, 8, 9 and 10]. The estimated frequency of speciation events accompanied by ploidy increase in angiosperms and ferns, has been reported as 15% and 31%,

respectively [11]. Many important crop species show evidence of genome duplication. Both potato and alfalfa are derived from autopolyploidy, while wheat, oat, cotton, coffee, and canola have allopolyploidy in their past evolutionary history [12]. Other crops, such as cabbage, soybean, and maize, are considered paleopolyploids due to their more recent diploidization. Genome duplication in angiosperms may be cyclic, with repeated rounds of polyploidy followed by diploidization, which can begin immediately after polyploidization [4].

The widespread occurrence, and re-occurrence, of polyploidy is thought to reflect the greater ability of neopolyploids to adapt to new environments compared to their parental taxa. Further study into changing life history traits resulting from epigenetic modification via neopolyploidy will help explain how this phenomenon has been a major force in the evolution and speciation of plants. Mechanisms that may contribute to the phenotypic changes associated with polyploidy have been extensively reviewed [1, 2•, 3•, 4, 5•, 6•, 7•, 8, 9 and 10]. Altered regulatory interactions, increased variation of dosage-regulated gene effects and expression, genetic changes (e.g. insertions, deletions, translocations, and gene conversions), and epigenetic changes (repression or derepression of gene expression) have been suggested to contribute to phenotypic changes resulting from the interaction between genomes that have undergone either autopolyploid or allopolyploid events. Owing to the rarity of these events, polyploidy must sometimes confer an immediate advantage to ensure the establishment of polyploids among their diploid parents. Although mechanisms affecting functional divergence of duplicate genes may confer a selective advantage to new polyploids over time, those that confer novel forms of

gene expression resulting in immediate phenotypic effects have been suggested as interesting avenues for further investigation into the success of polyploids [7• and 8].

The flowering time pathway

Plants rely on environmental and endogenous cues to initiate flowering. Seasonal shifts in temperature and day length are perceived by plants through the vernalization and photoperiod pathways, respectively. Together, along with the endogenous changes that permit or cause flowering (sometimes referred to as autonomous pathways), these pathways are often referred to as flowering pathways. Much of the current understanding of flowering has come from Arabidopsis research [13••]. Many recent reviews highlight specific aspects of flowering [14•, 15, 16, 17, 18•, 19• and 20]. Extensive research has contributed to making the flowering time network one of the most well-characterized networks in plants. Genetically and epigenetically altered regulatory interactions of pathways, such as the flowering time pathway, are likely responsible for the increased phenotypic variation observed in polyploids (see Figure 1). Here, we focus on how polyploidy uniquely contributes to the evolution of flowering time via epigenetic change.

In Arabidopsis and Brassica, late flowering is a dominant trait probably because flowering time is mainly dependent on the expression of a flowering repressor, FLC, in the autonomous pathway. Consequently, the interspecific hybrids and allotetraploids between an early-flowering and a late-flowering parent in Arabidopsis or Brassica are often late flowering. Data support a role of antisense

and noncoding RNAs in FLC expression [21]. Lui et al. found that a transposable element inserted in an intron of FLC in the commonly used Arabidopsis accession Landsberg erecta (Ler) is associated with repressive chromatin modifications to FLC-Ler that are mediated by siRNAs generated from homologous transposable elements in the genome [22]. It may be that different antisense and noncoding RNAs originating from different species regulate expression of flowering time genes. In plants such as corn and wheat that do not have FLC, the hybrids between an early-flowering and a late-flowering parent usually flower early. This is because flowering time is controlled by the *FLOWERING LOCUS T* (*FT*)-like gene that promotes early flowering in these plants.

The photoperiod pathway consists of photoreceptor, circadian clock, and circadian clock-regulated genes and promotes the floral transition in response to a long photoperiod [23]. In Arabidopsis, this pathway includes downstream circadian clock genes, including *FT* and *CONSTANS* (*CO*). *CO* transcription is mediated by circadian clock regulators such as *GIGANTEA* (*GI*) [24]. In long-day (LD) conditions, *CO* expression peaks during the light period, resulting in the activation of *FT* expression, leading to early flowering [25 and 26]. LD plants respond to lengthening days and flower in the spring or early summer, whereas short-day (SD) plants flower in late summer or autumn in response to shortening days and lengthening nights [27]. Several flowering genes were identified in temperate cereals such as wheat that has vernalization response and LD flowering behaviors. Bread wheat (*Triticum aestivum* L., 2n = 6x = 42) is an allohexaploid and quantitative LD plant, and SD conditions delay the heading time [28]. Although the same names are used,

the *VRN* genes in wheat are completely different from those in Arabidopsis. Although both winter and spring wheat are allohexaploid, winter wheat requires winter vernalization to flower and is sown in the fall, whereas spring wheat does not require vernalization and can be planted in the spring or fall. If all genes are present, as shown in genetic mapping studies, the expression of *VRN* genes and other genes in photoperiod and GA pathways is likely to be altered in winter and spring wheat through genetic and epigenetic mechanisms [10]. It will be interesting to understand how these genes are regulated in allohexaploid wheat in response to vernalization and to the domestication of winter and spring wheat.

Polyploidy and flowering time

Allopolyploids are formed through the combination of two or more genomes from related species, giving rise to the effects of genome doubling as well as hybridization [1, 2•, 3•, 4, 5•, 6•, 7•, 8, 9 and 10]. Hybridization can induce allelic interactions and epigenetic modifications of homoeologous loci in the same nuclei, leading to nonadditive gene regulation [9 and 10]. In addition, the hybridization effects may also be amplified through genome doubling as a consequence of dosage regulation. Thus, allopolyploidy has larger effects on gene expression and phenotypic variation than autopolyploidy [6•, 29 and 30••]. At gene expression levels, the regulatory networks are altered by the addition of genes that possess diverged cis-regulatory elements and/or trans-acting factors. Many genes in various biological pathways, including flowering time, are expressed nonadditively.

Two epistatically acting loci, namely *FRI* and *FLC*, largely control natural

variation of flowering time in Arabidopsis. Current data suggest a model for genetic and epigenetic interactions between homoeologous loci in a genetic pathway that mediates flowering time variation in Arabidopsis allotetraploids, which explains how new allopolyploid species combine compatible biological pathways by selecting and modifying the expression of orthologous loci originating from divergent species [10 and 30••].

Two Arabidopsis species, *A. arenosa* and *A. thaliana*, diverged ~6 million years ago and probably differ in flowering habits because of selective adaptation to cold and warm climate [31], respectively. In the genetic pathway, sequence evolution of *FRI* and *FLC* loci has led to a nonfunctional *AtFRI* in *A. thaliana* and cis-regulatory changes in *A. thaliana* and *A. arenosa FLC* loci. In synthetic allotetraploids, *A. arenosa FRI* interacts in trans with the downstream gene, *AtFLC*, making the synthetic allotetraploids late flowering in a dosage-dependent manner [30]. Interestingly, it was trans-activation of *AtFLC* (instead of *AaFLC*) that determined late flowering in the synthetic allotetraploids, which is related to a long promoter segment of *AtFLC* including cis-elements for vernalization [32]. On the contrary, the promoter regions of *AaFLC* loci are small and contain a minimal segment for *FLC* expression, and their expression levels are low. *AtFLC1* and *Arabidopsis suecica FLC1* (*AsFLC1*) with intact cis-regulatory elements (promoters and/or introns) are selectively associated with a strong winter-annual habit in natural *A. suecica* strains, whereas *AaFLC2* and *AsFLC2* expression appears to be dispensable. The effects of *AaFRI* on *AtFLC* and *AaFLC1/AsFLC1* upregulation are mediated by histone acetylation and methylation [30••], which is likely affected by

other genes in the autonomous and vernalization pathways. For example, VERNALIZATION INSENSITIVE3 (VIN3), a PHD-domain-containing protein, is induced only after a prolonged period of cold. VIN3-induced silencing involves changes in histone modifications of *FLC* chromatin [33]. The stable silencing of *FLC* also requires the DNA-binding protein VERNALIZATION1 (VRN1) and the polycomb-group protein VRN2. VRN2 represses *FLC* expression through histone H3 K27 dimethylation, and VRN1 acts downstream of VRN2 [34]. The effects of these vernalization-responsive genes and many genes in other pathways impacting flowering time in polyploids are largely unknown. The hypothetical effects of genetic and epigenetic loci duplication and network duplication on pathways such as the flowering time pathway can be predicted (see Figure 2). It is likely that genetic interactions as well as epigenetic modifications of key regulatory genes in flowering pathways are responsible for a wide range of flowering time variation observed in plant polyploids.

Conclusions and prospects

The genetic and genomic consequences of polyploidization have undergone much more investigation in recent years than other important areas of polyploid evolution [1, 2•, 3•, 4, 5•, 6•, 7•, 8, 9 and 10]. The epigenetic consequences of genome merger may reveal how polyploidy contributes to phenotypic changes, as seen in Arabidopsis [10, 35 and 36•], Brassica [37•• and 38•], Spartina [39], and wheat [40 and 41]: and epigenetic phenomena associated with polyploidy will continue to be an active area of research. Studying an array of natural and resynthesized polyploids in common

garden experiments in a variety of ecological contexts is another understudied area of polyploidy research. Ideally, a system that gives rise to extreme phenotypic variation while maintaining a large degree of genome stability following neopolyploidy would be most useful in studying the epigenetic effects of polyploidy on flowering time. A potentially powerful approach would be to resynthesize polyploids that generate a high degree of variation in flowering time while originating from progenitors that previously underwent stable polyploid events in nature. Selecting lines for polyploid resynthesis based on their successful polyploidization in nature could result in a greater chance of finding instances where whole genome duplications, and their consequent epigenetic changes, give rise to adaptive traits such as changes in flowering time.

Figures

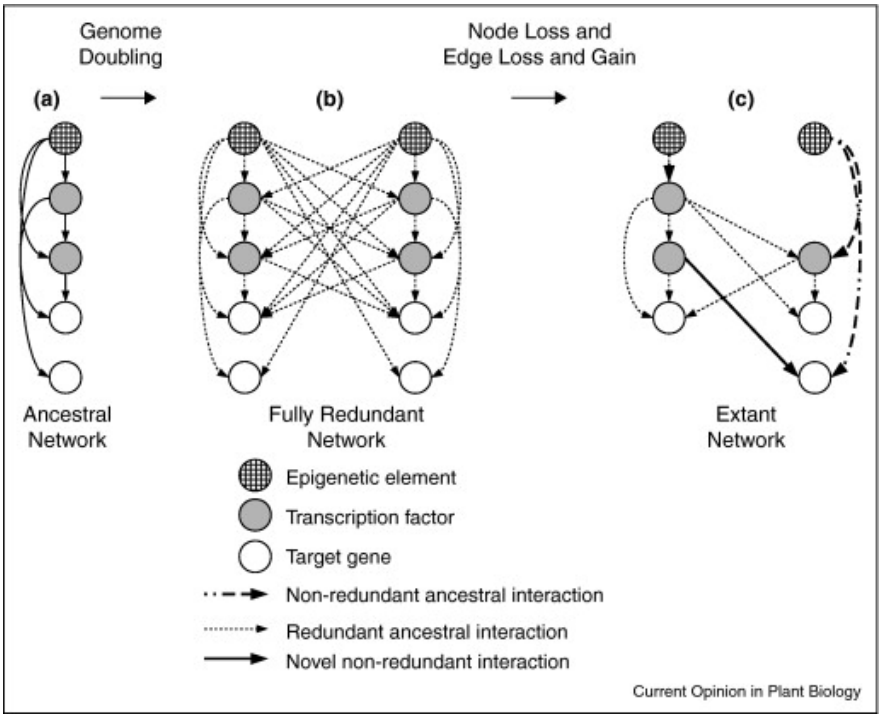


Figure 1. Network evolution as a consequence of genome doubling.

Diagrammatic representation of networks with genetic and epigenetic elements represented by circles and interactions by arrows. (a) Ancestral network with epigenetic element (that may regulate transcription factors or target genes via, for example, histone modification, chromatin remodeling, or DNA methylation), transcription factors, and target genes. (b) The fully redundant network after duplication by genome doubling. (c) The extant network following node/edge loss and gain. Nonredundant and redundant ancestral interactions remain within the network in addition to novel nonredundant interactions. Changes in gene expression of epigenetically regulated loci may contribute to novel phenotypes observed in polyploids.

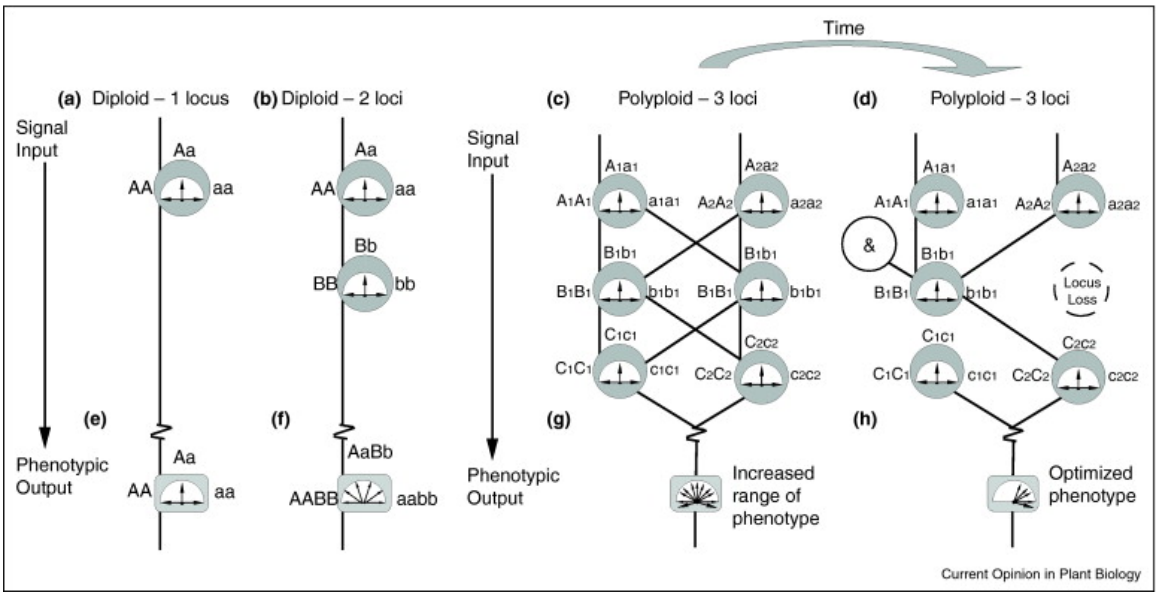


Figure 2. Hypothetical effects of the duplication of genetic and epigenetic loci on the regulation of flowering time. Circular rheostats (a)–(d) represent genetic and epigenetic loci ultimately regulating phenotype. Rectangular rheostats (e)–(h) represent the possible range of phenotypes observed in each hypothetical network. White half circles indicate the range of signal variation for each locus. Upper and lower case letters represent alternative alleles and epialleles. Black lines connect loci in pathway from their signal input to their respective possible phenotypic output. (a) Two alleles/epialleles at one locus regulate flowering time in a rheostat-like manner. (b) Two alleles/epialleles at two loci regulate flowering time, allowing for increased phenotypic variation within the range limits of parental expression. (c) Duplication of genetic and epigenetic loci within the context of a network as a consequence of recent genome doubling. Potentially, genetic and epigenetic loci could have complete redundancy in their interactions, as shown here. However, in allopolyploids where combining genomes may significantly differ from each other, cross-talk between pathways originating from previously separated genomes may

interact to a lesser degree. (d) Duplication of genetic and epigenetic loci within the context of a network following locus and interaction loss over time. & represents a locus or pathway that is acquired by the flowering pathway. In both (c) and (d), phenotypic variation can vary more widely depending on allele/epiallele context. Additionally in (c), numerous allele/epiallele combinations allow for the instant optimization of flowering time for a specific environment that includes phenotypes outside the range of the diploid parents. These instant optimizations and extreme phenotypes may explain why some species that undergo genome doubling have increased fitness compared to their diploid progenitors. (e)–(h) Overall possible ranges of phenotypic variation in flowering time observed in situation (a), (b), (c), and (d) shown as (e), (f), (g), and (h), respectively.

Acknowledgments

We apologize for being unable to cite all the relevant literature owing to space limitations of this article. Our research on polyploidy is supported by grants from the National Institutes of Health (GM067015 to ZJC) and the National Science Foundation (DBI 0624077 to ZJC, DBI 0638536 and DBI 0501712 to JCP).

Literature Cited

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

[1] Doyle JJ, *et al.* (2008) Evolutionary genetics of genome merger and doubling in plants. *Annu Rev Genet* 42:443-461.

[2] Hegarty MJ & Hiscock SJ (2008) Genomic clues to the evolutionary success of polyploid plants. *Curr Biol* 18(10):R435-444.
* While reviewing the impact of genetic and genomic changes and mechanisms in polyploid plants, the authors argue for more research with field-based experiments to determine whether genetic changes associated with polyploidy are adaptive.

[3] Leitch AR & Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. *Science* 320(5875):481-483.
* This review examines the highly plastic genome structure of polyploid plants and argues that this plasticity is a major driving force behind the divergence of angiosperms.

[4] Edger PP & Pires JC (2009) Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Res* 17(5):699-717.

[5] Chen ZJ (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* 15(2):57-71.
* This review calls attention to recent studies that have helped determine the roles of nonadditive gene expression, small RNAs, and epigenetic regulation in hybrid vigor, which may lead to better exploitation of hybrids and allopolyploids.

[6] Jackson S & Chen ZJ (2010) Genomic and expression plasticity of polyploidy. *Curr Opin Plant Biol* 13(2):153-159.
* This review emphasizes the central role that epigenetic mechanisms, including chromatin modifications and small RNAs, play in shaping the molecular and phenotypic novelty that may be selected in many polyploid plants.

[7] Soltis DE, Buggs RJA, Doyle JJ, Soltis PS, & (2010) What we still don't know about polyploidy. *Taxon* 59:1387-1403.
* This review summarizes several remaining questions in the study of polyploid plants. While evidence for molecular mechanisms associated with polyploidy has seen considerable interest in the past decade and has been reviewed elsewhere, this review seeks to promote discussion and research into new avenues of polyploid

research.

[8] Osborn TC*, et al.* (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* 19(3):141-147.

[9] Chen ZJ & Ni Z (2006) Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays* 28(3):240-252.

[10] Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu Rev Plant Biol* 58:377-406.

[11] Wood TE*, et al.* (2009) The frequency of polyploid speciation in vascular plants. *Proc Natl Acad Sci U S A* 106(33):13875-13879.

[12] Soltis DE*, et al.* (2009) Polyploidy and angiosperm diversification. *Am. J. Bot.* 96(1):336-348.

[13] Amasino R (2010) Seasonal and developmental timing of flowering. *Plant J* 61(6):1001-1013.
** This review highlights several contributions of Arabidopsis research to our understanding of the molecular basis of flowering. It provides an updated discussion on flowering in the context of major contributing pathways.

[14] Turck F, Fornara F, & Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu Rev Plant Biol* 59:573-594.
* The authors of this review highlight the molecular mechanisms that activate the CONSTANS/ FLOWERING LOCUS T module that promotes flowering in response to changes in day length and integrate evidence that FT protein is a major component of florigen.

[15] Zeevaart JA (2008) Leaf-produced floral signals. *Curr Opin Plant Biol* 11(5):541-547.

[16] Adrian J, Torti S, & Turck F (2009) From decision to commitment: the molecular memory of flowering. *Mol Plant* 2(4):628-642.

[17] He Y (2009) Control of the transition to flowering by chromatin modifications. *Mol Plant* 2(4):554-564.

[18] Kim DH, Doyle MR, Sung S, & Amasino RM (2009) Vernalization: winter and the timing of flowering in plants. *Annu Rev Cell Dev Biol* 25:277-299.
* This review provides attention to flowering time, specifically the process of vernalization in Arabidopsis. It also compares and contrasts vernalization in Arabidopsis with vernalization in cereal species.

[19] Michaels SD (2009) Flowering time regulation produces much fruit. *Curr Opin Plant Biol* 12(1):75-80.
* This review explains the detailed molecular framework for the regulation of flowering time that is established in Arabidopsis.

[20] Yant L, Mathieu J, & Schmid M (2009) Just say no: floral repressors help Arabidopsis bide the time. *Curr Opin Plant Biol* 12(5):580-586.

[21] Liu F, Marquardt S, Lister C, Swiezewski S, & Dean C (2010) Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. *Science* 327(5961):94-97.

[22] Liu J, He Y, Amasino R, & Chen X (2004) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. *Genes Dev* 18(23):2873-2878.

[23] Searle I & Coupland G (2004) Induction of flowering by seasonal changes in photoperiod. *EMBO J* 23(6):1217-1222.

[24] Imaizumi T & Kay SA (2006) Photoperiodic control of flowering: not only by coincidence. *Trends Plant Sci* 11(11):550-558.

[25] Yanovsky MJ & Kay SA (2003) Living by the calendar: how plants know when to flower. *Nat Rev Mol Cell Biol* 4(4):265-275.

[26] Valverde F*, et al.* (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303(5660):1003-1006.

[27] Thomas B & Vince-Prue D (1997) *Photoperiodism in plants* (Academic Press, London).

[28] Murai K, Miyamae M, Kato H, Takumi S, & Ogihara Y (2003) WAP1, a wheat APETALA1 homolog, plays a central role in the phase transition from vegetative to reproductive growth. *Plant Cell Physiol* 44(12):1255-1265.

[29] Veitia RA, Bottani S, & Birchler JA (2008) Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends Genet* 24(8):390-397.

[30] Wang J*, et al.* (2006) Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172(1):507-517.
** This study reports genome-wide gene expression analysis of Arabidopsis synthetic allotetraploids. Genome-wide, nonadditive gene regulation is suggested to provide a molecular basis for de novo variation and allopolyploid evolution.

[31] Sall T, Jakobsson M, Lind-Hallden C, & Hallden C (2003) Chloroplast DNA indicates a single origin of the allotetraploid Arabidopsis suecica. *J Evol Biol* 16(5):1019-1029.

[32] Sheldon CC, Conn AB, Dennis ES, & Peacock WJ (2002) Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. *Plant Cell* 14(10):2527-2537.

[33] Sung S & Amasino RM (2004) Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* 427(6970):159-164.

[34] Bastow R*, et al.* (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427(6970):164-167.

[35] Ni Z*, et al.* (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457(7227):327-331.

[36] Madlung A*, et al.* (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic Arabidopsis allotetraploids. *Plant Physiol* 129(2):733-746.
* This article reports how treatment of resynthesized Arabidopsis allotetraploids with a methylation disruptor caused novel phenotypes.

[37] Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, & Osborn TC (2007) Genomic changes in resynthesized Brassica napus and their effect on gene expression and phenotype. *Plant Cell* 19(11):3403-3417.
** The authors of this study found massive variation in genotype, gene expression, and phenotype among 47 independently-derived allopolyploid lines after five generations of selfing. Traits like flowering time were highly correlated with homoeologous exhanges among the parental chromosomes that contained flowering time genes such as FLC.

[38] Pires JC*, et al.* (2004) Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids (Brassicaceae). *Biol. J. Linn. Soc.* 82:675–688.
* In this study, the authors measured changes in genome structure and transcript levels of four sets of FLOWERING LOCUS C (FLC) genes among resynthesized Brassica lines and their diploid progenitors.

[39] Salmon A, Ainouche ML, & Wendel JF (2005) Genetic and epigenetic consequences of recent hybridization and polyploidy in Spartina (Poaceae). *Mol Ecol* 14(4):1163-1175.

[40] Liu B*, et al.* (2009) Rapid genomic changes in polyploid wheat and related species: implications for genome evolution and genetic improvement. *J Genet Genomics* 36(9):519-528.

[41] Shaked H, Kashkush K, Ozkan H, Feldman M, & Levy AA (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13(8):1749-1759.

Summary

Polyploidy is a complex mutational mechanism that provides not only an increased genetic reservoir for evolutionary forces to act upon, but it involves a unique amplification of the combinatorial possibilities between features of biological networks. Only recently, scientists have begun the process of uncovering biological networks on a genome-wide scale, such as protein-protein interaction networks, metabolic networks, genetic interaction networks, co-expression networks, co-function networks, and regulatory networks (reviewed in [1]). A full account of the complex cellular networks that underlie function remains a primary goal of systems biologists [2].

Polyploid cell lineages and entire polyploid organisms are most commonly formed by mitotic and meiotic disruptions, respectively. Polyploid cell lineages arise in organisms varied along many branches of the tree of life, and here, we highlight cases in arthropods, mammals, and plants. By reviewing routes of polyploid cell lineage formation via endoreduplication, the repression of telomerase, failures in cytokinesis, or the merging of cells, we show how many features of cellular biology are affected by multiplication of the genome.

The knowledge gained from studying whole genome duplication is of interest to those outside the polyploidy research community as it invokes many cellular changes for both normal and pathological states. In addition to providing a platform for discovery about cell cycle, mitosis and meiosis, we highlight in chapter one a model for the progression of some cancers from a polyploid to aneuploid state. Drawing a causal connection between polyploidy and the cancer phenotype is

difficult given the paucity or understanding about mutation and cancer and the practical fact that virtually all cancers are removed following detection. However, we discuss Barrett's esophagus, a slow developing premalignant neoplasm, which provides a safe opportunity for longitudinally observing neoplastic progression in a solid tumor. Identifying additional cases in which genome-wide observations can be made of lineage-specific cells would help address the progression of unchecked growth canonical to cancer.

A major contribution to chapter one is the link between the metabolic shifts to glucose fermentation from glucose respiration in cancer (Warburg effect) and yeast (the Crabtree effect) (also see [3] for review). This shift in cancer may be similarly attributable to polyploidy as in yeast. The links between these metabolic phenomenon and polyploidy deserve further analysis, and a first step may be to compare metabolite loss following polyploidy in both yeast and cancer cell lineages that have underwent this change in the cellular processing of glucose. In the case of Barrett's esophagus mentioned previously, recent evidence suggests a metabolic adaptation to microenvironments that may be similar to the changes typically associated with cancer [4].

At the organismal level, we define temporal periods of changes following polyploidy: neopolyploidy, mesopolyploidy, and paleopolyploidy. These terms are useful in thinking about the rate and series of events regarding biological network evolution following whole genome duplication in addition to what may be driving nascent polyploids to "dead ends" while older ones have been a "success." The timing of events are important in light of the insightful observation that in several

instances in plants, a whole genome duplication occurs, a novel trait evolves, and after a significant amount of time, one of two sister groups rapidly radiates. These data suggest a model, termed the Whole Genome Duplication Lag-Time model, in which the novel trait alone does not contribute to the successful diversification of the resulting crown group but rather subsequent phenomena such as migration or changing climactic conditions [5]. By recognizing the importance of differentiating polyploids into this spectrum, scientists can better conceptualize what may be occurring in lineages with varying ages of polyploidy, thus deepening our research agenda. As researchers continue to map biological networks of polyploids, we will want to compare the effects of genome doubling of various ages. We can then ask questions about the basic properties of network evolution such as what factors contribute the evolvabiltiy of genes, pathways, and traits.

In chapter two, we review the epigenetic effects following polyploidy on one such trait, flowering time. Flowering time represents one of many traits important for adaptation to different environments and has been shown to diverge in resynthesized polyploids [6]. Flowering time is regulated by a well-characterized set of integrated pathways [7]. Comparative analysis between related species has demonstrated alternative mechanisms for inducing flowering in different plant lineages. As in the case of key flowering repressors, major differences have been found in the regulation of the vernalization pathway [8]. In chapter two, we address how polyploidy may result in genetically and epigenetically altered regulatory interactions.

Specifically, we review a model for how these interactions mediate flowering time variation in *Arabidopsis* allotetraploids. The diploid progenitors, *A. arenosa* and *A. thaliana*, are adapted to cold and warm environments, respectively. In the warm-adapted *A. thaliana*, the ortholog controlling response to cold temperature is non-functional. Interestingly, this non-functional gene in *A. thaliana* interacted with its upstream floral regulator from the cold-adapted species in the allotetraploid and not the ortholog contributed from *A. arenosa*. The interaction between the two loci is mediated by histone modification. Indeed, many other epigenetic processes have been found to control flowering time and have been reviewed elsewhere [9-11].

Chapter two combines this evidence from Arabidopsis with hypothetical models for how two genomes, when combined via polyploidy, can result in epigenetic and genetic changes that affect phenotype. We suggest factors for selecting useful organismal models for further research of natural and resynthesized polyploids to gain insight into the immediate effects of whole genome duplication. This will aid in our understanding of why and how new polyploids are established and maintained while it appears most go extinct in nature [12, 13].

Together, these chapters synthesize research regarding the occurrence and implications of biological network evolution following polyploidy. As I leave my current position at the University of Missouri, I look forward to learning more about the fruits of comparative genomics and interactomics as scientists elucidate the molecular mechanisms controlling agronomically important traits, such as flowering time. I hope to witness when systems biology contributes to the engineering of better plants for our needs the way that traditional breeding and molecular genetics

has revolutionized agricultural production.

Literature Cited

1. Chae, L., et al., *Towards understanding how molecular networks evolve in plants.* Curr Opin Plant Biol, 2012. **15**(2): p. 177-84.

2. Aon, M., *Complex Systems Biology of Networks: The Riddle and the Challenge*, in *Systems Biology of Metabolic and Signaling Networks*, M.A. Aon, V. Saks, and U. Schlattner, Editors. 2014, Springer Berlin Heidelberg. p. 19-35.

3. Diaz-Ruiz, R., M. Rigoulet, and A. Devin, *The Warburg and Crabtree effects: On the origin of cancer cell energy metabolism and of yeast glucose repression.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2011. **1807**(6): p. 568-576.

4. Suchorolski, M.T., et al., *Warburg and Crabtree Effects in Premalignant Barrett's Esophagus Cell Lines with Active Mitochondria.* PLoS One, 2013. **8**(2): p. e56884.

5. Schranz, E., M., S. Mohammadin, and P.P. Edger, *Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model.* Current Opinion in Plant Biology, 2012. **15**(2): p. 147-153.

6. Pires, J.C., et al., *Flowering time divergence and genomic rearrangements in resynthesized Brassica polyploids (Brassicaceae).* Biological Journal of the Linnean Society, 2004. **82**(4): p. 675-688.

7. Srikanth, A. and M. Schmid, *Regulation of flowering time: all roads lead to Rome.* Cell Mol Life Sci, 2011. **68**(12): p. 2013-37.

8. Kim, D.H., et al., *Vernalization: winter and the timing of flowering in plants.* Annu Rev Cell Dev Biol, 2009. **25**: p. 277-99.

9. Yaish, M.W., J. Colasanti, and S.J. Rothstein, *The role of epigenetic processes in controlling flowering time in plants exposed to stress.* J Exp Bot, 2011. **62**(11): p. 3727-35.

10. Zografou, T. and F. Turck, *Epigenetic Control of Flowering Time*, in *Epigenetic Memory and Control in Plants*, G. Grafi and N. Ohad, Editors. 2013, Springer Berlin Heidelberg. p. 77-105.

11. Turck, F. and G. Coupland, *NATURAL VARIATION IN EPIGENETIC GENE REGULATION AND ITS EFFECTS ON PLANT DEVELOPMENTAL TRAITS.* Evolution, 2013: p. n/a-n/a.

12. Mayrose, I., et al., *Recently Formed Polyploid Plants Diversify at Lower Rates.* Science, 2011. **333**(6047): p. 1257.

13. Arrigo, N. and M.S. Barker, *Rarely successful polyploids and their legacy in plant genomes.* Current Opinion in Plant Biology, 2012. **15**(2): p. 140-146.