# A review of crop evolution through polyploidization

**By Kidist Bogale Kibret**

**A review paper submitted to Addis Ababa University, Ethiopia as part of the requirements for the degree in Applied Genetics**

**2005-05-14**

# Contents

## Abstract

This review article describes the basic genetics behind polyploid genomes in simple context that can be understood by any breeder without molecular biology expertise. It aims to address the importance of polyploidy in genomics and breeding studies of the major crops: Brassica, Cotton, cereals and the experimental model crop Arabidopsis. Finally, the review discusses the importance of ancient polyploid diploidization and the genome dynamics of newly formed polyploids for consideration in breeding projects.
**Keywords**: Polyploid crop evolution, diploidization, neopolyploids, paleopolyploids

## 1) Introduction

Polyploidy is the formation of a higher chromosome number by the addition of extra whole chromosome set present in one or more ancestral organisms. In short, polyploidy is the presence of three or more chromosome sets in an organism (Soltis et al, 2003).

There are four types of polyploids based on genetic and cytogenetic criteria: autopolyploids, segmental allopolyploids, true or genomic allopolyploids and autoallopolyploids. Allopolyploids form between different species, whereas autopolyploids form within species (Soltis et al, 2003). Autopolyploids contain multiple sets of the same or similar genomes in their nucleus. They originate mostly (from intraspecific crosses followed) by chromosome doubling because of unreduced gametes and their chromosomes may form multivalents during meiosis and exhibit a multisomic inheritance. One advantage of autopolyploidy is the capacity to maintain high levels of heterozygosity, with multiple alleles per locus, or more rarely, to reach homozygosity with multiple dosage of a given allele. The formation of multivalents during meiosis is often associated with sterility. Therefore, it is not surprising that many autopolyploids are perennial species that also propagate vegetatively (Levy and Feldman, 2002). Autopolyploidy is characterized by multivalents at meiosis or tetrasomic ratios. In the examples artificially produced, autopolyploids have slower development and reduced fertility (Soltis et al, 2003).

Allopolyploids (or amphiploids) contain two or more diverged homoeologous genomes derived from interspecific or intergeneric hybridization between species with diverged genomes (Levy and Feldman, 2002). True allopolyploids rarely have multivalent associations and tetrasomic ratios and they therefore resemble diploids to a large extent in their cytogenetic behaviour (Soltis et al, 2003).

The resulting hybrid is sterile but fertility is obtained after chromosome doubling. Allopolyploids are characterized by bivalent pairing, full fertility, and disomic inheritance (Levy and Feldman, 2002)**.**

Segmental allopolyploids resemble autopolyploids to a greater or lesser degree in possessing multivalents and tetrasomic ratios, but these will be less common than in autoploids (Soltis et al, 2003). Segmental allopolyploids contain both homologous and homoeologous chromosomal segments and, therefore, exhibit both bivalents and multivalents during meiosis. As a result, they have a mixed disomic/polysomic inheritance (Levy and Feldman, 2002).

Principal criteria for distinguishing between autopolyploids and amphiploids (allopolyploids) are chromosomal behavior, fertility, segregation ratios and morphology. These criteria will break down in individual cases.

The classification in summary is as follows:

I. Autopolyploids

1. Strict autopolyploid AAAA
2. Interracial autopolyploid AAAA

II. Amphiploids

3. Segmental allopolyploid $A_1A_1A_1A_1$
4. Genomic allopolyploid AABB
5. Autoallopolyploid AAAABB (Soltis et al, 2003). (Fig 1)

Polyploidy, resulting either from duplication of a single but complete genome (autopolyploidy) or from combination of two or more differentiated genomes (allopolyploidy), is a prominent mode of evolution in plants (Liu and Wendel, 2002; Soltis et al, 2004b)

It is well known that polyploidy is a major force in evolution, particularly in plants (Soltis and Soltis, 1995). It is difficult to overstate the importance of polyploidy in the evolutionay history of plants. While estimates vary regarding the proportion of angiosperms that have experienced one or more episodes of chromosome doubling at some point in their evolutionary history, perhaps 95% of pteridophytes have experienced at least one episode of polyploidization in their past. An estimate of the frequency of polyploid angiosperm species is as high as 80% (Liu and Wendel, 2002). Many crop plants, including wheat,

coffee, oat, potato, canola, soybean, sugarcane, tobacco and cotton are typical polyploids (Liu and Wendel, 2002; Singh, 1993).

Recent genomic studies have dramatically altered the polyploidy paradigm. It now appears that flowering plants and perhaps all eukaryotes possess genomes with considerable gene redundancy, much of which is likely the result of polyploidy or whole genome duplication. Recent evidences suggest that lineages may undergo repeated cycles of polyploidization followed by extensive diploidization (Soltis et al, 2003).

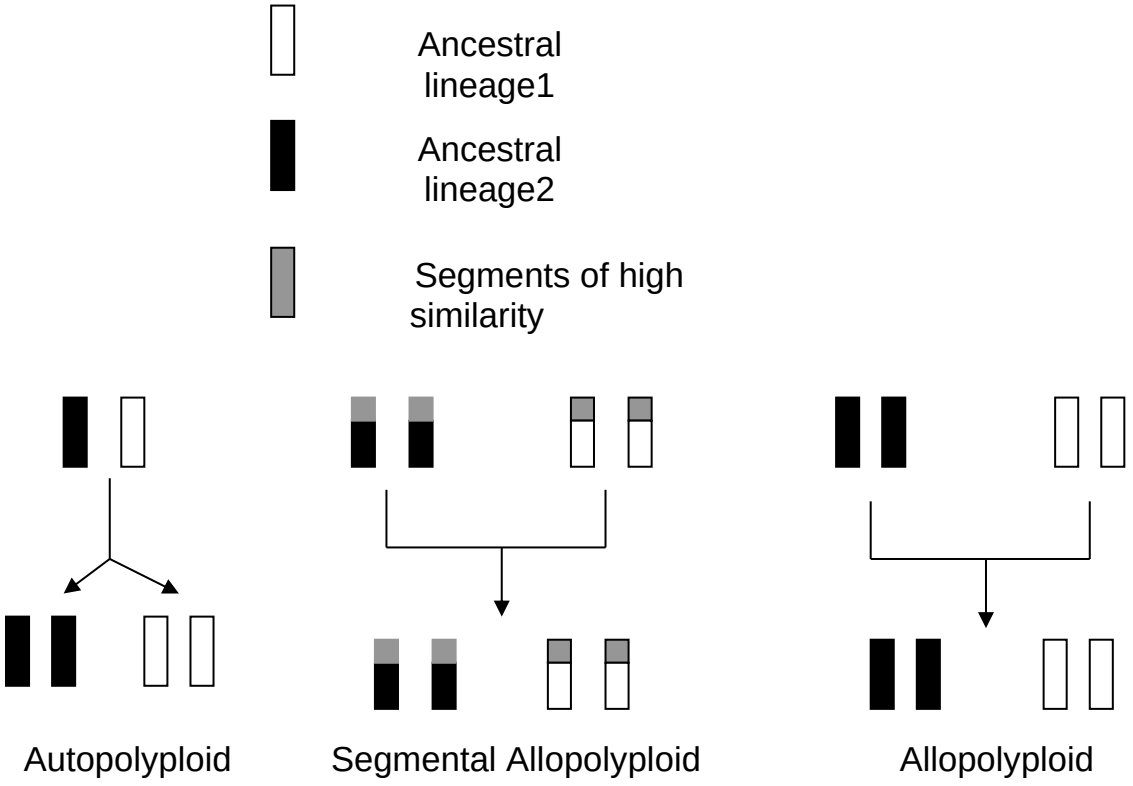


Figure 1 Spectrum of historical relatedness between progenitor genomes suggest three polyploidization categories. Source: Rauh, 2003

## 2) Recurrent formation of polyploid

According to the traditional tenets, polyploidization events were considered rare and each polyploid species was considered to have a single origin (Soltis and Soltis, 1995). According to Stebbins, (1971) hybridization of diploid species which are related enlarges the gene pools of polyploids. Only allopolyploidy was considered a major force in evolution (Stebbins, 1971). The presumed genetic uniformity of polyploids, coupled with their hypothesized reduced capacity for molding new genotypes, led to the extreme view

of polyploid species as evolutionary dead ends (Soltis and Soltis, 1999). The realization that most taxonomically recognized polyploid species are of multiple origins shattered the earlier perceptions of polyploids (Soltis and Soltis, 1995). Molecular approaches have provided a wealth of data that have dramatically reshaped views of polyploid evolution. Molecular data demonstrate that both autopolyploids and allopolypoloids exhibit a high frequency of recurrent formation (multiple origin). They reveal that multiple polyploidization events within species have significant genetic and evolutionary implications. The most important contributions of molecular data to the study of polyploid evolution is documentation that a single polyploid species may have separate, independent origin from the same diploid progenitor species. Multiple origins of polyploids have now been documented in bryophytes and in more than 40 species of ferns and in angiosperms (Soltis and Soltis, 1995). Polyphyletic polyploid species have been reported for mosses, ferns, and many angiosperms, and include both autopolyploids (e.g., *Heuchera grossulariifolia* and *Heuchera micrantha*) and allopolyploids (Soltis and Soltis, 2000). Nearly all polyploid species of plants that have been examined with molecular markers have been shown to be polyphyletic, having arisen multiple times from the same diploid species (Soltis and Soltis, 1999). Multiple origins of polyploids are the rule and not the exception (Soltis and Soltis, 1995).

*Trapogon mirus* and *Trapogon miscellus* formed independently 12 and 20 times respectively (Soltis et al, 1995). *Heuchera grossulariifolia* has five origins and *Draba norvegica* has 13 origins (reviewed by Rauh, 2003).Other examples include *Madia*, *Guttierrezia*, *Mimulus* and *Rubus* (Soltis et al, 2003). Recurrent polyploidy is also seen in arctic plant species: *Saxifraga osloensis, Parnassia palustris,* and *Vaccinium uliginosum* - a complex example of recurring tetraploidy. Recent molecular, ploidal and morphological analyses of populations from the entire distribution range have provided strong evidence for at least three and probably more independent origins of tetraploids. *Dupontia* (Poaceae) probably arose via hybridization and polyploidization. The higher polyploids have probably arisen within the genus itself independently. More complex histories of reticulations are found in *Cerastium*, consists of two highly divergent dodecaploid lineages, which probably have different polyploid origins and now are recognized as two distinct taxonomic species (*C. arcticum* and *C. nigrescens*) (Brochmann et al, 2004).

The frequent recurrence of polyploidization has major evolutionary implications. It suggests that polyploids are much more genetically dynamic than formerly envisioned.

(Soltis and Soltis, 1995). Multiple origins of a polyploid species not only affect patterns of genetic variation in natural populations, but also contribute to differential patterns of gene expression and may therefore play a major role in the long-term evolution of polyploids (Soltis et al, 2004a).

The multiple independent formations of polyploid species from heterozygous diploid progenitors may provide a significant source of genetic variation (Soltis and Soltis, 1999; 2000).

# 3) Factors promoting polyploidy

## 3.1 Modes of polyploid formation

Several mechanisms may lead to polyploidy in plants. Somatic doubling at the zygotic, embryonic, or meristematic stage of a plant's life cycle will ultimately lead to the production of polyploid tissues and possibly the generation of polyploid offspring. Although many examples of polyploidy via somatic doubling have been reported, this mechanism now seems less common than gametic nonreduction, or the production of unreduced gametes, as a means of polyploidy formation in natural populations (Soltis et al, 2003, Soltis et al, 2004b). Unreduced gametes have been reported in a number of species, most notably those that also produce polyploids. Both auto- and allopolyploids can arise in one step after unreduced gamete formation by the union of two unreduced gametes from the same plant or different plants (Soltis et al, 2004b). Alternatively, the production of either an auto or allotetraploid may involve a triploid bridge in which triploids are formed within a diploid population, and backcrossing to diploids (3x gamete x haploid gamete) produced by a single individual produces a tetraploid. This two-step method has been considered a significant pathway to polyploidy formation, but some have suggested that the one step process involving the union of two unreduced gametes may be more common than often considered. The triploid bridge may indeed offer a significant pathway to polyploidy because triploids may not be completely sterile (Soltis et al, 2003).

Hybrids between well differentiated races or species frequently exhibit superior vigor or viability or have enhanced physiological homeostasis. In general, allopolyploids are more

vigorous than diploids (Singh, 1993). They have greater buffering in their genotypes as compared with diploids, owing to the presence of numerous duplications (Soltis et al, 2003).

## 3.2 Apomixis

Many allopolyploids are apomictic. Apomixis arises in response to the near complete sterility of interspecific hybrids. Such sterile plants are able to reproduce asexually through apomixis (Singh, 1993). Several polyploid complexes are characterized by apomixis, which is defined broadly as the replacement of sexual by asexual reproduction. Sometimes this replacement is through the presence of vegetative buds or leafy rosettes in positions where flowers would normally be expected (Stebbins, 1971).

## 3.3 Increased Heterozygosity

Allopolyploids exhibit fixed heterozygosity, representing permanent hybrids and possessing the genes and gene products of both parental diploids. They have the capacity for expressing unique gene product combinations which confer novel genotypes. Permanent heterozygosity allows inbreeding or self fertilization without suffering inbreeding depression (Pikkard, 2001). They have the capacity to produce novel heterodimeric proteins not formed by either diploid parent. Fixed heterozygosity could increase the biochemical pathways and perhaps provide an advantage compared to diploid progenitors (Soltis et al, 2003). Populations of autopolyploids are expected to maintain higher levels of heterozygosity than do the diploids, and these higher heterozygosities can be attributed simply to polysomic inheritance. For example, assuming simple tetrasomic segregation, selfing of a heterozygous autotetraploid of genotype aabb is expected to produce progeny in the ratio of 1 aaaa: 34 heterozygotes (of various genotypes):1 bbbb, a huge increase over expectations for a diploid with disomic inheritance (i.e., 1aa:2ab:1bb). Polyploids exhibit polysomic inheritance; consequently, these polyploids likely also maintain higher levels of heterozygosity than do their diploid parents, simply because of their mode of inheritance (Soltis and Soltis, 2000).

Allotetraploids in *Tragopogon* had fixed heterozygosity at isozyme loci, representing the combination of divergent genomes. In the allotetraploid *Tragopogon mirus* and *Tragopogon miscellus*, 33% and 43%, respectively, of the loci examined were duplicated.

These values are typical of the levels of duplicated loci observed in many allotetraploid plants. An allotetraploid derivative of two allozymically similar parents would display lower apparent levels of duplicated loci and fixed heterozygosity than would a derivative of more genetically divergent parents. However, even in those cases where there is no apparent allelic divergence between the parental genomes, the chromosomal segment is still duplicated. All allopolyploid individuals are essentially heterozygous through nonsegregating, fixed heterozygosity (Soltis and Soltis, 2000).

### 3.4 Ecological factors

Views on the fitness of diploid hybrids have also focused on the availability of novel habitats for hybrid derivatives. While some models hold that hybrids are consistently less fit than their parents, the 'bounded hybrid superiority model' views hybrids as less fit than their parents in the parental habitats but more fit than either parent in other habitats. Biologists have long recognized the role of disturbance in providing new habitats for hybrids (the 'hybridized habitat'), but undisturbed areas may also have open habitats for hybrids. Several studies suggest evidence of habitat differentiation among cytotypes. Greater variability in polyploids for morphological, demographic, and phenotypic traits relative to their diploid progenitors is believed to contribute to habitat differentiation (Soltis et al, 2004b).

Although new polyploids may have reduced fertility, the fertility of early generation polyploids increases rapidly. The new polyploids must contend with the challenge of establishment among typically larger numbers of their diploid progenitors. A new polyploid may persist by replacing its diploid parent, either as a result of an unstable equilibrium between the cytotypes and the stochastic loss of a small diploid population or by out competing it. A new polyploid may coexist with its diploid parent as the result of habitat differentiation immediately following the origin of the polyploid. Replacement and habitat differentiation should not be mutually exclusive. For example, habitat differentiation could contribute to the initial establishment of a new polyploid, but because of higher fitness; this neopolyploid may be able to replace one or both of its diploid progenitors, at least in some areas. In general, greater variability in polyploids for morphological, demographic, and phenotypic traits relative to their diploid progenitors is believed to contribute to habitat differentiation (Soltis et al, 2003).

When the products of hybridization are exposed to a rapidly changing environment, in which many new ecological niches are being opened up, some of these new combinations are highly likely to be better adapted to these new conditions than are any genotypes present in the old established populations. Polyploidy serves the purpose of stabilizing these valuable new genotypes, both by reducing the amount of genetic segregation, and by eliminating the sterility which exists in hybrids between well differentiated species. In addition, many individual polyploid genotypes have phenotypes which are able to tolerate a wide range of environmental conditions. The increased size of certain organs, particularly seeds, which accompanies polyploidy may also help in the process of stabilization and establishment in new habitats, since it increases seedling vigor (Stebbins, 1971).

### 3.5 Effects of plant polyploidy on plant animal interactions

Recent studies indicate that plant polyploidy can have profound effects on interactions with animal herbivores and pollinators. Current data show marked differences in patterns of attack by herbivores on diploids vs. polyploids (Soltis et al, 2003).

Coordinated studies of the effects of polyploidy on interactions with insect herbivores and pollinators indicate that polyploidy certainly affects patterns of attack and visitation, respectively. Studies of *Heuchera grossulariifolia* from the northern rocky mountains of the USA have demonstrated that autotetraploid lineages of independent origin all experience higher levels of attack by the moth *Greya politella* than do sympatric or parapatric diploid plants (Thompson et al, 2004).

Polyploidy in *H. grossulariifolia* has major effects on visitation by pollinators. Visits by bees to plants of each ploidy differ greatly among bee species. For example, *Lasioglossum* bees constituted about one-quarter of the visits to diploids but only one-tenth of the visits to tetraploids. The moth *G. politella* visited tetraploid flowers five times more frequently than diploid flowers. The bee-fly *Bombyllius major* visited tetraploid flowers six times more frequently than diploid flowers. (Thompson et al, 2004).

## 4) Genome changes following polyploidization

Allopolyploidization is a revolutionary event through which a new species is formed in one step. It generates two genomic "shocks" on the newly formed allopolyploid species:

hybridity, in which two divergent genomes are joined together to form one nucleus; and polyploidy, resulting in duplicated genomes. In response to these two unanticipated shocks, the genomes of the newly formed allopolyploids react in a burst of irreversible genomic reorganizations and modifications. These changes include, among others, structural rearrangements on the chromosome level and sequence level, regulation of gene expression, activation of transposons' and amplification, reassortment, or elimination of highly repetitive sequences, and low-copy sequences (Ozkan et al, 2001).

The evolutionary fate of duplicate genes in diploid organisms has been a topic of interest for some time. In polyploids, because every gene is duplicated, the fate of these genes is even more intriguing. Recent studies have focused on the evolutionary dynamics of duplicated genes within polyploid species and the rates of molecular evolution in polyploids compared to their diploid relatives (Soltis et al, 2004b).

## 4.1 Diploidization

There is an increase in bivalent formation as a result of selection so that established polypoids become identical to diploids in terms of chromosome behavior, i.e., they become diploidized. The process due to which allopolyploids show diploid like behaviour, i.e., only bivalent formation, is known as diploidization (Singh, 1993). Old polyploids tend to be more diploid like than newly formed polyploids (Soltis et al, 2003).

Because most ancient polyploids have undergone an evolutionary process of chromosomal and perhaps genic dipoidization, their polyploidy history may be obscured at the cytological and classic genetics levels (Soltis et al, 2003).

Diploidization makes ancient rounds of chromosome doubling difficult to detect (Feldman et al., 1997). As a result, the polyploid nature of many plant genomes was not evident until the advent of comparative genomics and whole genome sequencing. Recent and prominent examples include maize and Arabidopsis- both species were traditionally recognized as diploids, but in fact their genomes harbor compelling evidence of historical cycles of genome doubling. As can be seen from these and other examples from plants, it is probably safe to state that there are no strictly diploid species in the plant kingdom (Liu and Wendel, 2002).

### 4.1.1 Genetic diploidization

Genetic diploidization refers to the phenomenon whereby expression levels in a polyploid are reduced to those of one of the diploid progenitors, by either gene silencing or dosage compensation. (Liu and Wendel, 2002).

#### *4.1.1.1 Epigenetic changes*

Genetic changes are based on alteration of the DNA sequence itself, resulting in permanent changes in DNA sequence or gene loss (Soltis et al, 2003), whereas the term epigenetic refers to changes in gene expression that do not entail a change in DNA sequence. Among the several mechanisms responsible for epigenetic phenomena, the most important appear to be DNA methylation and histone deacetylation (Liu and Wendel, 2002). It also includes RNA interference and dosage compensation (Soltis et al, 2003).

An important recent development is the widespread appreciation for the role that epigenetic changes may play in polyploid evolution (Soltis et al, 2003).

#### *4.1.1.2 Gene silencing*

Allopolyploidy often is accompanied by epigenetic gene silencing (Liu and Wendel, 2002). Several models of genome evolution, in which a polyploid genome gradually will undergo gene silencing and return to a diploid condition, have been presented (Soltis and Soltis, 2000).

Gene silencing is one of the processes by which expression levels in a polyploid are reduced during genetic diploidization to those of one of the diploid progenitors (Liu and Wendel, 2002). The stabilization of allopolyploids involves genetic diploidization. Duplicate genes are either silenced or expressed at reduced levels (Ozkan et al, 2001). Changes in cytosine methylation, histone modifications and positional effects from a higher order change in chromatin structure can play a role in silencing (Adams and Wendel, 2005).

Epigenetic gene silencing can affect a variety of genes with different biological functions. Although the silencing events accompany the onset of allopolyploid formation, reversibility for at least some of the expression changes is maintained in natural allopolyploids with the same genomic composition, suggesting evolutionarily stable

epigenetic transformations. This in turn suggests selection to stabilize the epigenetic regulatory response (Liu and Wendel, 2002).

Silencing arises immediately upon polyploid formation and it is epigenetically controlled since there is insufficient time for point mutations to accumulate (Adams and Wendel, 2005). Gene silencing by epigenetic instead of mutational means in a natural allopolyploid provides the potential for developmental or tissue specific reversibility, which may provide an evolutionarily adaptive regulatory flexibility (Liu and Wendel, 2002).

Some duplicated genes are silenced immediately upon allopolypoidy in some organs of the plant but remain expressed in other organs at varying levels (Adams and Wendel, 2005).

Transposable elements appear to be the causal agents for gene silencing in some instances (Soltis et al, 2003).

Gene silencing is in some case stochastic and in some cases directed. The scale of the phenomenon of duplicate gene silencing suggests that this is a significant aspect of polyploid evolution (Adams and Wendel, 2005).

Duplicate gene silencing in polyploids need not necessarily be an epigenetic phenomenon. Models for homoeologous gene silencing, involving repeats and long terminal repeats of retroelements have been proposed (Adams and Wendel, 2005).

## 4.2 Diploidization and homoeologous chromosomes (chromosomal diploidization)

Two sets of chromosomes in a diploid hybrid may be sufficiently different from one another that they fail to pair correctly at meiosis. Chromosome doubling and the creation of a polyploid thus solve the immediate problem of meiotic pairing of homologous chromosomes in a hybrid organism; each chromosome can pair with its own duplicate. However, the polyploid organism faces another immediate problem during meiosis: the two sets of homoeologous chromosomes (i.e., from different parental stock) may be sufficiently similar to one another that pairing of homoeologous chromosomes may occur, disrupting the correct pairing of the new truly homologous (i.e., duplicated) chromosomes (Feldman et al., 1997).

The stabilization of allopolyploids involves cytological or chromosome diploidization, in which meiotic pairing of homoeologous chromosomes (genetically related chromosomes of different genomes) is suppressed, thus leading to a pairing pattern similar to that of diploids, that is, exclusive bivalent pairing of homologous chromosomes (Ozkan et al, 2001). The process of chromosomal diploidization serves to accentuate differences between homeologous chromosomes in a polyploid organism, facilitating correct pairing of homologous chromosomes during meiosis (Feldman et al., 1997).

Because allopolyploidy usually entails the merger of genomes that are sufficiently homologous that homoeologous chromosomes may pair during meiosis, the earliest generations in allopolyploid formation must experience strong selection for exclusive bivalent pairing. This is evidenced in modern allopolyploids by the common observation that homoeologous pairing is much lower than one might expect from chromosome associations formed in hybrids between the extant parental diploids. One of the best examples of this phenomenon is wheat, where both tetraploids and hexaploids exhibit exclusive bivalent formation (Liu and Wendel, 2002).

Polyploidy-induced differential elimination of genome specific sequences facilitates homologous chromosome pairing (Shaked et al, 2001, Liu and Wendel, 2002).

Logically, polyploids should have larger C- values than diploids, with the C- values of polyploids increasing in direct proportion to ploidal level. This expectation holds true in synthetic polyploids and newly formed polyploids. For example, the newly formed allotetraploids *Tragopogon miscellus* and *T.mirus* have C- values that are additive of their diploid progenitors (Soltis et al, 2003). However, Leitch and Bennet, 2004 found that the mean 1C DNA amounts tended to decrease with increasing ploidal level. This shows DNA loss following polyploid formation. It can be explained by homoeologous recombination leading to DNA loss and elimination of specific DNA sequences can lead to reduction in DNA amounts in plants (Leitch and Bennett, 2004).

## 4.3 Functional fates of duplicated genes

Much of the genetic redundancy created by ancient polyploidy vanishes through gene loss but some duplicated gene pairs are retained over millions of years. This is evidence for a functional role of duplicated genes and divergence following duplication (Adams and

Wendel, 2005). At the gene level, there are several potential functional fates of duplicated genes; functional redundancy (both genes retain original function), subfunctionalization (combined activity of both genes), neofunctionalization (one gene copy retains original function) and pseudogene formation (second gene copy non functional) (Rauh, 2003).

Duplicate gene retention is possible by neo functionalization and subfunctionalization (Adams and Wendel, 2005).

Reciprocal silencing i.e. one copy silenced in some organs and the other copy in other organs) found in cotton polyploids provides an evolutionarily recent example of subfunctionalization (Adams and Wendel, 2005).

The genetic and epigenetic changes triggered by hybridization and genome duplication may have important phenotypic consequences and appear as a key feature determining the adaptive success of newly formed allopolyploid species (Salmon et al, 2005).

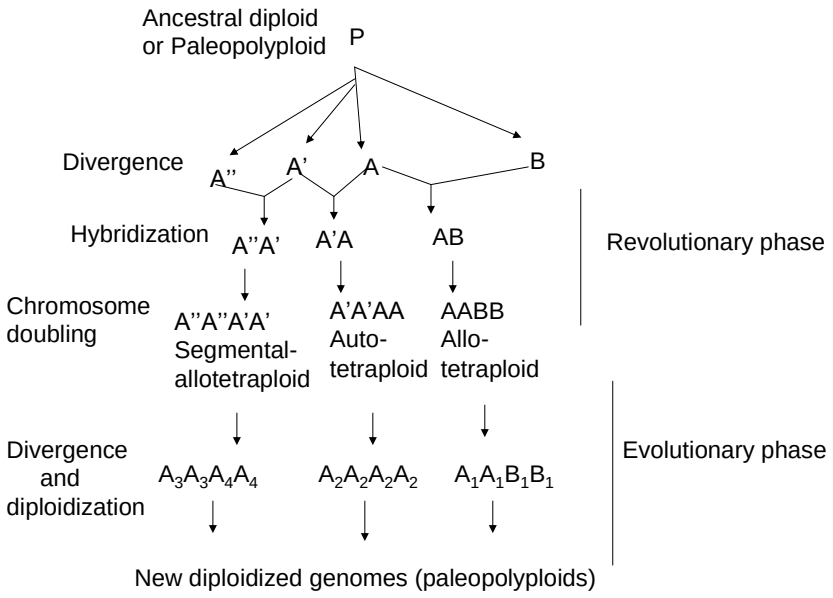


Figure 2 **Evolution** under polyploidy. Divergence from an ancestral species, which can be diploid or paleopolyploid, generates species with related genomes. In the example shown, genome A is very closely related to A' and the chromosomes of A and A' can still pair in an AA' hybrid. A and A" share partial similarity with some chromosomes that can pair and other chromosomes that are more diverged and do not pair. A and B are more distant and, therefore, their chromosome cannot pair and the hybrid AB is sterile. Upon chromosome doubling, hybrids AA', AA", and AB give rise to an autopolyploid, a segmental

allopolyploid, or a genomic allopolyploid, respectively. A series of genetic and epigenetic changes occur in the early phase of the hybrid and/or after chromosome doubling (the revolutionary phase). Later on (the evolutionary phase), the genomes of the polyploids keep diverging and the similarity to their diploid progenitors decreases but can still be detected by cytological analysis. Note that at this stage, lateral gene flow between species is possible in polyploids that share one genome. With further diploidization, the traces of the original genome duplication can be detected only by a thorough sequence analysis. By the end of this process, new diploid, or paleopolyploid, genomes are formed. Source: Levy and Feldman, 2002

# 5) Model plants for the study of polyploidy

Several groups of plants have served as models of the study of polyploidy.

Emphasis is placed upon the processes which take place in their origin and evolution, rather than in classification and categorization of individual polyploids or groups of them (Stebbins, 1971).

## 5.1 Ancient polyploids (paleopolyloids)

### 5.1.1 Crops of the grass family (Poaceae)

Most, if not all, grasses are polyploids; all major types of polyploids, namely autopolyploids, segmental allopolyploids, and allopolyploids, can be found in this family.

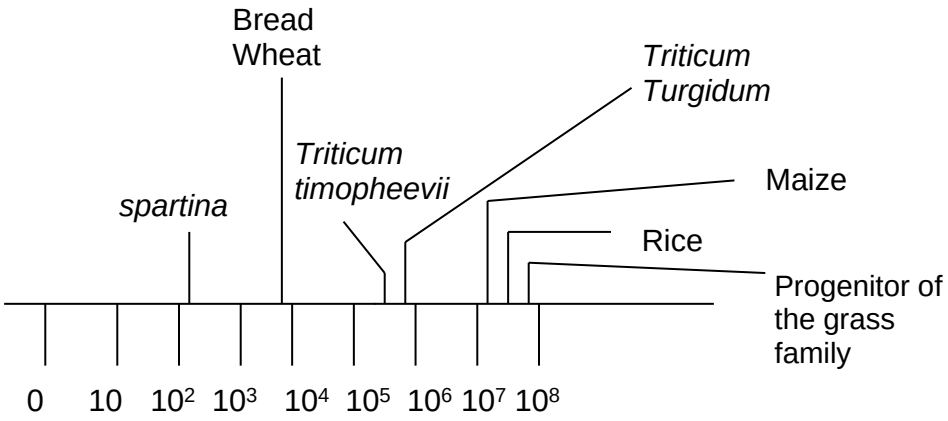


Figure 3. The estimated number of years (age) since polyploidy occurred for various grass species. Source: (Levy and Feldman, 2002)

A basic chromosome number of x = 12, is a derivative of lines that underwent genome duplication some time during their evolutionary history (Stebbins, 1971). All the primitive grass subfamilies have a chromosome number of x=12, this would imply the ancestor of the grasses was itself a polyploid. In accordance, all the diploid Poaceae species would be paleopolyploids. The family also contains a number of species (>60%), distributed in all clades, that are classified as polyploids, i.e., underwent an additional cycle of chromosome doubling (neopolyploids) (Levy and Feldman, 2002). In these neopolyploid species, the duplicated genomes did not diverge much from their "diploid" progenitors and chromosome number and cytological behavior are still indicative of genome duplication. Most of these species were derived from distant interspecific or intergeneric hybridizations, giving rise to new allopolyploid species and the remaining derived from intraspecific close interspecific hybridizations giving rise to autopolyploid cytotypes (Stebbins, 1971).

The grasses represent a range of genome size and structural complexity, with rice on one extreme. A diploid with n = 12 chromosomes ($2n$ = 24), rice has one of the smallest plant genomes, with only 0.9 pg of DNA per 2C nucleus (Fig.4). Other grass species exhibit far

larger genomes. Bread wheat, for example, is a hexaploid with n = 21 chromosomes (2*n* = 42) and a haploid DNA content of 33.1 pg. Genera like *Saccharum* (sugarcane) and *Festuca* are even more complicated, displaying wide variation in ploidy level and over 100 chromosomes in some species. As a diploid with n = 10 chromosomes (2*n* = 20) and a 2C genome content roughly 6-fold larger than rice, maize lies somewhere in the middle of grass genome size and structural complexity (Fig.4) (Gaut et al, 2000).

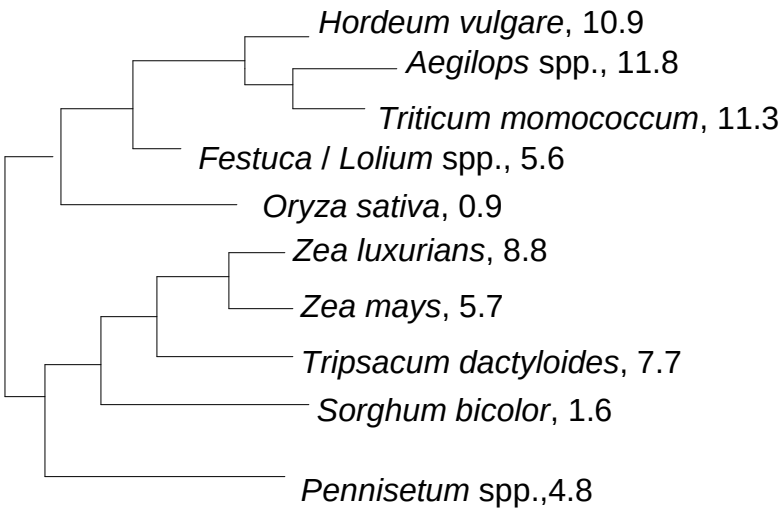


Figure 4 A phylogeny of diploid grass species. Numerical values next to species names represent the 2C genome content of the species, measured in picograms. Source: Gaut et al (2000)

Grass crops such as wheat (*Triticum* spp.), maize (*Zea mays*), and rice are well-studied genetic systems, each corresponding to a prototype polyploid (allopolyploid, segmental allopolyploid, and paleopolyploid, respectively); all types of synthetic polyploids can be readily made, enabling one to mimic and study the genetic and epigenetic events that occurred early at the onset of polyploidy. All the above makes grasses one of the best systems to study polyploidy (Levy and Feldman, 2002).

#### *5.1.1.1 Wheat*

Bread wheat (*Triticum aestivum*) represents one of the best examples of genome evolution through allopolyploidy. It is an allohexaploid (genome BBAADD) containing

n =21 chromosomes that arose from hybridization, followed by chromosome doubling, between a tetraploid (genome BBAA, n = 14) and a diploid (genome DD, n =7) species about 9,500 calibrated years ago. Two of its three diploid progenitors have been identified: *Triticum urartu* (genome AA) and *Aegilops tauschii* (genome DD). The origin of the B genome remains elusive. The diploid species whose genome is the closest to the B genome is *Aegilops speltoides* (Fig 5) The 21 pairs of bread wheat chromosomes were classified into seven homoeologous groups based on the ability of a trisome (three doses of a given chromosome) to compensate for the absence of each of the other two chromosomes of the same group. This work was the first evidence for syntenic relationships among the grass genomes (Levy and Feldman, 2002).

Polyploidy is widespread among the wheat (*Aegilops* and *Triticum* genera) group. For example, bread wheat (*T. aestivum*) is a hexaploid that arose from successive rounds of chromosome doubling after hybridization between various species of *Aegilops* and *Triticum* (Figure 5). The observation that chromosome doubling apparently occurred at least twice in the evolutionary history of bread wheat underscores the notion that polyploidy is a common and frequent occurrence in wheat (Soltis and Soltis, 1995).

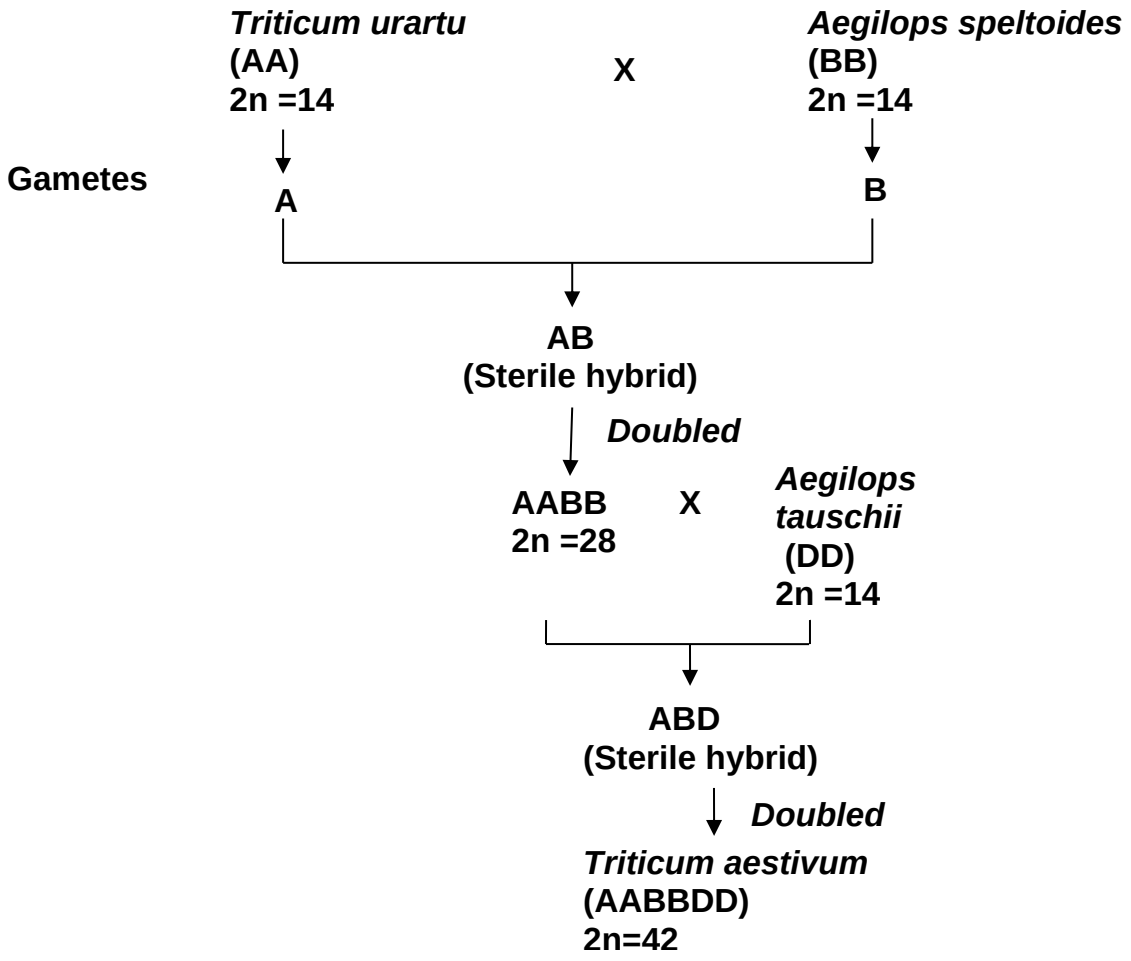


Figure 5 Recent views about the origin of bread wheat. Source: Levy and Feldman (2002)

Despite their genetic relatedness, homoeologous wheat chromosomes do not pair under normal conditions. However, in the absence of the homoeologous pairing suppressor gene Ph1, pairing and recombination between homoeologs can occur. Suppressing inter-genomic pairing is important in maintaining the disomic inheritance and, as a consequence, the fertility of wheat, as well as facilitating permanent heterozygosity between homoeoalleles (Levy and Feldman, 2002).

Ozkan et al (2001) studied two dominant genes, *Ph1* and, to a lesser extent, *Ph2*, were considered to be responsible for the cytological diploidization by suppressing homoeologous pairing. However, *Ph1*-like genes were not found in polyploids of the related genus *Aegilops*, which exhibits diploid-like meiotic behavior. Moreover, plants of hexaploid wheat deficient for *Ph1* exhibit much lower levels of homoeologous pairing than

was expected from the level of pairing in hybrids among the three diploid progenitors. This fact indicates that diploid-like meiotic behavior in polyploid wheat depends on factors other than *Ph1* (Ozkan et al, 2001).

Ozkan et al. (2001) analyzed the elimination of eight low-copy DNA sequences in diploid $F_1$ progeny and in derived allopolyploids formed from hybridization between various *Aegilops* and *Triticum* species. The sequences analyzed are present in all diploid wheat species but occur in only one genome in polyploid wheat, either as a single homologous pair (chromosome-specific sequence) or in several pairs on more than one chromosome (genome-specific sequence) (Ozkan et al, 2001).

Sequences such as GSSs (genome specific sequences) or CSSs (chromosome specific sequences) do not have homoeologous counterpart (Levy and Feldman, 2004). Feldman et al (1997) took GSS and CSS sequences isolated from a library of DNA obtained through micro dissection of the long arm of chromosome 5B to understand the origin of these sequences. The study showed that diploid relatives of wheat contained sequence homologous to sequences that are present in only one genome or chromosome of tetraploid or hexaploid wheat. This indicated that these sequences were lost in the course of wheat evolution (Feldman et al 1997).

Ozkan et al (2001) studied the same CSSs and GSSs used by Feldman et al (1997) and found that rapid elimination of CSSs and GSSs is a general phenomenon in newly synthesized allopolyploids. Elimination of GSSs was already initiated in $F_1$ plants and was completed in the second or third allopolyploid generation, whereas elimination of CSSs started in the first allopolyploid generation and was completed in the second or third generation. Sequence elimination started earlier in allopolyploids whose genome constitution was analogous to natural polyploids compared with allopolyploids that do not occur in nature. Elimination is a nonrandom and reproducible event whose direction was determined by the genomic combination of the hybrid or the allopolyploid. It was not affected by the genotype of the parental plants, by their cytoplasm, or by the ploidy level, and it did not result from intergenomic recombination. Allopolyploidy-induced sequence elimination occurred in a sizable fraction of the genome and in sequences that were apparently noncoding. This finding suggests a role in augmenting the differentiation of

homoeologous chromosomes at the polyploid level, thereby providing the physical basis for the diploid-like meiotic behavior of newly formed allopolyploids (Ozkan et al, 2001).

In a similar study, Shaked et al (2001) surveyed $F_1$ hybrids between diploid species from the wheat (*Aegilops* and *Triticum*) group and their derived allotetraploids and found that sequence elimination is one of the major and immediate responses of the wheat genome to wide hybridization or allopolyploidy, that it affects a large fraction of the genome, and that it is reproducible; that is, the same loci were always involved in the same patterns of sequence elimination. Their results show that sequence elimination was a widespread and immediate response to allopolyploidization among these wheat species. Furthermore, it often followed a nonrandom, reproducible pattern characterized by preferential elimination of sequences from one of the parental genomes. 15 – 14 % of the loci were eliminated (Shaked et al, 2001).

Elimination of non-coding sequences augments the differentiation of homoeologous chromosomes at the polyploid level, thus increasing the physical divergence between homoeologues and contributing to the diploid-like meiotic behaviour of polyploid wheat (Levy and Feldman 2004).

Nonrandom sequence elimination leads to diploidization and a more well-behaved meiotic system (Feldman et al. 1997). This can be seen from the data from wheat (Ozkan et al., 2001; Shaked et al., 2001) and *Brassica* species (Song et al., 1995).

Sequence elimination in synthetic amphploids leading might be a faithful repetition of the events that occurred in nature at early stages of allopolyploid wheat formation (Levy and Feldman, 2004).

Shaked et al (2001) studied epigenetic alterations. They analyzed alterations in patterns of cytosine methylation throughout the genome using methylation-sensitive amplification polymorphism (MSAP) (Shaked et al., 2001), a method derived from AFLP but using a pair of isoschisomer restriction enzymes, each with a different sensitivity to cytosine methylation (Levy and Feldman, 2004). Shaked et al (2001) analyzed a large and unbiased set of polymorphic loci in three diploid wheat species, the $F_1$ progeny of interspecific hybridizations, and derived synthetic allotetraploid progeny using amplified fragment length polymorphism (AFLP) and methylation-sensitive amplification polymorphism (MSAP) fingerprinting.

Ozkan et al. (2001) and Shaked et al. (2001) describe the genetic and epigenetic rearrangements that occur in $F_1$ interspecific and intergeneric hybrids between various species of the wheat (*Aegilops* and *Triticum*) group and in the subsequent generations (S1, S2, and S3) of the derived amphiploids. DNA elimination and methylation were found to be rapid and major responses of the genome to wide hybridization and to allopolyploidy in wheat (Shaked et al, 2001, Ozkan et al, 2001).

From the studies of wheat by Levy, Feldman and co workers in the past few years, it was found that allopolyploidization brings about rapid genome evolution through the instantaneous generation of a variety of cardinal genetic and epigenetic alterations comprising: (1) non-random elimination of coding and non-coding DNA sequences, (2) epigenetic changes such as DNA methylation of coding and non-coding DNA leading, among others, to gene silencing, and (3) activation of retroelements, which in turn alters the expression of adjacent genes (reviewed by Levy and Feldman, 2004).

These changes were similar to those that occurred twice in nature: first, at the transition from diploid to tetraploid wheat (~0.5 Mya) and, second, at the transition from tetraploid to hexaploid wheat (~9500 years ago) (reviewed by Levy and Feldman, 2004).

Transcriptional and post-transcriptional alterations of gene activity, including transcriptional activation of retroelements, led to novel expression patterns. These phenomena emphasize the plasticity of the genome with regard to both structure and gene expression. This plasticity in turn might improve the adaptability of the newly formed allopolyploids and facilitate their rapid and successful establishment in nature (reviewed by Levy and Feldman, 2004).

#### *5.1.1.2 Maize*

Maize is thought to be an ancient segmental allotetraploid. The evidence for this is based on chromosome number and molecular analysis: Maize has n = 10 chromosomes, compared with n = 5 for species in its tribe (*Andropogoneae*). Moreover, sequence analysis of 14 pairs of duplicated genes revealed two sets of gene pairs: those that diverged approximately 20 mya (million years ago) and those that diverged approximately 11 mya. A likely explanation for such data is a segmental allotetraploid origin, whereby some chromosomes did not recombine in the early stages of polyploidy and duplicated sequences on these chromosomal segments have

the age of the common ancestor (approximately 20 mya), whereas on chromosomes that did recombine, the duplicated sequences have the age of the polyploidy formation (approximately 11 mya). This analysis has been extended to a large number of loci (from the UMC98 map) using new computational approaches. It reveals that up to 80% of the maize genome remains organized in colinear regions (Levy and Feldman, 2002).

Gaut and Doebley (1997) compared the pattern of divergence to patterns predicted by four models of the evolution of the maize genome autotetraploidy, genomic allotetraploidy, segmental allotetraploidy, and multiple segmental duplications. Their analyses indicate that coalescent times for duplicated sequences fall into two distinct groups, corresponding to roughly 20.5 and 11.4 million years. This observation strongly discounts the possibility that the maize genome is the product of a genomic allotetraploid event, and it is also difficult to reconcile with either autotetraploidy or multiple independent segmental duplications. However, the presence of two (and only two) coalescent times is predicted by the segmental allotetraploid model. If the maize genome is the product of a segmental allotetraploid event, as these data suggest, then its two diploid progenitors diverged roughly 20.5 Mya, and the allotetraploid event probably occurred approximately 11.4 Mya (Gaut and Doebley, 1997).

The first hints of the complex organization of the maize genome came from cytological studies. Although maize is diploid, early studies by McClintock demonstrated the association of nonhomologous chromosomes during meiosis. Later studies documented the formation of bivalents and multivalents in maize haploids. Altogether, cytological observations suggested that the maize genome contains extensive regions of homology, probably reflecting chromosomal duplications. Evidence for chromosomal duplication also came from linkage information (Gaut et al, 2000).

The hypothesized origin of the maize genome states that (i) maize is the product of a segmental allotetraploid event, (ii) the two diploid progenitors (or "parents") of maize diverged ≈20.5 mya, (iii) the tetraploid event occurred between 16.5 and 11.4 mya, sometime after the divergence of *Sorghum* from one of the progenitor lineages, and (iv) the genome "rediploidized" before 11.4 mya. The hypothesis is based on a relatively small

number of DNA sequences i.e., only 14 pairs of duplicated sequences. Extant maize is a diploid, and thus the segmental allotetraploid hypothesis presumes that the maize genome rearranged and diploidized (Gaut et al, 2000).

Regions of chromosomal duplication in maize have been documented. Portions of chromosome 1 are duplicated on chromosomes 5 and 9, portions of chromosome 2 are also found on chromosomes 7, 10. A great number of species contain chromosomal duplications. Even species with streamlined genomes contain chromosomal duplications; for example, rice has a large duplication between chromosomes 11 and 12 and *Arabidopsis* also has at least one large chromosomal duplication. Other plant genomes with chromosomal duplications include sorghum, cotton, soybean, and *Brassica* species. Some of these genomes are degenerate polyploids like maize, but others may owe their chromosomal duplications to independent segmental events (Gaut et al, 2000).

#### *5.1.1.3 Rice*

According to Paterson et al., (2003) rice is a paleo-polyploid across its entire genome i.e. it has whole genome duplicatoin (Paterson et al., 2003). But according to Vandepoele et al. (2003), rice is ancient aneuploid that has experienced the duplication of one or a large part of one chromosome in its evolutionary past. Based on their own early assembly of the unfinished genome sequence, Vandepoele et al. (2003) reported the duplication of rice chromatin (Vandepoele et al 2003).

Wang et al. (2005) found evidence generally supporting the occurrence of whole-genome duplication. They found 10 duplicated blocks. A paleopolyploidy (PPP1) or "whole-genome duplication" in rice was suggested which shows ancient genome duplication (Wang et al, 2005).Yang et al, 2005 reported another potentially older large-scale duplication (PPP2) event that predates monocot-dicot divergence in the genome of rice (*Oryza sativa* L.), as inferred from the age distribution of pairs of duplicate genes based on recent genome data for rice. Their results suggest that paleopolyploidy was widespread and played an important role in the evolution of rice (Yang et al, 2005).

### 5.1.2 Cotton

*Gossypium* provides one of the textbook examples of speciation via allopolyploidy (Liu et al, 2001). Synthetic *Gossypium* polyploids show few changes in overall genome sequences, yet they display differential expression of genes in different tissue types (Liu et al. 2001; Adams et al. 2003).

Indications of genomic additivity and epigenetic stasis during allopolyploid formation of cotton provide a contrast to recent evidence from several synthetic plant allopolyploids, most notably wheat, *Brassica* and an *Arabidopsis* polyploid, in which rapid and unexplained genomic changes, as well as methylation changes, have been reported (Reviewed by Adams and Wendel, 2004). AFLP analysis was performed to evaluate nine sets of newly synthesized allotetraploid and allohexaploid plants, their parents, and the selfed progeny from colchicine-doubled synthetics. Using both methylation-sensitive and methylation insensitive enzymes and genomic Southern analysis on six sets of synthetic allopolyploids probed with five retrotransposons indicated genomic additivity and epigenetic stasis during allopolyploid formation. This provides a contrast to recent evidence from several model plant allopolyploids, most notably wheat (Ozkan et al 2001 and *Brassica* (Song et al 1995), where rapid and unexplained genomic changes have been reported. In addition, the data contrast with evidence from repetitive DNAs in *Gossypium*, some of which are subject to non-Mendelian molecular evolutionary phenomena in extant polyploids. These contrasts indicate polyploid speciation in plants is accompanied by a diverse array of molecular evolutionary phenomena, which will vary among both genomic constituents and taxa (Liu et al, 2001).

Adams and Cronn (2003) assayed expression of 40 homoeologous gene pairs (1 to 2 million years old) in synthetic tetraploid cotton. They documented silencing or unequal expression of one homoeolog for 10 of 40 genes examined in ovules of *Gossypium hirsutum*. Assays of homoeolog expression in 10 organs revealed variable expression levels and silencing, depending on the gene and organ examined (Adams and Cronn, 2003). Remarkably, silencing and biased expression of some gene pairs are reciprocal and developmentally regulated, with one homoeolog showing silencing in some organs and the other being silenced in other organs, suggesting rapid subfunctionalization (Adams and Cronn, 2003; Adams and Wendel, 2005). Analysis of a synthetic tetraploid revealed homoeolog expression and silencing patterns that sometimes mirrors those of the natural tetraploid. Both long-term and immediate responses to polyploidization were implicated. Their data suggest that some silencing events are epigenetically induced during the allopolyploidization process (Adams and Cronn, 2003).

Table 1 Diversity and Geographic Distribution of the Major Lineages of Gossypium. Source: Wendel and Cronn (2003)

| Genome group | Number of species | Geographic distribution |
|---|---|---|
| A | 2 | Africa, possibly Asia |
| B | 3 | Africa, Cape Verde Islands |
| C | 2 | Australia |
| D | 13 | Primarily Mexico, also Peru, Galapagos Islands, Arizona |
| E | 7+ | Arabian Peninsula, Northeast Africa, Southwest Asia |
| F | 1 | East Africa |
| G | 3 | Australia |
| K | 12 | NW Australia |
| AD | 5 | New World tropics and subtropics including Hawaii |

### 5.1.3 Brassica

The *Brassica* crop species have long been model system to study the molecular and phenotypic changes associated with both recent and ancient polyploidization events (Lukens et al, 2004).

The base diploid species *B. rapa* (n=10), *B. nigra* (n=8) and *B. oleracea* (n=9) have an ancient origin (Lukens et al, 2004).

In contrast to the diploid *Brassicas*, the other crop species within the genus i.e. *B. napus* (n=19), *B. juncea* (n=18) and *B. carinata* (n=17) are allotetraploids whose genomes were derived from the recent fusion of two of the base diploid genomes (Singh, 1993). *B. napus* (n=19) can be resynthesized by crossing *B. rapa* (n=10) and *B. oleracea* (n=9), rescuing the embryo and colchicine doubling the resultant amphihaploid. *Brassica juncea* and *B. carinata* can be resynthesized in a similar way by crossing *B. rapa* and *B. nigra* and *B.*

*oleracea* and *B. nigra*, respectively. Several molecular mapping analyses have confirmed the similarity between the base diploid progenitor genomes and the polyploid genomes. For example, linkage groups N1 to N10 of *B. napus* correspond to the A (*B. rapa*) genome, and N11 to N19 correspond to the C (*B. oleracea*) genome (Lukens et al, 2004).

The genomes of the base diploid species are very similar to each other, and all have a high level of genomic redundancy. Researchers have proposed that different mechanisms generated this redundancy. A recent explanation is that the base *Brassica* genome is comprised of three ancestral genomes that are *Arabidopsis thaliana* -like in structure (Lagercrantz, 1998).

The study by Lagercrantz indicates that the evolution of genomes in the *Brassicaceae* family involves an unusually high rate of chromosomal rearrangements. Recent polyploidization is probably an important factor contributing to the rapid rearrangements of *Brassicaceae* chromosomes (Lagercrantz, 1998).

Comparative mapping has indicated that present-day diploid species in the *Brassica* genus and related genera have descended from a common hexaploid ancestor and are, thus, degenerate hexaploids. It seems likely that the replications have occurred through amphidiploidization the same way as novel amphidiploids such as *B. napus*, *B. juncea*, and *B. carinata* are derived from hybridization between different diploids (Lagercrantz and Lydiate 1996). Lagercrantz studied Chromosome organization and evolution in the *Brassicaceae* family using comparative linkage mapping. The data support that modern diploid *Brassica* species are descended from a hexaploid ancestor, and that the *A. thaliana* genome is similar in structure and complexity to those of each of the hypothetical diploid progenitors of the proposed hexaploid. Thus, the *Brassica* lineage probably went through a triplication after the divergence of the lineages leading to *A. thaliana* and *B. nigra*. These duplications were also accompanied by an exceptionally high rate of chromosomal rearrangements (Lagercrantz, 1998).

Comparative mapping of *B. oleracea* (n = 9), *B. nigra* (n = 8), and *B. rapa* (n = 10) support that changes in chromosome numbers caused by chromosome fusion or fission are frequent and also have occurred recently in the *Brassicaceae* family (Lagercrantz and Lydiate 1996).

Song et al. (1995) observed extensive and rapid genome changes after the production of synthetic polyploids derived from various combinations of three *Brassica* species, and concluded that homoeologous recombination was a likely cause of the changes. The results of Song et al. (1995) suggested that the frequency of genomic change associated with polyploidization is correlated positively with the degree of divergence between parental diploid genomes. They examined polyploids derived from two sets of crosses: *B. rapa* (A genome) used as the male or the female parent with *B. nigra* (B genome) to produce AB and BA polyploids, and *B. rapa* as the male or female parent with *B. oleracea* (C genome) to produce AC and CA polyploids. Of the 80 to 90 sequence fragments examined, the AB and BA polyploids showed about twice the number of changes as did the AC and CA polyploids, and the A and B genomes were more highly divergent than the A and C genomes. They detected extensive changes in the nuclear genome of each synthetic polyploidy during each of five generations. Thus, *B. napa* (contributor of the A genome) is more closely related to *B. oleracea* (contributor of the C genome) than to *B. nigra* (contributor of the B genome) (Song et al. 1995).

### 5.1.4 Arabidopsis

*Arabidopsis* genome size is small and genome sequence of *Arabidopsis* is available. Thus it is convenient to study its genome (Blanc et al, 2003). One of the most surprising discoveries revealed by its sequence is that it contains numerous large duplicated chromosomal segments (Blanc et al, 2003).

Blanc et al (2003) concluded that the *Arabidopsis* lineage underwent at least two distinct episodes of duplication. One was a polyploidy that occurred much more recently than estimated previously, before the *Arabidopsis/Brassica rapa* split and probably during the early emergence of the crucifer family (24–40 Mya). An older set of duplicated blocks was formed after the monocot/dicot divergence, and the relatively low level of overlap among these blocks indicates that at least some of them are remnants of a larger duplication such as a polyploidy or aneuploidy (Blanc et al, 2003).

Comai et al (2000) made extensive study of gene silencing and novel expression on synthetic allotetraploid lines obtained from a cross between *Arabidopsis thaliana* (2n=2x= 10) and *Cardaminopsis arenosa* (2n=4x=32). These lines exhibited a great deal of variation in morphology, fertility and flowering time. The high level of phenotypic instability

observed in the $F_2$ generation can only be explained by non-Mendelian genetic processes. These phenotypic instabilities in newly synthesized allopolyploids implicate one or more genetic or epigenetic phenomena. Gene silencing can take place in the $F_2$ generation after allopolyploid hybridization of *A. thaliana* and *C. arenosa*, affecting both putative euchromatic genes and a repeated gene related to transposons. The study showed that 0.4% of the genes in these allotetraploid lines are silenced. The results strongly suggested epigenetic silencing (Comai et al 2000).

Blanc and Wolfe (2004) have reported functional divergence of genes that have been duplicated by polyploidy during the evolutionary history of *Arabidopsis*. They showed that more than half of the gene pairs formed by the most recent polyploidy have significantly different expression patterns. They also provide evidence that 62% of the recently duplicated gene pairs have undergone functional diversification. They found conceted evolution of some duplicate genes (Blanc and Wolfe, 2004)

Madlung et al (2005) produced synthetic allopolyploids using autotetraploid *A. thaliana* and *A. arenosa* as parents. *Arabidopsis suecica* (2n=2x=26) is a natural allopolyploid whose maternal ancestor is *A. thaliana* (2n=2x=10) and whose paternal ancestor is *A. arenosa* (2n=4x=32). They used allopolyploids of *Arabidopsis* to explore potential sources of instability. The newly synthesized *Arabidopsis* allopolyploids, which display phenotypic instability and low fertility, displayed several, possibly related mechanisms that can remodel genomes. They detected transcriptional activity of several transposons although their transposition was limited. These neoallopolyploids display similarity to the natural *A. suecica* in genome constitution and in remodeling of DNA methylation and epigenetic gene silencing. The synthetic allotetraploids displayed instability of selected transposons, genomic rearrangements, and chromosomal abnormalities involving the formation of bridges and chromosomal breaks (Madlung et al, 2005).

Meiosis was analyzed cytogenetically in autotetraploids of *Arabidopsis*, including both established lines and newly generated autotetraploid plants. The new tetraploids showed high multivalent frequencies, exceeding the theoretical 66.66%. The established lines showed fewer multivalents than the new autotetraploids did, but the extent of this reduction was strongly line and chromosome dependent. The established autotetraploid lines have

undergone a partial diploidization of meiosis, but not necessarily genetical diploidization, since their formation (Santos et al, 2003).

The studies conducted on the model plants demonstrated extensive genomic changes immediately following polyploid formation. Collectively, these and other recent studies have drawn attention to the 'dynamic' nature of polyploids (Soltis & Soltis, 1995), and emphasized the relatively poorly understood and sometimes non-Mendelian mechanisms that may characterize gene and genome evolution in polyploids (Liu and Wendel, 2002).

The precise mechanisms whereby DNA elimination occurs or the sensing and signaling leading to the epigenetic alterations and changes in gene expression are unknown and promise to be exciting fields of research. Last, but not least, a major challenge will be to understand if and how the various genetic and epigenetic alterations increase fitness of the nascent allopolyploid so that, equipped with the advantages of allopolyploidy (e.g. permanent heterosis between homoeoalleles), it can out-compete its two parents (Levy and Feldman, 2004).

## 5.2 New polyploids (neopolyploids)

The immediate consequences of polyploidization can best be studied in polyploids of clear and recent ancestry (Soltis et al, 2004a). The discovery of new polyploid species in the wild soon after its origin presents the opportunity to examine numerous phenomena concerning polyploid speciation and evolution (Abbott and Lowe, 2004). *Tragopogon mirus, T. miscellus, Spartina anglica* and *Senecio cambrensis* are allopolyploids that formed within the last 150 years. Furthermore, their diploid parents have been well-documented. As such, these species represent model systems for the study of recent allopolyploidy (Soltis et al, 2004b). In *Tragopogon*, new polyploid populations are still being formed, offering a replicated natural experiment of genome merger (Adams and Wendel, 2005).

*Spartina anglica* arose during the end of the 19$^{th}$ century in England by hybridization between the indigenous *Spartina maritima* and the introduced East American *Spartina alterniflora* and following genome duplication of the hybrid (*S.townsendii*). This system allows investigations of the early evolutionary changes that accompany stabilization of a new allopolyploid species in natural populations. Various molecular data indicate that *S. anglica* has resulted from a unique parental genotype. This young species contains two

distinctly divergent homoeologous genomes that have not undergone extensive change since their reunion. No burst of retroelements has been encountered in the $F_1$ hybrid or in the allopolyploid, suggesting a 'structural genomic stasis' rather than 'rapid genomic changes'. However, modifications of the methylation patterns in the genomes of *S. townsendii* and *S. anglica* indicate that in this system, epigenetic changes have followed both hybridization and polyploidization (Ainouche et al, 2004).

Two new polyploid species of *Senecio* have originated in the British Isles in recent times following hybridization between native *S. vulgaris* (2n = 40) and introduced *S. squalidus* (2n = 20). One of these is the allohexaploid *S. cambrensis* (2n = 60), the other is the recombinant tetraploid *S. eboracensis* (2n = 40) (Abbott and Lowe, 2004). Abbott and Lowe (2004) reviewed evidence that suggests *S. cambrensis* may have undergone rapid genome evolution since its origin, and comment on the risks of extinction to each species due to chance factors operating during the early establishment phase. The discovery of both species soon after their origin provides an unparalleled opportunity to examine two different but related forms of speciation following hybridization between the same parent species. Further detailed study of the ecology and genomics of *S. cambrensis* and *S. eboracensis* will help improve understanding of the process of polyploid speciation in plants (Abbott and Lowe, 2004).

*Tragopogon mirus* Ownbey and *T. miscellus* Ownbey are allopolyploids that formed repeatedly during the past 80 years following the introduction of three diploids (*T. dubius* Scop., *T. pratensis* L. and *T. porrifolius L.*) from Europe to western North America. These polyploid species of known parentage are useful for studying the consequences of recent and recurrent polyploidization. Analyses of rDNA sequence data indicate that the parental diploids are phylogenetically well separated within *Tragopogon* (a genus of perhaps 150 species), in agreement with isozymic and cpDNA data. They have demonstrated concerted evolution of rDNA. Concerted evolution is ongoing, but has not proceeded to completion in any polyploid population examined; rDNA repeats of the diploid *T. dubius* are typically lost or converted in both allopolyploids, including populations of independent origin. Chromosome numbers of the diploid progenitors is 2n = 12. The diploid chromosome complements are additive in both allopolyploids (2n = 24); there is no evidence of major chromosomal rearrangements in populations of either *T. mirus* or *T. miscellus*.

Approximately 5% of the genes examined in the allopolyploid populations have been silenced, and an additional 4% exhibit novel gene expression relative to their diploid parents (Soltis et al, 2004a).

Molecular investigations of these new polyploids are providing evidence of gene silencing, alterations in cytosine methylation, and other manifestations of parental 'non-additivity' (Adams and Wendel, 2005).

To study the consequences of hybridization and genome duplication on polyploid genome evolution and adaptation, Salmon et al (2005) used independently formed, recent (150 years old) hybrid–allopolyploid system (*Spartina x townsendii* and *Spartina x neyrautii*) that originated from natural crosses between *Spartina alterniflora*, an American introduced species, and the European native *Spartina maritima*. They studied methylation changes using MSAP. 30% of the parental methylation patterns are altered in the hybrids and the allopolyploid. This high level of epigenetic regulation might explain the morphological plasticity of *Spartina anglica* and its larger ecological amplitude. Hybridization rather than genome doubling seems to have triggered most of the methylation changes observed in *Spartina anglica*. Sequence elimination was also reported. Genetic and the epigenetic changes following hybridization appear to have generated major phenotypic effects. (Salmon et al, 2005).

## 6. Summary

Polyploidy is a major force in evolution, particularly in plants. There are many factors that promote polyploidy. Polyploidy formation is possible due to the formation of unreduced gametes and triploid bridges. Since polyploids are hybrids between well differentiated races or species, they exhibit superior vigor than diploids. Fixed heterozygosity is another advantage of diploids over polyploids. If present, apomixis can overcome sterility. This in turn leads to complete fertility like diploids. Polyploids get established either in habitats of diploid parents or in new habitats. Pollinators in some cases tend to visit polyploids more than diploids thus favoring the reproduction of polyploids.

Polyploids are genetically dynamic. Both autopolyploids and allopolyploids exhibit a high frequency of recurrent formation. The allopolyploid genome undergoes many changes following hybridization and polyploidization. There are several models of genome

evolution in which an allopolyploid genome gradually will undergo gene silencing and return to diploid condition. Suppressed homoeologous chromosome pairing also leads to diploidization. With the use of molecular techniques, many plants which were earlier thought to be diploids like maize have been found to be ancient polyploids whose genomes have diploidizsed. The genomes of newly formed allopolyploids react in a burst of irreversible genomic reorganizations and modifications.

Many plant species have been used as model for the study of plant genome evolution. Ancient polyploids include maize (segmental allopolyploid) and rice (ancient aneuploid). *Gossypium* provides one of the textbook examples of speciation via allopolyploidy. The small genome size of Arabidopsis and the availability of its sequence make it convenient for genome study. Newly formed polyploids have been used for the study of numerous phenomena concerning polyploid speciation and evolution and artificially synthesized polyploids have been used to study genome changes immediately after polyploid formation.

Molecular techniques have revealed many aspects not known by the studies done traditionally. But still, much work needs to be done to understand the whole picture of polyploidy and genomic changes associated with it. Most of the work has been done in crop plants. There is still much work to be done in case of wild plants. More plants should be studied as the genomic consequences of polyploidy are not the same in all plants. For example, genomic nonadditivity recently has been reported for both natural and artificial allopolyploid plants (eg Wheat and *Brassica*). But analyses of newly synthesized *Gossypium* allopolyploids showed epigenetic stasis in contrast to other model plants and in another study they showed silencing in different organs. These studies indicate that more data is needed to completely understand the polyploidy paradigm.

Future work needs to address not only the immediate changes following polyploid formation but also the intermediate and longer-term changes occurring over many generations and how these changes help in the establishment of polyploids.

## Acknowledgements

I would like to thank my advisor, Dr. Kifle Dagne, Addis Ababa University, Ethiopia for his corrections and for providing me guidance how to write review papers. I am grateful to Dr. Sileshi Nemomessa, Addis Ababa University, Ethiopia for providing me with guidance how to write a review paper.

## References

Adams, K.L., Cronn, R., Percifield, R., and Wendel J.F. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci.* **100:** 84649–4654

Abbott, R.J. and. Lowe, A.J. (2004). Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biol. J. Linn. Soc.* **82:** 467

Adams, K.L., and. Wendel, J.F. (2004). Exploring the genomic mysteries of polyploidy in cotton. *Biol. J. Linn. Soc.* **82:** 573–581.

Adams, K.L and Wendel, J.F (2005). Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* **8:**135–141

Ainouche, M.L., Baumel, A., and Salmon, A. (2004). *Spartina anglica* C. E. Hubbard: a natural model system for analysing early evolutionary changes that affect allopolyploid genomes. *Biol. J. Linn. Soc.* **82:** (4).

Blanc, G. and Wolfe, K.H. (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**:1679-1691.

Blanc, G., Hokamp, K., and. Wolfe, K.H. (2003). A Recent Polyploidy Superimposed on Older Large-Scale Duplications in the *Arabidopsis* Genome. *Genome* **13**(2)**:** 137-144

Brochmann , C. , Brysting, A. K., Alsos , I. G., Borgen, L., Grundt, H. H,. Scheen, A.C., and Elven, R. (2004). Polyploidy in arctic plants. *Biol. J. Linn. Soc.* **82**(4).

**Comai, L., Tyagi, A.P., Winter, K., Davis, R.H., Reynolds, S.H. , Stevens, Y., and Byers, B.** (2000). Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* Allotetraploids. *Plant Cell* **12:** 1551-1568

Comai, L. (2000). Genetic and epigenetic interactions in allopolyploid plants. *Plant Mol. Biol.* **43**:387–399.

Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A.A., and Vega, J.M. (1997). Rapid elimination of low-copy DNA sequences in polyploid wheat: A possible mechanism for differentiation of homeologous chromosomes. *Genetics* **147:** 1381–1387

Gaut, B.S., and Doebley, J.F. (1997). DNA sequence evidence for the segmental allotetraploid origin of maize (duplicated loci / chromosomal evolution). *Proc. Natl. Acad. Sci.* **94**: 6809-6814

Gaut, B. S., d'Ennequin, M. T, Peek, A S. and Sawkins M.C. (2000). Maize as a model for the evolution of plant nuclear genomes. *Proc. Natl. Acad. Sci.* **97:** 7008–7015.

Lagercrantz, U., and Lydiate, D. J. (1996). Comparative Genome Mapping in *Brassica. Genetics* **144:** 1903-1910

Lagercrantz, U. (1998). Comparative Mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* **150**: 1217-1228

Leitch, I. J and Bennett, M. D. (2004). Genome downsizing in polyploid plants. *Biol. J. Linn. Soc.* **82:** 4 P. 651

Levy, A., and Feldman, M. (2002). The impact of polyploidy on grass genome evolution. *Plant Physiol.* **130:** 1587–1593.

Levy, A. and Feldman, M. (2004). Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. *Biol. J. Linn. Soc.* **82**(4).

Liu, B., and Wendel, J.F. (2002). Non-Mendelian phenomena in allopolyploid genome evolution. *Current Genomics*. **3(**6**)**.

Liu, B., Brubaker, C.L., Mergeai, G., Cronn, R.C., and Wendel, J.F. (2001). Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* **44:** 321–330.

Lukens, L. N., Pablo, A. Q, Udall, J., Piresm, J. C., Schranz, E and Osborn, T. C. (2004). Genome redundancy and plasticity within ancient and recent *Brassica* crop species. *Biol. J. Linn. Soc.* **82**: 665–674.

Madlung, A., Tyagi, A.P., Watson, B., Jiang, H., Kagochi, T., Doerge, R.W., Martienssen, R. and Comai, L. (2005). Genomic changes in synthetic *Arabidopsis* polyploids. *Plant Journal* **41**: 221–230.

Ozkan, H., Levy, A., and Feldman, M. (2001). Allopolyploidy-Induced Rapid Genome Evolution in the Wheat (*Aegilops–Triticum*) Group. *Plant Cell* **13:** 1735-1747.

Paterson A.H., Bowers J.E., Peterson D.G., Estill J.C., Chapman B.A. (2003). Structure and evolution of cereal genomes. *Curr. Opin. Genet. Dev.* **13**: 644–650.

Pikaard, C.S. (2001). Genomic change and gene silencing in polyploids. *Trends Genet*. **17:** 675–677.

Rauh, A.L (2003). Evolutionary dynamics of duplicated genes in plants. *Mol. Phylogenet. Evol.* **29**: 396–409.

Salmon, A., Ainouche, M.L. and Wendel, J.F. (2005). Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (*Poaceae*). *Mol. Ecol.* **14**(4).

Santos, J. L., Alfaro, D., Sanchez, M.E., Armstrong, S. J., Franklin, F. C. and Jones, G.H. (2003). Partial Diploidization of Meiosis in Autotetraploid *Arabidopsis thaliana*. *Genetics* **165:** 1533-1540.

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

Singh, B.D. (1993). *Plant Breeding: Principles and Methods*. Kalyani publishers, New Delhi.

Soltis, D.E., Soltis, P.S., Pires, J. C., Kovarik, A., Tate, J. A., and Mavrodiev, E. (2004a). Recent and recurrent polyploidy in *Tragopogon* (*Asteraceae*): cytogenetic, genomic and genetic comparisons. *Biol. J. Linn. Soc.* **82** (4).

Soltis, D.E., Soltis, P.S., and Tate, J.A. (2004b). Advances in the study of polyploidy since Plant speciation. *New Phytol.* **161:** 173.

Soltis, D. E., Soltis, P.S. and Tate, J.A. (2003). Advances in the study of polyploidy since plant speciation. *New Phytol.* **161**: 173-191.

Soltis, D. E. & Soltis, P. S. (1995). The dynamic nature of polyploid genomes. *Proc. Natl. Acad. Sci.* **92**: 8089-8091.

Soltis, P.S., and Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proc. Natl. Acad. Sci.* **97:** 7051–7057.

Soltis, D. E. & Soltis, P. S. (1999). Polyploidy: recurrent formation and genome evolution. *Trends Ecol. Evol*. **14**: 348–352.

Stebbins, G.L. (1971). *Chromosomal Evolution in Higher Plants*. Addison-Wesley, New York.

Song, K., Lu, P., Tang, K., and Osborn, T.C. (1995). Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. *Proc. Natl. Acad. Sci.* **92:** 7719–7723.

Thompson, J.N., Nuismer, S.L., and Merg, K. (2004). Plant polyploidy and the evolutionary ecology of plant/animal interactions. *Biol. J. Linn. Soc.* **82**: 511–519.

Vandepoele, K., Simillion, C., and Van de Peer, Y. (2003). Evidence That Rice and Other Cereals Are Ancient Aneuploids. *Plant Cell* **15:** 2192-2202.

Wang, X., Shi, X., Hao, B.L., Ge, S., and Luo, J. (2005). Duplication and DNA segmental loss in rice genome and their implications for diploidization. *New Phytol.* **165:** 937–946.

Wendel, J.F. and Cronn, R.C. (2003). Polyploidy and the evolutionary history of cotton. *Advances in Agronomy* **78**

Yang , Z., Guo-hua X., Xing-yi , G., and Long-jiang , F. (2005). Two ancient rounds of polyploidy in rice genome. *J Zhejiang Univ. SCI.* **6B:** 87-90.