

Biosystematic relationships and the formation of polyploids

Richard J. A. Buggs^{1,2}, Pamela S. Soltis^{2,3}, and Douglas E. Soltis^{1,3}

¹Department of Biology, University of Florida, Gainesville, FL

²Florida Museum of Natural History, University of Florida, Gainesville, FL

³The Genetics Institute, University of Florida, Gainesville, FL

Running title: Biosystematic relationships and polyploidy

Author email addresses: buggs@ufl.edu, psoltis@flmnh.ufl.edu, dsoltis@botany.ufl.edu

Abstract

Since the discovery of polyploidy (the presence of more than two chromosome sets in a nucleus) a century ago, scientists have investigated and speculated about the factors which could promote polyploid formation. One of the oldest and most enduring ideas is that hybridization promotes whole genome doubling. First suggested by Ø. Winge, this concept was developed and refined by the major plant evolutionary biologists of the last century, and especially in the biosystematic thinking of J. Clausen, D. D. Keck and W. M. Hiesey. In the past few years, this issue has been revisited by various authors, using molecular systematic methods to study patterns of parental divergence in relation to the formation of polyploids. Progress in molecular genetics and genome evolution also allows re-appraisal of the mechanistic arguments put forward by earlier researchers. In surveying and critically appraising this field over the past century, in the light of recent progress, we conclude that there is not currently persuasive evidence that hybridization between divergent parents serves as a driver for polyploidization.

Biosystematics *sensu stricto* is an approach to systematics that arose in the mid-twentieth century, seeking to take into account biological information about taxa. According to J. Clausen, D. D. Keck and W. M. Hiesey (1945, p. iii): “Genetics, cytology, comparative morphology, and ecology (including distributional and physiological ecology) are among the fields that furnish the critical data. Together, when applied to the study of organic evolution they comprise *biosystematics*.” Thus, biosystematics sought a global biological approach to systematics. It gave a more dynamic view of taxa than had been common before, through the incorporation of morphological data at the population level using numerical taxonomy (phenetics). This approach has now been largely eclipsed by the cladistic approaches of molecular systematics. The exclusive use of molecular phylogenetics is not truly biosystematic, as it concentrates only on genetic data and phylogenetic history. In this contribution we describe 90 years of discussion on an important issue in plant evolution that has taken place in the context of changing approaches to systematics. We discuss how these changing approaches, and especially the move from biosystematics to molecular systematics, have influenced and developed the ideas of biologists.

Polyploidy, the presence of more than two chromosome sets in a nucleus, was discovered in the early years of the twentieth century (Gates 1909; Lutz 1907; Strasburger 1910). How was this novel phenomenon to be explained? One explanation, first put forward by Winge (1917), was that polyploidization is initiated by hybridization. Since that time, the issue of how the systematic relationship between the parents of polyploids can affect the polyploidization process has been actively investigated and discussed. Today there is an ongoing debate about whether or not the extent of parental

divergence provides a drive for polyploidy (Buggs *et al.* 2008; Buggs *et al.* 2009; Chapman & Burke 2007; Paun *et al.* 2009).

Understanding the extent to which parental divergence affects polyploid formation is important because polyploidization is a very common mechanism of speciation in plants, the vast majority of extant plant species showing evidence for recent or ancient polyploidy (Doyle *et al.* 2008; Fawcett *et al.* 2009; Soltis *et al.* 2009). There has recently been a renewed interest in the contribution which hybridization makes to evolution both via homoploid hybrid speciation and allopolyploidy (Mallet 2007; Rieseberg *et al.* 2003; Rieseberg & Willis 2007; Soltis & Soltis 2009). In addition, the interaction of systematic relationships with polyploidization is important from a theoretical point of view, because any causal connection could be seen as making the evolutionary process more predictable.

There is a long history of critical thought on the relationship of parental divergence and the opportunity for hybridization and polyploidization, particularly from the biosystematic era of the mid-twentieth century. This literature and the important insights therein are easily overlooked in this molecular era. Below, we provide an historical context reviewing the thinking of workers in this field beginning nearly a century ago and tracing the evolution of the issue.

Ø. Winge

Winge (1917) was the first to develop a theory linking the formation of polyploids and hybridization. He suggested that “occasional hybridization might be the cause” (p.

13) of polyploidy and directly linked the probability of chromosome doubling to the relatedness of the hybridizing parents. He drew up “a scheme showing the different degrees of physiogenetic likeness between gametes – and thus also between their chromosomes – endeavouring at the same time to ascertain what results we can expect in each case from the fusion of gametes.” In this scheme, chromosomes of gametes from within the same species will pair readily with each other, producing viable progeny. He termed this “philozygoty” (literally: “loving to be yoked [i.e. paired]”). Chromosomes of gametes derived from different species or races exhibit “pathozygoty” (literally: “submitting to be yoked”), and do not readily pair. Mild cases of pathozygoty would cause reduced fertility without causing sterility, but in more extreme cases, “indirect chromosome union” (Figure 1) would be required to restore fertility:

“If the chromosomes are to find a partner, then each of the chromosomes in the zygote must divide, for thus indirectly to produce a union of chromosomes, and we must assume that this is realized in the hybrid zygotes which have any possibility at all of propagating – in accordance with what we know from experience as to the behaviour of pairs of chromosomes. The hybrid sporophyte thus produced will then have 4x chromosomes, taking the number for each of the parent gametes as x... We may doubtless assume that most of the species exhibiting ‘double’ chromosome numbers are hybrids formed in this manner.” (p. 199).

A third condition, “misozygoty” (literally: “hating to be yoked”) occurs between chromosomes of “systematically widely differing organisms” (p. 201) and results in a failure of fertilization.

Winge (1917) thought that polyploidization occurred in the sporophyte after the union of gametes rather than through unreduced gamete formation. He did observe that “imperfect reduction” of pollen “is associated with the hybrid nature” but he did not link this to polyploidization. He was vague as to exactly what caused the division of the chromosomes (which he also called “cleavage”) in the sporophyte, but assumed that it simply must happen. Later, Darlington (1937) would claim that Winge had the teleological idea that the need for a partner would stimulate chromosome doubling, but Winge does not say this directly in his 1917 work, in which he also rejects teleological ideas (p. 195).

Winge was also interested in finding law-like processes or underlying principles in evolution. In a discussion of hybrid organisms, he reports that he had once thought that “the sterility of the offspring is in inverse proportion to the power of conjugation in the chromosomes” (p. 250). However, on further investigation he realized there were many exceptions to this, and he concluded that in genetics it is not possible to apply “laws in the mathematical sense” (p. 250).

In 1927, Kihara and Ono proposed that there were two types of polyploids: autopolyploids whose chromosomes sets come from the same species, and allopolyploids which are derived from interspecific hybrids (Kihara & Ono 1927). They briefly referred to Winge’s (1917) hypothesis that hybridization causes polyploidy, but the concept of autopolyploidy would seem to contradict the idea that this hypothesis is universal.

C. D. Darlington

Darlington (1937) objected to Winge's idea of "indirect chromosome union" as being teleological and suggested that polyploidization is due to gametic doubling that occurs when meiosis has proved "unworkable". Darlington (1937) also developed Winge's earlier idea of a law by proposing an inverse relationship between the fertility of a diploid hybrid and that of a tetraploid to which it gives rise (Figure 2). He reasoned that (1) at low parental divergences, homoploid hybrids will be fertile because chromosomes will be able to pair at meiosis, but allopolyploids will be of low fertility because pairing will occur between both duplicated chromosomes and homeologous chromosomes from each parent, causing multivalent formation and uneven segregation. In contrast, (2) at high parental divergences, homoploid hybrids will be sterile due to failure of chromosome pairing, but allopolyploids will be fertile due to consistent bivalent formation at meiosis. This has sometimes been called "Darlington's rule".

In formulating this rule, Darlington assumed that "there is a correlation between the genetic differentiation of the chromosomes of species and their structural differentiation...all interspecific hybrids are also structural hybrids." (Darlington 1937, p. 197). However, earlier in the same book Darlington suggests that this correlation will not always hold and cautions against simplistic application of chromosomal and genetic differences in systematics:

Structural changes themselves have no necessary genetic effect, but they may condition genetic isolation of stocks, and hence a later genetic differentiation between species. But this differentiation need not follow. We therefore find every relationship between structural and numerical and genotypic differences in the chromosomes on the one hand and systematic differences on the other...A comparison of the mitotic chromosomes of different forms is

therefore of little value in placing them systematically unless we know the type of variation that prevails in the groups in question. If the forms have different chromosome numbers they are probably intersterile or yield sterile hybrids, a fact which the systematist may or may not take into consideration. (Darlington 1937, p. 79-80)

The variable relationship between chromosomal differences and systematic differentiation underlines the difficulties involved in the classification of polyploids, problems which have persisted to the present day. The fact that differentiation in chromosome structure between species may not correlate with genetic differences is a problem for approaches that rely solely on molecular/genetic data to address the association between systematic relationships and polyploidy, an issue to which we shall return below.

J. Clausen, D. D. Keck and W. M. Hiesey

The biosystematists J. Clausen, D. D. Keck and W. M. Hiesey (1945) devoted a chapter of their book *Experimental Studies on the Nature of Species, Volume II: Plant Evolution Through Amphiploidy and Autoploidy with Examples from the Madiinae* to “Biosystematic relationships and amphiploidy”. Here, they reviewed current thought in biosystematics, culminating in a section on the requirements for success in amphiploids (polyploids with an interspecific hybrid history, also known as allopolyploids). They concluded that the parents of amphiploids “should be closely enough related to produce a vigorous F₁ hybrid, but remotely enough so that the balance between their combined genomes can be perpetuated” (Clausen *et al.* 1945, p. 69). Clausen *et al.* based their

argument almost entirely on their biosystematic thinking, which in many ways contrasts with the systematic perspectives of today. Below we briefly review their ideas using their original subheadings.

The principles governing biosystematic investigations.— The idea of “balance” is central to the biosystematic approach of Clausen *et al.* (1945). Clausen *et al.* held that natural species exist in a delicate balance both internally between their genes, and externally with their environment. Hybridization between species disrupts this harmonious balance, resulting in reduced fitness. Under this view, they note that slow evolutionary change and paleontological stasis were to be expected, whereas the free interchange of genes between ecotypes and subspecies was surprising. Clausen *et al.* (1945) emphasized that the morphology of living organisms is controlled by both genetic and ecologic factors, and that morphology was therefore a good indicator of the “more basic genetic-physiologic” character of species.

The biosystematic units.—Clausen *et al.* (1945) argued that “an abundance of experimental data” shows that “natural units of various ranks exist” (p. 63). They demarcated these “natural units” on the basis of hybridization experiments, recognizing four biosystematic ranks. (1) *Ecotypes* are separated by differences in their external balance with the environment, but have the same internal balance so are not separated by endogenous factors. They can freely hybridize. (2) *Ecospecies* are separated by differences in both their internal and external balance. Ecospecies possess endogenous barriers to hybridization, but limited interchange of genes can occur. (3) *Cenospecies* are

also separated by differences in both their internal and external balance, but endogenous barriers to hybridization can be overcome by polyploidization. (4) *Comparia* are species that are separated by such large differences in internal balance that they cannot produce even a sterile hybrid. Clausen *et al.* acknowledged that these ranks were points on a continuum of evolutionary change, but considered them “important evolutionary nodes”. For example, reticulate evolution is only possible below the level of the *comparium*.

The evolutionary process.—Clausen *et al.* (1945) argued that through mutation, recombination and selection, an ecotype can ultimately evolve into an ecospecies, then a cenospecies, then a *comparium*. Hybridization also played an important role in Clausen *et al.*'s view of the evolutionary process. They speculated that a *comparium* with many hybridizing units of low rank was “in its most active and expansionist stage of its development” (p. 66), but that over time a *comparium* could become depleted and eventually consist of a single ecotype. The balance between evolutionary forces would therefore determine the success or failure of a *comparium*.

Applications of biosystematic principles.—Hybridization is clearly fundamental to the biosystematic approach of Clausen *et al.* (1945) who viewed their system (reviewed above) as a biological classification. In many cases they found their classification to correspond well with purely morphological classification, but there were significant exceptions, such as the grasses *Lolium perenne* and *Festuca pratensis* which were placed in different tribes but can hybridize and therefore belong to the same cenospecies. (significantly, *Lolium* is now known to be nested within *Festuca* so the ability to

hybridize is not surprising; see www.mobot.org/MOBOT/research/APWeb/.) Clausen *et al.* concluded that the tribes “may be entirely artificial” and the classification of such groups needed revision.

The requirements for success in amphiploids.—Emphasis on “balance” and hybridization in the biosystematic thinking of Clausen *et al.* logically led them to formulate principles for the success of allopolyploids. They argued that the parents of an allopolyploid should be closely enough related to produce an F₁ hybrid which has “harmonious and vigorous development”. For the allopolyploid to succeed through subsequent generations, however, the “original balance” within the two genomes “must remain unchanged”. Recombination could not be permitted between the two genomes: “since the balances that determine success or failure are intricate and delicate, they may be upset by slight genetic interchanges.” To prevent recombination, the chromosomes had to be different enough not to pair at meiosis. This generally meant that the two parental species had to be distantly related congeners. The requirement for success in allopolyploids was therefore itself a delicate balance between being too closely related and too distantly related.

In the Introduction to their book, Clausen *et al.* (1945) also give empirical survey evidence for a link between relatedness of parents and success in allopolyploids. They report

...a critical survey of the relationships between the parents, and of the circumstances surrounding the origin of some of the better-studied amphiploids [allopolyploids] reported in the literature, in the search for fundamental principles

that govern the successful production of amphiploids. From this survey it appeared that the success and constancy of amphiploids is linked with the degree of relationship found between their parents. (Clausen *et al.* 1945 p. 2)

They give no more details of this survey, however, so it is hard to assess directly its results. Their ideas about biosystematic relationships and polyploidy therefore stand or fall on whether or not their biosystematic ideas – and particularly their emphasis on “balance” – were correct.

G. L. Stebbins

Stebbins appears to have agreed with the arguments of Darlington (1937) and Clausen *et al.* (1945), noting that the parental species of polyploids should be “sufficiently closely related to each other so that they can produce vigorous F₁ hybrids, but strongly enough differentiated so that their chromosomes are nearly or entirely incapable of pairing with each other” (Stebbins 1950 p. 357). Stebbins initially viewed polyploidization as a rare event, resulting in polyploid species that were genetically uniform. Because of this, and buffering of the evolutionary process due to multiple genomes, Stebbins’ early view saw polyploids as evolutionary dead ends (Stebbins 1950; Stebbins 1971) and considered autopolyploidy to be maladaptive and very rare in natural populations (Stebbins 1950; Stebbins 1971). In later years he modified this view somewhat as he expanded his view of autopolyploidy to include crosses between ecotypes and different diploid genotypes within a species; such crosses could yield autopolyploids that were superior to their parental diploids (Stebbins 1985). Stebbin’s

secondary contact hypothesis for the origin of polyploidy relies upon the idea that frequent opportunities for hybridization lead to high incidence of successful polyploidy (Stebbins 1984, 1985). He argued that “successful polyploids among natural populations are usually or almost always the result of increased heterozygosity accompanying either interracial or interspecific hybridization” (Stebbins 1985, p. 824).

V. Grant

When Grant surveyed the literature to identify factors promoting allopolyploidy, he did not mention a putative role of relationship between parental species, but he referred to more general factors that could differentiate parental genomes (Grant 1981). He considered it crucial for successful allopolyploids that the parental genomes differ due to chromosomal re-patterning as this causes reduced chromosome pairing in hybrids, which “sets the stage” for non-reduced gamete formation and hence allopolyploidy. He also outlined evidence that certain species or genera had a proclivity for allopolyploidization due to variation in certain genes controlling meiosis. Overall, his emphasis seems to have been not on a simple correlation between parental divergence and allopolyploid formation, but on mechanisms promoting polyploidy which sometimes will, but sometimes will not, be affected by parental divergence.

Grant (1981) considered that in hybridizing plant groups there is an inclusive unit—the syngameon—composed of interbreeding individuals above the species level (Lotsy 1925), similar to the *cenospecies* of Clausen *et al.* (1945). Grant defined the

syngameon as “the sum total of species or semispecies linked by frequent or occasional hybridization in nature” (Grant 1981; p. 234). His view of the evolutionary potential and fate of syngameons stressed the importance of maintaining sexually reproducing diploid species: without these, a syngameon would be an evolutionary dead-end (Grant, 1981). In other words, Grant essentially agreed with Stebbins in thinking that polyploids by themselves had little evolutionary potential.

M. A. Chapman and J. M. Burke

Huge progress was made in the study of polyploidy between 1981 and 2007, but Chapman and Burke (2007) were the first authors to explicitly bring the discussion of the effect of parental divergence on polyploidization into the molecular era of systematics. They provided the first study that uses DNA sequences to assess the genetic distance (as a proxy for evolutionary divergence) between the parental species of homoploid and polyploid hybrid species. They calculated Kimura’s two-parameter (K2P) genetic distance between DNA sequences from the internal transcribed spacer (ITS) region of nuclear ribosomal RNA genes of 12 species pairs that have given rise to homoploid hybrid species and 26 species pairs that have given rise to allopolyploid species. They compared all hybrid versus allopolyploid parental pairs and found a significantly larger genetic sequence divergence between the parents of allopolyploids. They concluded that “the extent of evolutionary divergence between hybridizing taxa plays an important role in determining the outcome of hybrid speciation” (p. 1778). In their discussion, Chapman and Burke (2007) suggested that this relationship might be due to a greater probability of

unreduced gamete formation in a hybrid between more distantly related parents (cf. Grant 1981), or higher fitness of allopolyploids from divergent crosses, due to preferential pairing of homologous chromosomes and fewer meiotic abnormalities (cf. Darlington 1937).

R. J. A. Buggs, P. S. Soltis, E. V. Mavrodiev, V. V. Symonds, and D. E. Soltis

Buggs *et al.* (2008) used molecular phylogenies of eight genera that contain polyploids to test the hypothesis that closely related parents are less likely to form a successful polyploid than more divergent parents. This hypothesis is a major corollary of the idea that phylogenetic divergence drives polyploidization. They used node-based and clade-based methods of calculating the phylogenetic distance between parental pairs to compare these with expected divergences based on the null hypothesis that hybridization would occur successfully at random between all species of a genus. They found that the phylogenetic divergence between parents of polyploids was not significantly different from the divergence expected under this null hypothesis. The same analysis on homoploid hybrids (including unstable hybrids) in the same genera found a lower divergence between the parents of homoploid hybrids than the null expectation. They concluded that “contrasting patterns of divergence between the parents of polyploids and homoploid hybrids are...determined by the restriction of homoploid hybrid formation to low parental divergence, rather than the restriction of polyploid formation to high parental divergence” (p. 87).

O. Paun, F. Forest, M. F. Fay and M. W. Chase

Paun *et al.* (2009) conducted an additional analysis that combined and improved some of the approaches of Chapman and Burke (2007) and Buggs *et al.* (2008). For 16 homoploid hybrids and 32 allopolyploids, they calculated uncorrected p-distances and K2P distances between parental pairs using ITS and/or low-copy nuclear gene sequences. They converted each of these distances to a genetic divergence index (GDI) by dividing parental divergence by the average genetic distance between all pairs in each genus based on the same molecular markers. The GDI gave very similar results for both distance measures, and parents of polyploids were found to be significantly more divergent than parents of hybrids. Fitting a heuristic model to their data, Paun *et al.* (2009) suggested that at a GDI of around 0.75, there is an equal probability of a hybrid being homoploid or allopolyploid, but above this point, allopolyploidy is more likely, and below this, homoploidy is more likely. They concluded that “parental divergence drives ploidy”, discussing this in terms of some of the mechanisms suggested by earlier researchers (reviewed above).

Although Paun *et al.* (2009) calculated the average divergence between all species pairs in each genus, they did not use this as a null hypothesis for the expected divergence between parents of allopolyploids as in Buggs *et al.* (2008). Therefore, Buggs *et al.* (2009) carried out a two-tailed paired *t*-test on the genetic distances between parental pairs and the average genetic distance between all species pairs in their respective genera, using the data from Table S1 of Paun *et al.* (2009). They found a significant difference between the expected and observed values for homoploid hybrids ($t = 3.427$, d.f. = 15, P

< 0.01), but no significant difference for allopolyploids ($t = 1.533$, d.f. = 31, $P > 0.1$). Significantly, this result agreed with that of Buggs *et al.* (2008): homoploid hybrid formation occurs at low parental divergence, but polyploid formation fits a model of random hybridization.

Discussion

Systematics has developed considerably since a link was first suggested between diploid parental divergence and polyploid formation. In particular, our knowledge of molecular genetics has exploded in the period between Grant (1981) and Chapman and Burke (2007), bringing new approaches to systematics and a new knowledge of genome evolution. For example, today biologists agree with Clausen *et al.* (1945) that an organism's phenotype is produced by the interaction between genotype and environment, but unlike Clausen *et al.* we now base most systematic investigations of relationship directly on gene sequence data, under the assumption that this approach provides a more direct measure of the phylogenetic relatedness of organisms (Cantino *et al.* 1999; de Queiroz & Gauthier 1990; de Queiroz & Gauthier 1992; Forey 2002; Keller *et al.* 2003). Modern systematics relies upon monophyletic groups inferred in molecular phylogenies (e.g. Judd *et al.* 2007), whereas the biosystematics of Clausen *et al.* demarcated "natural units" on the basis of hybridization experiments in conjunction with other data. The current view that taxonomic ranks are arbitrary (e.g. de Queiroz & Gauthier 1992; Wiley 1981) contrasts with the view of Clausen *et al.* that "an abundance of experimental data" shows that "natural units of various ranks exist" (p. 63), though it agrees with their view

that higher taxonomic ranks such as tribes are arbitrary. The connection between systematic relationship of progenitor diploids and polyploidy requires thorough reexamination in light of the genetic/phylogenetic revolution of the past 20 years.

Two types of argument have been made in support of the idea that divergence between parental diploids is responsible for driving polyploidy. The first type is somewhat *a priori*, based on theories of cytogenetic interactions as proposed by early cytologists. These cytogenetic interaction arguments are essentially about the processes involved in polyploid evolution. The second type of argument is more *a posteriori*, based on surveys of existing polyploids, as carried out in recent studies to determine which diploid combinations have produced existing polyploids and homoploid hybrids. We will discuss these two approaches in turn.

Process-based arguments

Our knowledge of genome evolution has developed considerably in the past two decades due to the emergence of molecular genetics. Unlike earlier views that the genome is finely balanced and easily perturbed (Clausen *et al.* 1945), today we know from molecular genetic studies that the genome is a dynamic entity that is able to undergo many duplications, deletions and insertions without significant detriment (Leitch & Leitch 2008; Lonngig & Saedler 2002; McClintock 1984). In fact, perturbation of the genome following hybridization is now seen as often beneficial, and the low fitness of hybrids is often considered to be due to specific genes (Arnold 2006; Lexer & Widmer

2008) or chromosomal differences (Rieseberg 2001) rather than a disruption of internal “balance” (Clausen *et al.* 1945).

Molecular studies have also prompted a more dynamic view of polyploidy than that held by earlier researchers (reviewed by Soltis and Soltis 1999). Darlington (1937), Clausen *et al.* (1945) and Stebbins (1950) made arguments based on the premise that polyploids would best succeed if they maintained the integrity of the two diploid genomes within them. However, there is abundant recent evidence that homeologous recombination does occur in allopolyploids. It has been found in mapping populations of allotetraploid *B. napus* (Gaeta *et al.* 2007; Udall *et al.* 2005), and ESTs of *Gossypium* allotetraploids (Salmon *et al.* 2010). Chromosome painting techniques have revealed intergenomic rearrangements in allopolyploids: *Nicotiana tabacum* (Kenton *et al.* 1993; Lim *et al.* 2004), *Tragopogon miscellus* (Lim *et al.* 2008), *Avena maroccana* and *Avena sativa* (Hayasaki *et al.* 2000). Homeologous recombination may be an important component of the dynamic nature of polyploid genomes and contribute to their evolutionary potential (Gaeta and Pires 2010).

Chromosomal mis-pairing in polyploids has costs, as early cytologists realized (Darlington 1937). However, as Ramsey and Schemske (1998) have argued, earlier researchers may have underestimated the rates of polyploid production that occur in nature, which could provide multiple opportunities for rare successful rearrangements to occur. We now know that polyploidization occurs more frequently than was previously appreciated and that independent origins can generate different genotypes (reviewed in Soltis & Soltis 1999; Soltis & Soltis 2000). Thus, in most individuals, homeologous pairing will reduce fertility, but for the population as a whole, the occurrence of

homeologous recombinants may introduce valuable new variation. On the other hand, we also know that prevention of homeologous recombination may be provided for by specific genes, such as *Ph1* in wheat (Cifuentes *et al.* 2010; Mestiri *et al.* 2010), so structural differences between chromosomes are not essential to prevent pairing.

Views of autopolyploidy have changed dramatically over the past century, and consideration of autopolyploids is crucial to the subject of parental divergence and polyploid formation. The relatively recent recognition that autopolyploidy is prevalent in nature tends to contradict Winge's hypothesis that hybridization between species is the cause of polyploidy (Winge 1917). However, as noted, autopolyploidy was considered rare for many decades and of little evolutionary importance. More recently we have come to a new view that autopolyploids are formed very frequently in nature (e.g. Ramsey & Schemske 1998; Soltis & Soltis 1993) even though they are rarely classified as separate species (Soltis *et al.* 2007). It seems that even though autopolyploids may initially have frequent meiotic abnormalities (Darlington 1937), normal meiosis is quickly selected for (Ramsey & Schemske 1998), and may be to some extent under genetic control (Cifuentes *et al.* 2010). Thus autopolyploids are likely to have more evolutionary potential than was once thought (Parisod *et al.* 2010).

Many factors relating to parental divergence may affect polyploid formation, making the polyploidization process a complex one. Under an allopatric model of speciation, more divergent parents are less likely to meet in space and time, reducing rates of hybridization between distantly-related species. Unreduced gamete formation is more likely to occur in hybrids, increasing the likelihood of polyploidization (reviewed in Ramanna & Jacobsen 2003; Ramsey & Schemske 1998). Transgressive segregation may

occur more in hybrids between more distant parents (Stelkens & Seehausen 2009). Differences in small RNA sequences may produce post-fertilization barriers (Martienssen 2010). Parental divergence may also affect resistance to inbreeding depression, though there are currently few published empirical or theoretical studies in this area (Birchler *et al.* 2003; Pannell *et al.* 2004).

One crucial factor is the degree to which chromosome structural differences occur between hybridizing species. As Darlington (1937) pointed out, the extent of these differences does not necessarily correlate with other differences. A few chromosomal rearrangements could, for example, differentiate between two populations without any changes at the base pair level in an ITS sequence. Thus, biologically relevant parental divergence is present in one character but invisible in another which is easier to characterise. Moreover, in many cases the extent of chromosomal differentiation may be more important for successful polyploidization than differentiation at commonly sequenced genetic loci.

In addition, there are many factors that may be involved in polyploid formation that do not depend upon parental divergence and have a confounding effect. For example, certain species or genera have a proclivity for successful allopolyploidization due to variations in genes controlling meiosis (Bretagnolle & Thompson 1995; Cai & Xu 2007; Grant 1981; Ramanna & Jacobsen 2003; Ramsey & Schemske 1998). Chance environmental effects may also influence the frequency of genome doubling (Hagerup 1932; Mable 2004; Ramsey & Schemske 1998). Allotetraploid *Leucaena leucocephala* appears to have formed multiple times due to movement by humans and serendipitous backyard hybridization (Hughes *et al.* 2007), and several classic cases of recent

polyploidization occurred because human introductions brought previously allopatric species together (Abbott & Lowe 2004; Ainouche *et al.* 2004; Harris & Ingram 1992; Ownbey 1950; Soltis *et al.* 2004). These all make it impossible to make simple generalizations about the relationship between parental divergence and polyploidy based on studies of a single step in the complex process of polyploid origin and establishment. They call for a modern biosystematic approach to the issue that seeks to take all of these factors into account.

Pattern-based arguments

The molecular phylogenetic approach to systematics has permitted new tests of the relationship between parental divergence and polyploidy, in the surveys of genetic distance by Chapman and Burke (2007) and Paun *et al.* (2009), and the phylogenetic survey of Buggs *et al.* (2008). In some ways, these methods bring a new objectivity to bear on the questions and have been interpreted as supporting the hypothesis that parental divergence drives polyploidy (Paun *et al.* 2009). Buggs *et al.* (2009) argue that the three surveys all provide approximately the same pattern (see Figure 3A): polyploids tend to be formed from parents of higher divergence than those of homoploid hybrid species, but this pattern reflects random crossing leading to polyploids and the restriction of homoploid hybrids to lower divergences (Buggs *et al.* 2008, 2009).

Whilst arguing that the surveys as they stand do not support the hypothesis that parental divergence drives polyploidy, Buggs *et al.* (2008, 2009) suggested that the

surveys have severe problems in their sampling. In the recent surveys, Chapman and Burke (2007) and Paun *et al.* (2009) excluded autopolyploids from their analysis, and in the eight genera studied by Buggs *et al.* (2008) only one autopolyploid was known. Buggs *et al.* (2008, 2009) argue that autopolyploids should be included, as there is a continuum between allo- and autopolyploidy that can only be artificially demarcated (Wendel & Doyle 2005). We do not know the true frequency in nature of cryptic autopolyploids and cryptic allopolyploids between closely related parents (Soltis *et al.* 2007). There has therefore been some debate over whether or not autopolyploids should be included in surveys, and what effect their inclusion would have on the conclusions regarding which conditions promote polyploidization. Buggs *et al.* (2008, 2009) argue that the inclusion of autopolyploids would give a negative relationship of polyploid frequency versus parental divergence (Figure 3B). In contrast, Paun *et al.* (2009) argue that autopolyploids should not be included in these comparisons, as they do not correspond to any speciation process at the homoploid level (i.e. homoploid hybrids cannot form if parental divergence is zero), and if they were included, a bimodal distribution would result (Figure 3C). This debate is likely to continue until we have better estimates of the natural frequency of autopolyploids and allopolyploids derived from closely related parental species. This has only been possible in a rigorous fashion since the common use of molecular markers and the development of coalescent theory that helps to examine alternative hypotheses regarding multiple origins of polyploidy versus introgression between polyploids and parental taxa. However, independent origins of polyploids and allopolyploids from closely related parental species are still hard to detect (Rieseberg and Willis 2007).

There is a further problem in the sampling strategies of these surveys because they rely upon delimitation of species. As the work of earlier biosystematists shows, species delimitation in hybridizing plant groups is difficult (see above). Due to the common use of different species concepts and types of information, the classification of certain groups as species is not standardized and somewhat arbitrary, even today. Chapman and Burke (2007) and Paun *et al.* (2009) make a direct comparison between homoploid hybrid species and allopolyploid species, as if these “species” were comparable evolutionary units, but as Buggs *et al.* (2009) discuss, this is not the case. Polyploid and hybrid species may not be exactly comparable evolutionary units: while polyploid species typically have post-zygotic reproductive isolation from their parents, homoploid hybrids are likely to be more interfertile with their parents than the parents are with each other. This example illustrates the general problem that recognized species are not equivalent units and may best be treated as arbitrary units, like the rest of the taxonomic hierarchy (Mishler and Theriot, 2000). Plants currently named as species can correspond to ecotypes, ecospecies, cenospecies, or comparia of Clausen *et al.* (1945) and may or may not correspond to biological entities. Whilst systematics has made great strides over the past decades, the species problem remains, and recognizing that “species” are not equivalent is important for teasing apart relevant elements of parental divergence important for polyploid formation.

Many different ecological, chromosomal or genic factors can be involved in differentiation between two populations. Biosystematics sought to take all of these factors into account, though studying all these factors is an immense task. The molecular and phylogenetic approaches of Chapman and Burke (2007), Paun *et al.* (2009) and Buggs *et*

al. (2008) assume that these different forms of differentiation will generally correlate with differentiation in nucleotide sequence divergence in a few genes. This assumption may be a useful approximation and a starting point for investigation, but a more holistic, biosystematic approach to the issue will be needed for final answers to the issue of parental divergence and polyploidy.

Concluding remarks

The connection between systematic relationships and polyploidy has been discussed for 90 years, through a time of much progress in biology. In the modern context, we have to avoid two problems. On the one hand, we cannot ignore the work of previous generations as obsolete: earlier scientists had many correct insights and molecular methods in systematics may benefit from an interaction with previous approaches, such as the biosystematics of Clausen *et al* (1945). Indeed, today we need a new biosystematic approach that takes into account many more components of species divergence than a few DNA sequences if we are to understand the factors that promote or allow polyploidy formation. On the other hand, we cannot cite arguments made by earlier researchers without taking into account the context in which their theories were developed, as some of the less obvious premises of their arguments may now be questioned (such as the view that homeologous recombination is always deleterious).

Two types of argument have been made in support of the idea that divergence between parents is responsible for driving polyploidy. Arguments based on theories of cytogenetic interactions as proposed by early cytologists are now outdated, and they tend

to underestimate the effect of other factors involved in successful establishment of a polyploid, such as the role of selection in weeding out meiotic abnormalities. They generally emphasize one part of the polyploidization process without taking into account the complexity of the whole process. Arguments based on surveys of existing polyploids do not provide evidence for parental divergence driving polyploidy because the polyploids arguably fit a pattern of random crossing within a genus, with homoploid hybrid species restricted in the parental divergence that they can tolerate. In addition, surveys of patterns are subject to problems of sampling and a strong case can be made that they should include autopolyploids and cryptic allopolyploids result from closely related parents.

In our view, there is not currently persuasive evidence that hybridization between divergent parents serves as a driver for polyploidization. There is not a clear connection between either phylogenetic relatedness or genetic divergence and the occurrence of polyploidy. Future work needs to combine the approaches of both the biosystematists of the mid-twentieth century and current molecular systematics to take into account the many different factors that are involved in differentiation and determine whether these are relevant to polyploid formation.

References

- Abbott RJ, Lowe AJ (2004) Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society*, **82**, 467-474.
- Ainouche ML, Salmon A, Baumel A, Yannic G (2004) Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytologist*, **161**, 165-172.
- Arnold ML (2006) *Evolution through genetic exchange* Oxford University Press, Oxford.

- Birchler JA, Auger DL, Riddle NC (2003) In search of the molecular basis of heterosis. *Plant Cell*, **15**, 2236-2239.
- Bretagnolle F, Thompson JD (1995) Tansley Review No. 78. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist*, **129**, 1-22.
- Buggs RJA, Soltis PS, Mavrodiev EV, Symonds VV, Soltis DE (2008) Does phylogenetic distance between parental genomes govern the success of polyploids? *Castanea*, **73**, 74-93.
- Buggs RJA, Soltis PS, Soltis DE (2009) Does hybridization between divergent progenitors drive whole-genome duplication? *Molecular Ecology*, **18**, 3334-3339.
- Cai X, Xu SS (2007) Meiosis-driven genome variation in plants. *Current Genomics*, **8**, 151-161.
- Cantino PD, Bryant HN, Queiroz KD, *et al.* (1999) Species names in phylogenetic nomenclature. *Systematic Biology*, **48**, 790-807.
- Chapman MA, Burke JM (2007) Genetic divergence and hybrid speciation. *Evolution*, **61**, 1773-1780.
- Cifuentes M, Grandont L, Moore G, Chèvre A-M, Jenczewski E (2010) Genetic regulation of meiosis in polyploid species: new insights into an old question. *New Phytologist*, **186**, 29-36)
- Clausen J, Keck DD, Hiesey WM (1945) Experimental studies on the nature of species. II. Plant evolution through amphiploidy and autopolyploidy, with examples from the Madiinae. *Carnegie Inst. Wash.*, **564**.
- Darlington CD (1937) *Recent Advances in Cytology* Blakiston, Philadelphia.
- de Queiroz K, Gauthier J (1990) Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systematic Zoology*, **39**, 307-322.
- de Queiroz K, Gauthier J (1992) Phylogenetic taxonomy. *Annual Review of Ecology and Systematics*, **23**, 449-480.
- Doyle JJ, Flagel LE, Paterson AH, *et al.* (2008) Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics*, **42**, 443-461.
- Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences, USA*, **106**, 5737-5742.
- Forey PL (2002) PhyloCode: pain, no gain. *Taxon*, **51**, 43-54.
- 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**, 3403-3417
- Gaeta RT, Pires JC (2010) Homoeologous recombination in allopolyploids: the polyploid ratchet. *New Phytologist* **186**:18-28.
- Gates RR (1909) The behavior of the chromosomes in *Oenothera lata* x *O. gigas*. *Botanical Gazette*, **48**, 179.
- Grant V (1981) *Plant Speciation*, 2nd edn. Columbia University Press, New York.
- Hagerup O (1932) Über polyploidie in *Beziehung* zu klima, ökologie und phylogenie. *Hereditas*, **16**, 19-40.
- Harris SA, Ingram R (1992) Molecular systematics of the genus *Senecio* L .1. Hybridization in a British polyploid complex. *Heredity*, **69**, 1-10.

- Hayasaki M, Morikawa T, Tatumoto I (2000) Intergenomic translocations of polyploid oats (genus *Avena*) revealed by genomic in situ hybridization. *Genes & Genetic Systems*, **75**, 167-171.
- Hughes CE, Govindarajulu R, Robertson A, *et al.* (2007) Serendipitous backyard hybridization and the origin of crops. *Proceedings of the National Academy of Sciences, USA*, **104**, 14389-14394.
- Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2007) *Plant Systematics: A Phylogenetic Approach*, Third Edition edn. Sinauer, Sunderland, MA.
- Keller R, Boyd R, Wheeler Q (2003) The illogical basis of phylogenetic nomenclature. *The Botanical Review*, **69**, 93-110.
- Kenton A, Parokony AS, Gleba YY, Bennett MD (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Molecular and General Genetics MGG*, **240**, 159-169.
- Kihara H, Ono T (1927) Chromosomenzahlen und systematische Gruppierung der *Rumex*-Arten. *Zeitschrift für Zellforschung und mikroskopische Anatomie*, **4**, 475-481.
- Leitch AR, Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. *Science*, **320**, 481-483.
- Lexer C, Widmer A (2008) The genic view of plant speciation: recent progress and emerging questions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3023-3036.
- Lim KY, Matyasek R, Kovarik A, Leitch AR (2004) Genome evolution in allotetraploid *Nicotiana*. *Biological Journal of the Linnean Society*, **82**, 599-606.
- Lim KY, Soltis DE, Soltis PS, *et al.* (2008) Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). *PLoS ONE*, **3**, e3353.
- Lonnig WE, Saedler H (2002) Chromosome rearrangements and transposable elements. *Annual Review of Genetics*, **36**, 389-410.
- Lotsy JP (1925) Species or linneon. *Genetica*, **7**, 487-506.
- Lutz AM (1907) A preliminary note on the chromosomes of *Oenothera lamarckiana* and one of its mutants, *O. gigas* *Science*, **26**, 151-152.
- Mable BK (2004) 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. *Biological Journal of the Linnean Society*, **82**, 453-466.
- Mallet J (2007) Hybrid speciation. *Nature*, **446**, 279-283.
- Martienssen (2010) Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of *Arabidopsis*. *New Phytologist*, **186**, 46-53
- McClintock B (1984) The significance of responses of the genome to challenge. *Science*, **226**, 792-801.
- Mestiri I, Chagué V, Tanguy A-M, Huneau C, Huteau V, Belcram H, Coriton O, Chalhoub B, Jahier J. Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. *New Phytologist*, **186**, 86-101
- Mishler BD, Theriot EC (2000) The phylogenetic species concept (sensu Mishler and Theriot): monophyly, apomorphy, and phylogenetic species concepts. Pages 44–54. In: *Species Concepts and Phylogenetic Theory: A Debate* (eds. Wheeler QD & Meier R), pp. 44-54. Columbia University Press, New York, NY.

- Ownbey M (1950) Natural hybridization and amphiploidy in the genus *Tragopogon*. *American Journal of Botany*, **37**, 487-499.
- Pannell JR, Obbard DJ, Buggs RJA (2004) Polyploidy and the sexual system: what can we learn from *Mercurialis annua*? *Biological Journal of the Linnean Society*, **82**, 547-560.
- Parisod C, Holderegger R, and Brochmann C (2010) Evolutionary consequences of autopolyploidy. *New Phytologist*, **186**, 5-17.
- Paun O, Forest F, Fay MF, Chase MW (2009) Hybrid speciation in angiosperms: parental divergence drives ploidy. *New Phytologist*, **182**, 507-518.
- Ramanna MS, Jacobsen E (2003) Relevance of sexual polyploidization for crop improvement – a review. *Euphytica*, **133**, 3-8.
- Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, **29**, 467-501.
- Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16**, 351-358.
- Rieseberg LH, Raymond O, Rosenthal DM, *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 1211-1216.
- Rieseberg LH, Willis JH (2007) Plant speciation. *Science*, **317**, 910-914.
- Salmon A, Flagel L, Ying B, Udall JA, Wendel JF. Homoeologous nonreciprocal recombination in polyploid cotton. *New Phytologist*, **186**, 123-134.
- Soltis DE, Albert VA, Leebens-Mack J, *et al.* (2009) Polyploidy and angiosperm diversification. *American Journal of Botany*, **96**, 336-348.
- Soltis DE, Soltis PS (1993) Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences*, **12**, 243-273.
- Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution*, **14**, 348-352.
- Soltis DE, Soltis PS, Pires JC, *et al.* (2004) Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. *Biological Journal of the Linnean Society*, **82**, 485-501.
- Soltis DE, Soltis PS, Schemske DW, *et al.* (2007) Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon*, **56**, 13-30.
- Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 7051-7057.
- Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology*, **60**, 561-588.
- Stebbins GL (1950) *Variation and Evolution in Plants* Columbia University Press, New York.
- Stebbins GL (1971) *Chromosomal Evolution in Higher Plants* Edward Arnold, London.
- Stebbins GL (1984) Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. *Botanica Helvetica*, **94**, 1-13
- Stebbins GL (1985) Polyploidy, hybridization, and the invasion of new habitats. *Annals of the Missouri Botanical Garden*, **72**, 824-832.
- Stelkens R, Seehausen O (2009) Genetic distance between species predicts novel trait expression in their hybrids. *Evolution*, **63**, 884-897.

- Strasburger E (1910) Chromosomenzahl. *Flora*, **100**, 398-446.
- Udall JA, Quijada PA, Osborn TC (2005) Detection of chromosomal rearrangements derived from homeologous recombination in four mapping populations of *Brassica napus* L. *Genetics*, **169**, 967-979.
- Wendel J, Doyle J (2005) Polyploidy and evolution in plants. In: *Plant Diversity and Evolution* (ed. Henry RJ), pp. 97-117. CABI, Wallingford, UK.
- Wiley EO (1981) *Phylogenetics: The theory and practice of phylogenetic systematics*. John Wiley and Sons, New York.
- Winge Ø (1917) The chromosomes: their numbers and general importance. *Comptes Rendus des Travaux du Laboratoire Carlsberg*, **13**, 131-275.

Figure 1. Winge's (1917) schematic view of the occurrence of doubled chromosome numbers through hybridization: *a*, two dissimilar gametes each with nine chromosomes; *b*, the hybrid zygote; *c*, "indirect chromosome binding" in the zygote; *d*, the hybrid gametes produced on reductive division with 18 chromosomes each. Reproduced from Winge (1917) p. 200.

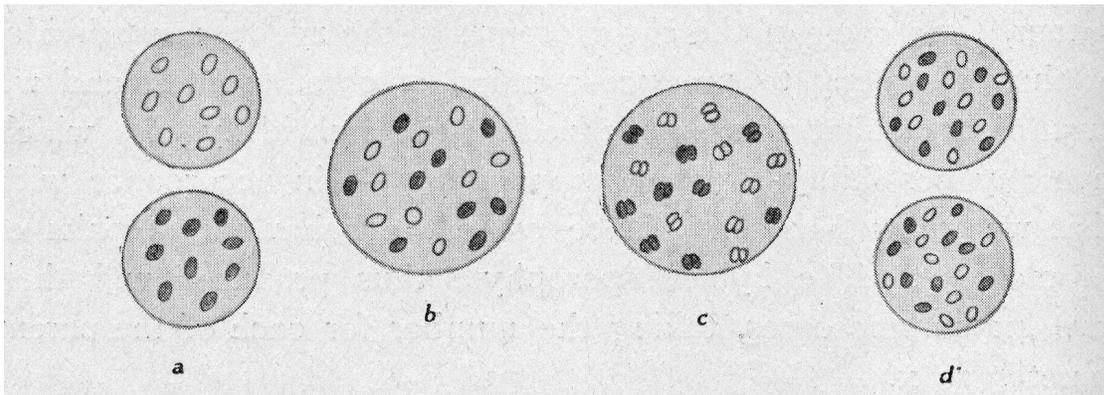


Figure 2. Illustration of Darlington's rule, showing the basis of an inverse correlation of the fertility of a diploid and tetraploid. The fertile diploid gives a tetraploid with quadrivalents having inherently irregular division. The fertile tetraploid is derived from a diploid having segregation of dissimilar chromosomes. Reproduced from Darlington 1937 p. 196.

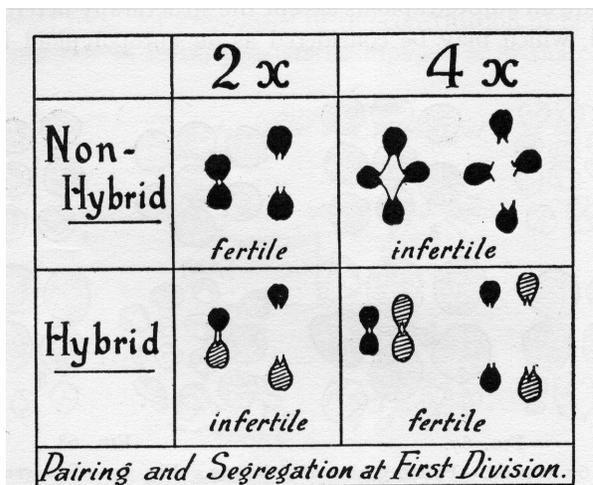


Figure 3. Suggested relationships of the genetic divergence between parents and the frequency of polyploids (dashed lines) versus homoploid hybrids (solid lines). A. The approximate distributions found by Chapman and Burke (2008), Buggs *et al.* (2008) and Paun *et al.* (2009). Buggs *et al.* (2008, 2009) show that the mean of the polyploid distribution fits the mean divergence expected if crossing occurs at random between all diploid species in a genus. B. The distribution predicted by Buggs *et al.* (2008, 2009) if autopolyploids and cryptic allopolyploids were included. C. The distribution predicted by Paun *et al.* (2009) if autopolyploids were included.

