New Phytologist

Research

Hybrid speciation in angiosperms: parental divergence drives ploidy

Ovidiu Paun, Félix Forest, Michael F. Fay and Mark W. Chase

Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

Author for correspondence:
Ovidiu Paun
Tel: +44 2083325378
Email: o.paun@kew.org

Received: *24 October 2008*
Accepted: *16 December 2008*

New Phytologist (2009) **182**: 507–518
doi: 10.1111/j.1469-8137.2009.02767.x

Key words: adaptation, allopolyploidy, homoploid hybrid, hybridization, parental divergence, polyploidy, speciation.

Summary

- Hybridization and polyploidy are now hypothesized to have regularly stimulated speciation in angiosperms, but individual or combined involvement of these two processes seems to involve significant differences in pathways of formation, establishment and evolutionary consequences of resulting lineages. We evaluate here the classical cytological hypothesis that ploidy in hybrid speciation is governed by the extent of chromosomal rearrangements among parental species.
- Within a phylogenetic framework, we calculate genetic divergence indices for 50 parental species pairs and use these indices as surrogates for the overall degree of genomic divergence (that is, as proxy for assessments of dissimilarity of the parental chromosomes).
- The results confirm that genomic differentiation between progenitor taxa influences the likelihood of diploid (homoploid) versus polyploid hybrid speciation because genetic divergence between parents of polyploids is found to be significantly greater than in the case of homoploid hybrid species.
- We argue that this asymmetric relationship may be reinforced immediately after hybrid formation, during stabilization and establishment. Underlying mechanisms potentially producing this pattern are discussed.

Introduction

At first, hybridization might seem 'a reversal in the process of evolutionary divergence' (Grant, 1981, p. 195), but in fact hybridization appears to regularly stimulate plant speciation: the combination of different genomes in hybrid lineages has extensive evolutionary and ecological implications, potentially facilitating evolutionary innovation and adaptive radiation (Andersson, 1949; Stebbins, 1950; Grant, 1981; Arnold, 1997; Barton, 2001; Rieseberg *et al.*, 2003; Seehausen, 2004; Mallet, 2007; Paun *et al.*, 2007).

Hybrid speciation refers to the mode of origin of a new species in which gene flow between species plays a major role. More than 25% of plant species seem to be involved in hybridization with other species (Mallet, 2005), but its frequency seems to vary considerably between groups, for example being more prevalent in rapidly radiating lineages (see Ellstrand *et al.*, 1996). Closely related species are most likely to hybridize, but this phenomenon often persists for millions of years after initial diversification (Mallet, 2005). The rate of hybrid speciation is definitely much lower than the statistic of 25% owing to many disadvantages that early generation hybrids need to overcome to achieve successful establishment (e.g. reduced fertility and viability, lack of reproductive and ecological isolation from the parents, lack of mates of the same type, hybrid dysgenesis and necrosis, etc.). An estimate from five regional floras indicated that *c.* 11% of species are putative hybrids (Ellstrand *et al.*, 1996).

Homoploid hybrid speciation ('recombinational', *sensu* Grant, 1981) appears to be facilitated by several factors, for example, availability of a suitable ecological niche or an available fitness peak, and rapid chromosomal evolution (Rieseberg, 1997; Mallet, 2007). To be evolutionarily successful, even fertile and 'stable' homoploid hybrids must be reproductively isolated from the parental species either by sorting genic or chromosomal sterility factors that already differentiate the parental species (Grant, 1981; Rieseberg, 2001; Wu, 2001) or by prezygotic barriers, such as spatial/temporal isolation and/or divergence into a new ecological niche (Rieseberg, 1997; Gross & Rieseberg, 2005). Indeed, hybrids may combine characteristics from both parents and/or exhibit transgressive traits that allow ecological distinctiveness (Andersson, 1949; Arnold, 1997; Rieseberg *et al.*, 2003; Seehausen, 2004).

Homoploid hybrid speciation seems to proceed at a rapid tempo (Rieseberg *et al.*, 1996; Rieseberg, 1997; Buerkle & Rieseberg, 2008), and diploid hybrid genomes are likely to be stabilized quickly, for example after 10–60 generations in the case of *Helianthus anomalus* (Ungerer *et al.*, 1998). However, Buerkle & Rieseberg (2008) have recently shown that in *Helianthus* recombination continues to shape genomic composition of homoploid hybrid species for hundreds of generations. Even at this scale, diploid hybrid speciation can still be considered one of the fastest modes of speciation.

Hybrid speciation may, however, happen much more suddenly when combined with polyploidy, which immediately provides a hybrid with a high degree of post-zygotic reproductive isolation from its progenitors: backcrossing to either parent will produce nonviable or mostly sterile offspring of odd-numbered ploidy (triploids, pentaploids etc.: Stebbins, 1950, p. 308; Grant, 1981; Ramsey & Schemske, 1998). Allopolyploidy can be the product of gametic nonreduction (frequently via a 'triploid bridge'; Ramsey & Schemske, 1998), and, more rarely, can also result from somatic chromosome doubling of a homoploid hybrid or polyspermy (Thompson & Lumaret, 1992; Ramsey & Schemske, 1998; Mallet, 2007). Of all these pathways, nonreduction during meiosis seems to be the most frequent route to polyploidy, as parents of spontaneous polyploids often produce a substantial number of unreduced gametes (see reviews by Thompson & Lumaret, 1992; Ramsey & Schemske, 1998).

Even if allopolyploidy can be viewed as abrupt or saltational speciation (Mallet, 2007), most neopolyploids will fail to become established because of meiotic abnormalities (Ramsey & Schemske, 2002) and/or their isolation, resulting in a frequency-dependent minority cytotype disadvantage (Husband, 2000). However, the latter may be overcome with the help of perenniality, asexual reproduction, assortative mating and loss of self-incompatibility barriers. Originating in sympatry (or parapatry) with progenitors, allopolyploids still require niche divergence to escape direct competition with parental taxa (Coyne & Orr, 2004). The co-joined genomes in polyploids usually have to face a complicated process of reorganization before full stabilization: chromosomal rearrangements within parental genomes, loss of low-copy DNA sequences, epigenetic effects on expression in duplicated genes and activation of transposable elements (for reviews see Comai, 2005; Chen, 2007; Paun *et al.*, 2007). Such genomic responses also have the potential to induce novel expression patterns, which together with permanent heterozygosity (potentially resulting in hybrid vigour) and gene redundancy, might result in significant shifts in morphology, breeding system and ecological tolerances, and, finally, in elevated evolutionary flexibility and major 'jumps' in evolution (De Bodt *et al.*, 2005; Comai, 2005; Otto, 2007; Paun *et al.*, 2007).

Speciation via polyploidy is likely to be a major mode of sympatric speciation in plants. A model-based estimated frequency of polyploid (usually allopolyploid) speciation in angiosperms points to at least 2–4% of recent speciation events (Otto & Whitton, 2000). However, recent direct estimates indicate that 15–25% of angiosperm speciation events are accompanied by an increase in ploidy (Wood & Rieseberg, 2005). Moreover, up to 70% of extant flowering plant species are currently polyploids (Otto & Whitton, 2000); in fact all angiosperms have descended from polyploid ancestors and are paleopolyploids (N.B. except probably *Amborella*; De Bodt *et al.*, 2005; Cui *et al.*, 2006; Soltis *et al.*, 2009). Meyers & Levin (2006) suggested that the abundance of polyploids may result from a simple ratcheting mechanism; they argued that in evolution chromosome number can double but not halve. However, genome size (as DNA amount and chromosome number) can decrease, for example as observed in *Nicotiana* sect. *Suaveolentes* where multiple chromosome fusions have resulted in chromosome number reduction (Chase *et al.*, 2003).

Because polyploidy and hybridization have been so central to plant evolution, it is important to identify processes responsible for origins of hybrid species and those that promote shifts in ploidy, changing the possible outcomes of hybrid speciation. The interest here is not simply limited to predicting results of hybrid evolution and of polyploid dynamics, it is also of great importance for our understanding of evolutionary processes that result in isolation between species, including those that influence establishment of new taxa and maintain biodiversity.

A relevant hypothesis was proposed in the early twentieth century: the level of (structural) differentiation between ancestor genomes influences ploidy of successful hybrids (Winge, 1917; Darlington, 1937; Stebbins, 1950). Winge (1917, as cited by Darlington, 1937), for example, considered that polyploid formation after somatic doubling of homoploid hybrids would be stimulated by the need for a partner with which chromosomes could pair. Therefore, higher chromosomal differentiation between parents would increase the chance of shifts in ploidy. Decades later, Grant (1981, pp. 247–248 and 320) referring also to the initial formation of an allopolyploid, stated that pre-existing chromosomal rearrangements within parental genomes 'upset the course of meiosis in the hybrid', resulting in reduced pairing, and that the latter 'sets (the stage) for (gamete) nonreduction and amphiploid formation'. Other authors extended this idea by referring more to the moment of polyploid establishment, rather than initiation. Darlington (1937, p.136), for example wrote: 'the characteristic properties of hybrids depend not on the properties of the parents, but on the differences between these properties'. He considered that a 'differential affinity' between parental chromosomes governs long-term successful pairing in structural hybrids and polyploids (pp. 160, 172, 199): 'The greater the (parental) dissimilarities, the more regularly do the identical chromosomes pair in the allotetraploid derived, and therefore the less frequent are the multivalents in the tetraploid'. In 1945 Clausen *et al.* (as cited by Buggs *et al.*,

2008) reached the conclusion that the 'success and constancy' of allopolyploids must be linked with the 'degree of relationships' found between their parents. Even Stebbins (1950, p. 354) referred to the genetic relationships of the parental diploid species to each other as one of the factors promoting development of allopolyploidy in plants.

The potential cause–effect relationship between the level of chromosomal (genetic) divergence of the parents and ploidy of hybrids has recently been revisited by Chapman & Burke (2007) and Buggs *et al.* (2008), who, from different perspectives, reached partly contradictory conclusions. Based on 11 cases of homoploid hybrids (plus a misclassified polyploid *Eupatorium*) and 26 cases of allopolyploids, Chapman & Burke (2007) demonstrated that, in angiosperms, parental nuclear ribosomal internal transcribed spacer (ITS) divergence is significantly greater for allopolyploids than for homoploid hybrids. However, the method employed disregarded the variable substitution rates expected across such unrelated cases even in the same molecular marker (a caveat also discussed by the authors) and it included in the analysis hybrids formed by parents with different basic chromosome numbers (e.g. *Arabidopsis suecica*, *Spiranthes diluvialis* and *Symphyotricum ascendens*) or even different ploidies (*Artemisia douglasiana*, *Primula scotica* and *Rubus maximus*). Hybrid speciation starting from such parental pairs is particularly prone to result in allopolyploids and might follow special routes and rules (Ramsey & Schemske, 1998). By contrast, Buggs and collaborators (2008) took a molecular phylogenetic approach to the issue, but relied on subjectively defined clades as a measure of genetic divergence. Moreover, the latter study considered any naturally occurring hybrid individual reported in the literature within eight selected plant genera and, therefore, focused on polyploid formation not evolutionary success (effective speciation). Long-term success in meiosis is key to operation of the mechanism that governs ploidal shifts, which means that only taxa that appear to be valid species in their own right should be included in the calculations. Therefore, by including ephemeral homoploid hybrids, sterile triploids and neopolyploids, Buggs *et al.* (2008) did not directly evaluate the classical cytological hypothesis and failed to find convincing evidence showing that ploidal increase in established species is determined by the phylogenetic distance between progenitor species. However, their findings point to a restriction of homoploid formation to parental pairs less divergent than expectation if crossings were random between all species pairs in a genus.

In the light of this recent debate, we approach the potential relationship between ancestor divergence–descendant ploidy by uniting methodologically the two recent studies mentioned earlier. Like Chapman & Burke (2007) we use the extent of genetic divergence between parental pairs as a surrogate for chromosomal differentiation, but we attempt to standardize the method by taking into account the rate of evolution in the respective marker(s) and genus from a phylogenetic approach. We extend the sampling to more cases, but we include only diploid parental pairs with identical base chromosome numbers and only fertile, successful hybrids that have a long species history.

Materials and Methods

Selection of taxa

This analysis is based on 50 case studies (Table 1) chosen from the literature following several criteria: (1) the hybrid status for the respective species has been documented with some certainty by molecular means in addition to (at least) morphology; (2) an extensive and representative molecular phylogenetic analysis for the genus including the parental taxa was already available; (3) the parents are diploids and have the same chromosome number; and (4) the hybrids are natural and stable, with proven evolutionary success (neopolyploids and unnamed suspected hybrids were excluded). Owing to methodological constraints, we did not consider in our analysis intergeneric hybrids (e.g. allotetraploid *Triticum turgidum*), hybrids produced by more than two parents (e.g. homoploid hybrid *Iris nelsonii*) or hybrids from genera of uncertain delimitation (e.g. *Tarasa* and *Brassica*). The last exclusions were followed to try to minimize the influence of clearly artificial taxonomies.

We classified the data into three categories: (1) homoploid hybrids ($n = 16$); (2) allopolyploids ($n = 32$); and (3) two cases of both diploid and polyploid hybrids formed by the same parental species (Table 1). We counted the parental pairs, so we considered just once the instances where more than one homoploid or polyploid hybrid was formed by the same parental pair. In this way, we were able to identify half as many homoploid hybrids as allopolyploid species. This pattern may result from a biological rarity of homoploid hybrid speciation versus allopolyploid speciation, but it also mirrors a more general problem, namely that detecting and rigorously documenting homoploid hybrid species is much more difficult than detecting polyploid ones (Rieseberg, 1997).

Molecular data and statistical analyses

Some DNA sequence matrices were obtained directly from authors of published analyses (see the Acknowledgements); for the others, DNA sequences were collected from GenBank (http://www.ncbi.nlm.nih.gov) and realigned using CLUSTAL W (http://www.ebi.ac.uk/Tools/clustalw/; Chenna *et al.*, 2003). Based only on ingroup taxa (species within genera), we calculated with PAUP* 4.0b10 (Swofford, 2003) all intrageneric pairwise genetic distances, using both uncorrected p-distances (P) and Kimura's (1980) two-parameter (K2P) distances. Uncorrected P is the observed number of changes between two sequences, with no correction for multiple changes.

Table 1 Details of hybrids included in this analysis (taxonomy used follows the original papers; P-GDI, uncorrected P-derived parental genetic divergence index)

Hybrid	Parental pair	P-GDI	Reference hybrid	Reference phylogeny/sequences
Homoploid hybrids				
Achillea roseoalba	*Achillea setacea* × *Achillea asplenifolia*	0.47	Guo *et al.* (2004, 2005)	Guo *et al.* (2004)
Actinidia persicina, Actinidia zhejiangensis	*Actinidia hemsleyana* × *Actinidia eriantha*/*Actinidia styracifolia*	0.85	Li *et al.* (2002)	Li *et al.* (2002); Chat *et al.* (2004)
Argyranthemum lemsii/*sundingii*	*Argyranthemum broussonetii* × *Argyranthemum frutescens*	0	Brochmann *et al.* (2000); Fjellheim *et al.* (2009)	J. Francisco-Ortega *et al.* (unpublished, L77739, L77784-5, L77788-99, L77801)
Arisaema ehimense	*Arisaema tosaense* × *Arisaema serratum*	0	Maki & Murata (2001)	Renner *et al.* (2004)
Berberis bidentata	*Berberis darwinii* × *Berberis trigona*	0.07	Bottini *et al.* (2007)	Kim *et al.* (2004)
Encelia virginiensis	*Encelia actoni* × *Encelia frutescens*	0.46	Allan *et al.* (1997)	Fehlberg & Ranker (2007)
Gossypium bickii	*Gossypium sturtianum* × *Gossypium australe*	0.52	Seelanan *et al.* (1999)	Seelanan *et al.* (1997, 1999); Liu *et al.* (2001)
Helianthus anomalus, Helianthus deserticola, Helianthus paradoxus	*Helianthus annuus* × *Helianthus petiolaris*	0.41	Rieseberg *et al.* (1996, 2003)	Schilling *et al.* (1998)
Hippophae goniocarpa	*Hippophae rhamnoides* ssp. *sinensis* × *Hippophae neurocarpa* ssp. *neurocarpa*	1.02	Sun *et al.* (2002); Wang *et al.* (2008)	Sun *et al.* (2002); Wang *et al.* (2008)
H. goniocarpa subsp. *litangensis* (*Hippophae litangensis*)	*H. rhamnoides* subsp. *yumanensis* × *H. neurocarpa* subsp. *stellatopilosa*	1.25	Sun *et al.* (2002)	Sun *et al.* (2002); Wang *et al.* (2008)
Hyobanche glabrata	*Hyobanche sanguinea* × *Hyobanche rubra*	0.86	Wolfe & Randle (2001)	Wolfe & Randle (2001)
Lithophragma thompsonii	*Lithophragma tenellum* × *Lithophragma parviflorum*	0.64	Kuzoff *et al.* (1999)	Kuzoff *et al.* (1999)
Paeonia anomala, Paeonia emodii	*Paeonia veitchii* × *Paeonia lactiflora*	0.38	Sang *et al.* (1997); Pan *et al.* (2007)	Sang *et al.* (1997)
Penstemon clevelandii	*Penstemon spectabilis* × *Penstemon centrathifolium*	0.35	Wolfe *et al.* (1998)	Wolfe *et al.* (2006)
Scaevola kilaueae	*S. coriacea* × *S. chamissoniana*	0.29	Howarth & Baum (2005)	Howarth *et al.* (2003)
Scaevola procera	*Scaevola gaudichaudii* × *Scaevola mollis*	0.12	Howarth & Baum (2005)	Howarth *et al.* (2003)
Allopolyploids				
Achillea alpina, Achillea wilsoniana	*Achillea asiatica* × *Achillea acuminata*	1.41	Guo *et al.* (2006)	Guo *et al.* (2004)
Actinidia callosa var. *strigillosa*	*Actinidia callosa* × *Actinidia chinensis*	1.03	Li *et al.* (2002; Chat *et al.* (2004)	Li *et al.* (2002); Chat *et al.* (2004)
Actinidia cylindrica var. *reticulata*	*A. cylindrica* var. *cylindrica* × *Actinidia eriantha*	1.06	Chat *et al.* (2004)	Li *et al.* (2002); Chat *et al.* (2004)
Arachis hypogaea	*Arachis duranensis* × *Arachis ipaensis*	0.5	Jung *et al.* (2003)	M. D. Bechara *et al.* (unpublished, AY615215-67)
Centaurium bianoris	*Centaurium maritimum* × *Centaurium tenuiflorum* var. *acutiflorum*	1.77	Mansion *et al.* (2005); Guggisberg *et al.* (2006)	Mansion *et al.* (2005)
Centaurium × *tenuiflorum*	*C. tenuiflorum* subsp. *acutiflorum* × *Centaurium erythraea* subsp. *erythraea*	1.28	Mansion *et al.* (2005)	Mansion *et al.* (2005)
Clarkia delicata	*Clarkia epilobioides* × *Clarkia unguiculata*	0.84	Ford & Gottlieb (2002)	Levin *et al.* (2004, AY271529-38); Chapman & Burke (2007, EF017398, EF017400-1, EF017404)
Clarkia similis	*C. epilobioides* × *C. modesta*	0.92	Ford & Gottlieb (2002)	Levin *et al.* (2004, AY271529-38); Chapman & Burke (2007, EF017398, EF017400-1, EF017404)

New Phytologist (2009) **182**: 507–518
www.newphytologist.org

Table 1 *continued*

Hybrid	Parental pair	P-GDI	Reference hybrid	Reference phylogeny/sequences
Coffea arabica	*Coffea eugenioides* × *Coffea canephora*	1.37	Maurin *et al.* (2007)	Maurin *et al.* (2007)
Dactylorhiza armeniaca	*Dactylorhiza euxina* × *Dactylorhiza incarnata*	0.5	Hedrén (2001)	Pillon *et al.* (2006)
Dactylorhiza angustata, Dactylorhiza baltica, Dactylorhiza majalis, Dactylorhiza traunsteineri	*Dactylorhiza fuchsii* × *D. incarnata*	1.57	Pillon *et al.* (2007)	Pillon *et al.* (2006)
Dactylorhiza urvilleana	*D. euxina* × *D. saccifera/D. fuchsii*	1.36	Hedrén (2001)	Pillon *et al.* (2006)
Draba ladina	*Dactylorhiza tomentosa* × *Dactylorhiza aizoides*	1.39	Widmer & Baltisberger (1999)	Koch & Al-Shehbaz (2002)
Erythronium elegans, Erythronium quinaultense	*Erythronium montanum* × *Erythronium revolutum*	0.82	Allen (2001)	Allen *et al.* (2003)
Gossypium barbadense, Gossypium darwinii, Gossypium hirsutum, Gossypium mustelinum, Gossypium tomentosum	*Gossypium arboreum/Gossypium herbaceum* × *Gossypium raimondii*	1.63	Liu *et al.* (2001); Senchina *et al.*(2003); Wendel & Cronn (2003); Cronn & Wendel (2004)	Seelanan *et al.* (1997, 1999); Liu *et al.* (2001)
Helianthus ciliaris	*Helianthus arizonensis* × *Helianthus laciniatus*	0.48	Timme *et al.* (2007)	Schilling *et al.* (1998)
Hepatica henryi	*Hepatica falconeri* × *Hepatica asiatica*	1.07	Weiss-Schneeweiss *et al.* (2007)	Weiss-Schneeweiss *et al.* (2007)
Hepatica transilvanica	*Hepatica nobilis var. nobilis* × *H. falconeri*	0.69	Weiss-Schneeweiss *et al.* (2007)	Weiss-Schneeweiss *et al.* (2007)
Leucaena confertiflora	*Leucaena trichandra* × *Leucaena cuspidata*	1.32	Hughes *et al.* (2007)	Hughes *et al.* (2007)
Leucaena diversifolia	*Leucaena pulverulenta* × *Leucaena trichandra*	1.3	Hughes *et al.* (2007)	Hughes *et al.* (2007)
Leucaena leucocephala	*Leucaena pulverulenta* × *Leucaena lanceolata*	1.42	Hughes *et al.* (2007)	Hughes *et al.* (2007)
Leucaena pallida	*Leucaena pueblana/Leucaena matudae* × *Leucaena lempirana*	1.1	Hughes *et al.* (2007)	Hughes *et al.* (2007)
Lithophragma bolanderi (4x)	*L. bolanderi (2x)* × *Lithophragma glabrum*	0.85	Kuzoff *et al.* (1999)	Kuzoff *et al.* (1999)
Nicotiana arentsii	*Nicotiana undulata* × *Nicotiana wigandioides*	0.78	Chase *et al.* (2003); Clarkson *et al.* (2004)	Chase *et al.* (2003); Clarkson *et al.* (2004)
Nicotiana nesophila, Nicotiana nudicaulis, Nicotiana repanda, Nicotiana stocktonii	*Nicotiana sylvestris* × *Nicotiana obtusifolia (Nicotiana trigonophylla)*	1.13	Chase *et al.* (2003); Clarkson *et al.* (2004)	Chase *et al.* (2003); Clarkson *et al.* (2004)
Nicotiana clevelandii, Nicotiana quadrivalvis (Nicotiana bigelovii)	*N. obtusifolia (N. trigonophylla)* × *Nicotiana attenuata*	1.11	Chase *et al.* (2003); Clarkson *et al.* (2004)	Chase *et al.* (2003); Clarkson *et al.* (2004)
Nicotiana rustica	*Nicotiana paniculata* × *N. undulata*	1.19	Chase *et al.* (2003); Clarkson *et al.* (2004)	Chase *et al.* (2003); Clarkson *et al.* (2004)
Nicotiana tabacum	*N. sylvestris* × *Nicotiana tomentosiformis*	1.21	Chase *et al.* (2003); Clarkson *et al.* (2004)	Chase *et al.* (2003); Clarkson *et al.* (2004)
Oryza eichingeri, Oryza minuta	*Oryza punctata* × *Oryza officinalis/Oryza rhizomatis*	0.75	Ge *et al.* (1999)	Ge *et al.* (1999)
Stylosanthes aff. *calcicola*	*Stylosanthes calcicola* × *Stylosanthes viscosa*	0.9	Vander Stapen *et al.* (2002)	Vander Stapen *et al.* (2002)
Tragopogon castellanus	*Tragopogon lamottei* × *Tragopogon crocifolius*	1.14	Buggs *et al.* (2008)	Mavrodiev *et al.* (2005)
Tragopogon tuberosus	*Tragopogon* sect. *Collini* × *Tragopogon pusillus*	1.19	Buggs *et al.* (2008)	Mavrodiev *et al.* (2005)
Cases of homoploid and polyploid hybrids with the same parental pair				
Paeonia cambessedesii (2x), P. russi (4x)	*Paeonia lactiflora* × *Paeonia mairei*	0.71	Sang *et al.* (1997)	Sang *et al.* (1997)
Stephanomeria diegensis (2x), S. elata (4x)	*Stephanomeria exigua* subsp. *deanei* × *Stephanomeria virgata*	0.62	Lee *et al.* (2002)	Lee *et al.* (2002)

Table 2 Nonparametric comparisons of parental divergence indices for homoploid versus polyploid hybrid species using the Mann–Whitney test as calculated in SPSS

Type of data	Genetic distance	*N*	Mann-Whitney U	Z	*P*
Overall	P	48	55.0	–4.34	< 0.0001
Overall	K2P	48	55.0	–4.40	< 0.0001
Nuclear only	P	45	79.0	–3.15	< 0.0001
ITS only	P	45	76.0	–3.59	< 0.0001

The indices are calculated using either the uncorrected-p (P) or Kimura's (1980) two-parameter (K2P) distance.

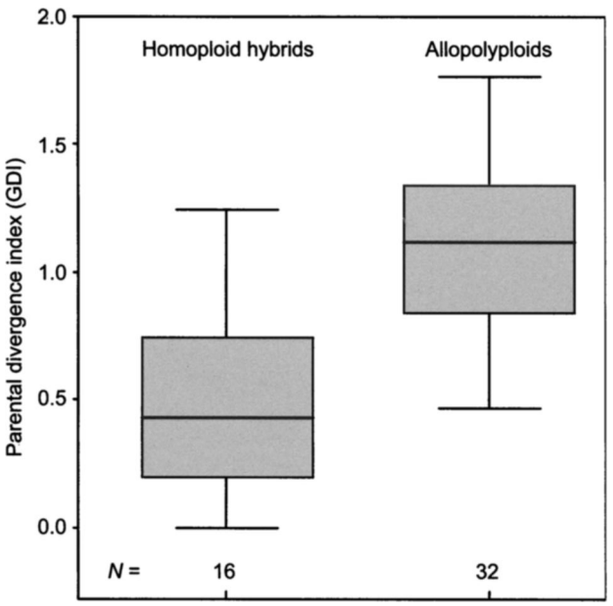

Fig. 1 Box plots of the distribution of genetic divergence index (GDI) of parental pairs for homoploid and polyploid hybrids. The two groups have an asymmetric dispersion range, with the parents of allopolyploids being more divergent than those producing diploid hybrid species (Mann-Whitney U test, $P < 0.0001$, see text). The difference in sample sizes probably reflects the greater difficulty of identifying homoploid hybrids.

By contrast, the K2P model addresses this problem by considering equal base frequencies but different rates for transitions and transversions (Kimura, 1980).

Because species-level phylogenetic analyses use molecular markers exhibiting different substitution rates, we standardized our data among cases by calculating for each parental pair a genetic divergence index (GDI). For each instance, the genetic distance between parental pairs (Pd) was divided by the average genetic distance (Av) in the genus based on the same molecular markers. Under this definition, GDI is always positive; if GDI > 1.0, then Pd is higher than Av. When multiple sequences were available for a given taxon, an average of the genetic distance for all possible parental pairs was used in further analyses.

To check for potential bias in our analysis created by uneven sampling in phylogenetic trees, we performed a nonparametric, one-tailed Spearman rank order correlation of Av with the number of taxa included in each tree for both homoploid and allopolyploid species.

All statistical analyses were performed using SPSS 15.0 (SPSS, Chicago, IL, USA). As Pd, Av and GDI are not expected to be normally distributed, we treated our data as nonparametric.

Results

The two genetic measures applied in this study, P and K2P, gave significantly congruent results (Spearman's correlation coefficient based upon ranks rho = 1, $P < 0.0001$, independently for Pd, Av and GDI). As expected, K2P values of Pd and Av were generally slightly higher then those calculated with P (see the Supporting Information, Table S1). However, GDI values based on the two genetic distances were identical up to the second decimal, confirming the value of our standardizing approach. In the following tests, we generally focused on the P-derived GDI because of the simpler assumptions.

Nonparametric comparisons of GDI values (calculated overall, exclusively on nuclear data, or just with nuclear ribosomal ITS data) for homoploid versus polyploid hybrid species using the Mann–Whitney test indicated statistically significant asymmetric relationships ($P < 0.0001$, Table 2). Parents of polyploids are generally more divergent than the average intrageneric distance (i.e. GDI > 1), whereas for most homoploid hybrids GDI is < 0.5 (Fig. 1). In addition, in all cases of direct comparisons between homoploid and polyploid hybrids in the same genus (i.e. *Achillea*, *Actinidia*, *Gossypium*, *Helianthus* and *Lithophragma*; Table 1) parents of polyploids are more divergent.

A histogram (Fig. 2a) illustrating frequency distributions of classes of parental GDI for homoploid hybrid species and allopolyploids indicates that both categories have unimodal but distinct distributions. The relationships between frequency of occurrence and degree of chromosomal divergence of parental pairs for allopolyploids and homoploid hybrids (Fig. 2b) meet at a GDI ≈ 0.75, indicating an equal probability of a hybrid formation with and without a change in ploidy when Pd is *c.* three-quarters of Av.

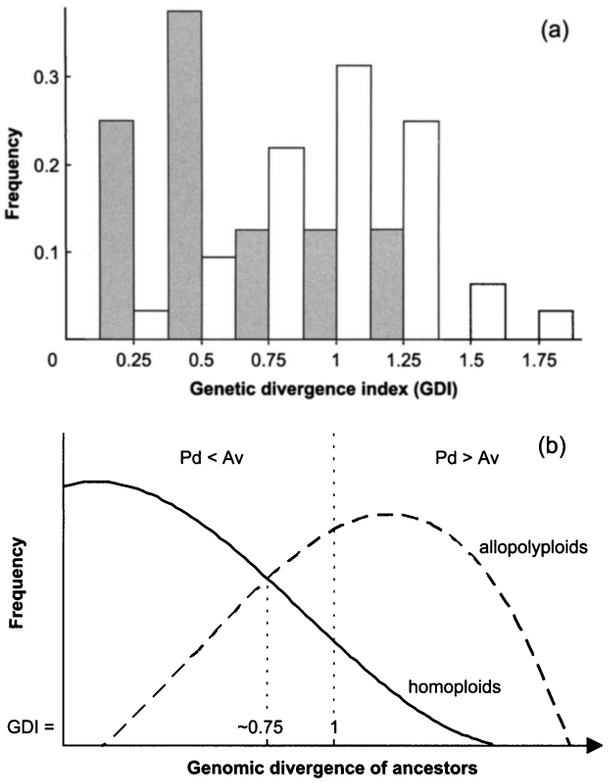

Fig. 2 (a) Histogram illustrating the different frequency distribution of parental genetic divergence index (GDI) classes for homoploid hybrid species (tinted bars) and allopolyploids (open bars). Values on the *x*-axis show the limits of the GDI classes, with a 0.25 increment. (b) Hypothetical relationships between frequency of occurrence and degree of genomic divergence of parental pairs for allopolyploids (goodness-of-fit to the data $R^2 = 0.815$) and homoploid hybrids (goodness-of-fit $R^2 = 0.804$), derived from (a). There is an equal probability of hybrid formation with and without a change of ploidy when Pd is three-quarters of Av (GDI ≈ 0.75). Pd, parental genetic distance; Av, average genetic distance in the genus being studied.

The nonparametric, one-tailed Spearman rank order correlation of Av with the number of taxa included in each phylogenetic analysis was not significant (Spearman's rho = 0.069, $P = 0.322$).

Discussion

By comparing frequency distributions of parental genetic distance (used here as a surrogate for chromosomal differentiation) for homoploid and allopolyploid hybrid species, we demonstrate the relevance of progenitor divergence as a determinant of ploidy in resulting hybrid species: although the range of genetic divergence between the parents of homoploid hybrids is similar to those of allopolyploids, the actual values of divergence are significantly higher in the latter (Fig. 1).

Our standardized approach, integrating each parental pair within its generic context, has several advantages: it makes our method independent of assumptions implied by specific genetic distances or defining clades (cf. Buggs *et al.*, 2008), it allows us to include molecular markers and more cases (cf. Chapman & Burke, 2007) and gives our analysis greater predictive power. The last derives from our suggestion that species pairs with a divergence smaller than three quarters of the average divergence between species within the genus (i.e. GDI ≤ 0.75) have chromosomes mostly displaying colinearity of genes. Most homoploid hybrids (75%) included here were formed between such parental pairs, but this category included just 12.5% of allopolyploids (Fig. 2a). Furthermore, if a parental pair has a divergence greater than three-quarters of the average in a given genus, most of their corresponding chromosomes are likely to be sufficiently heterologous to act as homeologs, and hybridization is most likely to result in an increase in ploidy.

Two cases (in *Paeonia* and *Stephanomeria*; see Table 1) identified in the literature for which the same parental pair has successfully produced both homoploid and allopolyploid hybrid species substantiate our results. Their calculated divergence index (Table 1) is indeed close to the estimated value for which there should be an equal probability of hybrid formation with and without ploidy change (i.e. GDI ≈ 0.75, Fig. 2b). Our results parallel those of Chapman & Burke (2007): their analysis indicated that parents of allopolyploids are, on average, more than twice as divergent as parents of homoploid hybrids, a significant relationship that is also visible in GDI (Fig. 1).

Assumptions, limitations and alternatives

The general premise that genetic divergence provides the best available surrogate for differentiation of chromosome sets is often employed (Edmands, 2002). As early as 1937, Darlington (p. 197) hypothesized 'a correlation between genetic differentiation of the chromosomes of the species and their structural differentiation'. Indeed, both genetic distance and magnitude of difference in genomic rearrangements between two species are expected to be proportional to the evolutionary time since common ancestry.

Calculating an average genetic distance within each genus adds a subjective component to our analyses. We cannot eliminate taxonomic inconsistencies created by differences in taxonomic practice among authors working on different taxa. We also start from the assumption that the modern taxa studied here are closely related to the actual progenitors of the hybrid species and that after hybridization genetic divergence between these species has remained largely unchanged. More appropriately, these taxa should be considered as the closest living descendants of the donor species. However, most of the cases included here, both homoploid and polyploid hybrids, are likely to be relatively recent, as with time such cases become increasingly difficult to detect (Chase *et al.*, 2003; Clarkson *et al.*, 2004).

Underlying processes: allopolyploids

A theoretical model for polyploid speciation along a continuous variation of genomic divergence between diploid progenitors was developed by Sang and collaborators (2004). They treated the origin of a successful polyploid lineage as a function of (1) polyploid formation (production of polyploid individuals from diploid populations), further broken up into probability of unreduced gamete production and frequency of hybridization, and (2) successful establishment of polyploid populations. Such a model would imply that frequency of polyploid formation would have a negative exponential distribution on parental genomic divergence. The overall probability of successful polyploid speciation, however, would have a unimodal distribution on parental divergence (Sang *et al.*, 2004), with established autopolyploids being much less frequent than allopolyploids, despite the fact that autopolyploids occur spontaneously in nature at relatively high rates (see Ramsey & Schemske, 1998).

The main route to allopolyploid speciation is represented by the fusion of an unreduced gamete with a haploid gamete resulting in a 'triploid bridge' (Ramsey & Schemske, 1998; Husband, 2000), and after self-fertilization or backcrossing to diploids a new allotetraploid may originate. Alternatively, allopolyploid speciation can also result after fusion of two unreduced gametes, with better chances in dense hybrid zones, marginal or disturbed habitats and/or other limiting conditions (e.g. temperature variation; Thompson & Lumaret, 1992; Ramsey & Schemske, 1998). Such unreduced gametes are thought to be rare (Mallet, 2007) and will be unsuccessful and lost, especially if enough haploid gametes are produced. However, poor chromosome pairing in unbalanced diploid F_1 hybrids leads to asynapsis at the first meiotic division, and such organisms form unreduced gametes with much greater frequency. Ramsey & Schemske (1998) reported that unreduced gametes are produced at rates *c.* 50 times higher in hybrids than in nonhybrid lineages. As predicted by Grant (1981, see the Introduction), it seems plausible that greater levels of parental divergence increase meiotic abnormalities at the homoploid level and, thus, the rate of nonreduction. For example, in *Lilium*, most gametes produced by intersectional hybrids are unreduced (van Tuyl *et al.*, 1989). The same trend is expected in the case of allopolyploidy resulting from somatic chromosome doubling in meristematic tissues of a diploid hybrid or in a zygote/young embryo (e.g. *Primula kewensis*; Ramsey & Schemske, 1998). Such somatic doubling may be directly triggered by structural differences between parental homeologs (Winge, 1917; see the Introduction).

However, both somatic chromosome doubling and unreduced gametes require spontaneous occurrence of diploid hybrids that are at least partly fertile and self-compatible (Sang *et al.*, 2004), thus limiting the possible maximal extent of parental divergence. Such a bounded distribution of parental divergence, between lower chances of gamete nonreduction or somatic doubling at minimal parental divergence and full diploid hybrid sterility at increased divergence, may result in apparently random occurrence of allopolyploidization events as suggested by Buggs *et al.* (2008). In addition, rare allopolyploid formation events via unreduced gametes produced directly in nonhybrid parents will not follow the rules presented above and may blur the trends concerning formation of allopolyploid individuals.

Independent of formation, a nascent allopolyploid must become established and expand its population/s in order to produce a new species. Establishment of the polyploid lineage will depend not only on stochastic events, such as the availability of appropriate environments, but also on its degree of viability, fertility, heterozygosity (hybrid vigour) and fitness. Darlington (1937, p. 196) hypothesized 'a negative correlation between the fertility of diploids and that of the tetraploids to which they give rise.... The greater the dissimilarities in the diploid, the more regularly do the identical chromosomes pair in the allotetraploid derived from it, and therefore the less frequent are the multivalents in the tetraploid'. A diploid hybrid with reduced chromosome pairing will exhibit a high degree of sterility, which doubling overcomes. Fertility of allopolyploids might increase with parental divergence, because of fewer meiotic abnormalities. Ramsey & Schemske (2002) provided evidence that the degree of allopolyploid fertility is positively correlated with frequency of bivalents, but not with other configurations. They also concluded that allopolyploids generated by semisterile diploid hybrids are generally much more fertile than their progenitors, an attribute partly reflecting genic incompatibilities independent of and in addition to meiotic behaviour. A significant increase in effective population size seems to be the consequence of selection on fertility acting on neopolyploids, which rapidly increases pollen viability and seed set. The picture here is more complex than at homoploid level, and it is still unclear how parental incompatibilities will behave at the polyploid level, when they result in sterility/reduced fitness or breakdown. This is also the result of a lack of study; allopolyploids do not facilitate gene flow between diverging parental taxa and, hence, they are not considered in current research and debates regarding reproductive isolation and genetics of speciation (see for example Widmer *et al.*, 2009).

It is expected that heterozygosity (resulting in heterosis) generally provides increased fitness and adaptive potential, through enhancing the potential for spatial, temporal and functional variation in gene expression (Flagel *et al.*, 2008; Leitch & Leitch, 2008). The proportion of homeologous loci that are stably heterozygous should be positively correlated with genomic divergence between progenitors. In addition, alterations of gene expression in allopolyploid genomes have the potential to trigger fitness differences in the parental environment, which will be available to selection. A few case studies have indicated that the extent of genomic alterations and changes in gene expression may depend on the degree of

divergence between the parental diploid genomes. For example, Song *et al.* (1995) observed fewer rearrangements in the allopolyploid genome formed from closely related *Brassica rapa* and *Brassica oleracea*, but many more in the allopolyploid combining more divergent *B. rapa* and *Brassica nigra*. Another such example is *Nicotiana* where allopolyploids *Nicotiana arentsii* and *Nicotiana rustica* show only minimal genetic changes, but *Nicotiana tabacum* (resulting from widely divergent *Nicotiana sylvestris* and *Nicotiana tomentosiformis*) exhibits intergenomic translocations (Lim *et al.*, 2004). Rapid genomic repatterning will also increase genetic variability available to new polyploid populations. In addition, extensive changes in gene expression seem to be triggered by wide hybridization rather than polyploidy (Paun *et al.*, 2007); recent studies show that genome duplication can, in fact, have widespread ameliorating effects on altered levels of gene expression arising from hybridization (as, for example in *Senecio*; Hegarty *et al.*, 2006). In allopolyploid cotton, however, a significant proportion of expression novelty is likely triggered by polyploidy after long-term evolutionary processes on duplicated genes (Flagel *et al.*, 2008). Finally, shifts in breeding system, such as the breakdown of self-incompatibility, seem to be initiated by polyploidization alone, whereas others, such as apomixis, usually seem to be triggered by the effects of combining hybridization and polyploidy (see Paun *et al.*, 2006). In conclusion, the frequency distribution of allopolyploids along the continuum of ancestral genomic divergence (Fig. 2b) will have an optimum between low fertility, heterozygosity and fitness at minimal parental divergence (as compared to diploid progenitors and homoploid hybrids) and low probability of polyploid formation towards maximal divergence of progenitors (due to increased prezygotic and postzygotic barriers).

Autopolyploids and allopolyploids

Buggs *et al.* (2008) argued that the low frequency of allopolyploids in lower parental divergence classes in the study of Chapman & Burke (2007) was caused by exclusion of autopolyploids from analysis of the latter. Similarly, our study does not include any autopolyploids, but we definitely expect that parental divergence in this case should not span such a high interval (at least until GDI = 1, i.e. Pd = Av) to make the polyploid distribution fit a negative linear/exponential function (Fig. 2). The overall polyploid distribution would instead be bimodal, indicating the presence of different phenomena influencing the frequency of the two types of polyploids (see also Sang *et al.*, 2004). We hypothesize that the adaptive valley between autopolyploid and allopolyploid frequency may have resulted from more or less equal combinations of bivalents and quadrivalents in meiosis. However, we regard inclusion of autopolyploids in analyses considering hybrid speciation as inappropriate because autopolyploid formation should correspond, at the diploid level, to frequent events of intraspecific processes, and these do not necessarily have a significant contribution to speciation.

Underlying processes: homoploid hybrids

The formation of homoploid hybrid individuals is partly shaped by stochastic events (e.g. shifts in distribution ranges) but is directly limited by the strength and nature of reproductive isolation between species pairs. Plant species are typically isolated by many prezygotic and postzygotic barriers and their complex interactions (Coyne & Orr, 2004). Recent studies have suggested that prezygotic isolation is usually much stronger than postzygotic isolation (Lowry *et al.*, 2008; Widmer *et al.*, 2009; but see Cozzolino *et al.*, 2004), because of factors such as distribution, immigrant inviability, phenological differences, pollinator specificity, mating system and pollen competition. Significant trends obvious in hybrid formation without a change in ploidy (Buggs *et al.*, 2008) represent, to a great degree, the entire speciation process in homoploid hybrids (this study). Studying hybrid formation, Buggs and colleagues (2008) reached the conclusion that parents of homoploid hybrids are less divergent than would be expected with random crossing. We found a similar, but more significant pattern (Fig. 2): the probability of production of a diploid hybrid is highest if the parental divergence is less than or equal to (more or less) half the average in the genus and decreases as reproductive barriers (both prezygotic and postzygotic) become stronger between diploid parents with increased genomic divergence. With greater parental divergence, genic and/or chromosomal incompatibility will occur with greater probability. Examples of characterized genic incompatibilities include: genes involved in hybrid necrosis or weakness (Bomblies & Weigel, 2007) and cytonuclear incompatibility (Chase, 2007). Models have suggested that the level of incompatibility between species should increase with evolutionary time at least as rapidly as the square of the divergence time between the two species (snowballing effect; Orr, 1995). Therefore, the degree of fertility of a diploid hybrid decreases rapidly with increased genomic divergence. Classical models of chromosomal speciation have stated that after a particular level, meiotic mismatches of parental chromosomes or karyosignificant reduction in fitness (White, 1978; but see Lowry *et al.*, 2008). In addition, chromosomal divergence may also increase the strength of genic incompatibilities by suppressing recombination and therefore maintaining the effects of linked isolation genes (Rieseberg, 2001).

The diploid form of hybrid speciation occurs mainly in sympatry or parapatry, and hence the greatest challenge of the nascent taxon is to achieve reproductive isolation. Isolation from progenitors often occurs as a byproduct of the process that stabilizes the hybrid lineage and may be based on ecological factors (e.g. habitat divergence; Gross & Rieseberg, 2005), sorting pre-existing sterility factors and/or chromosomal

rearrangements (Grant, 1981). All these pathways are potentially influenced by the extent of parental differentiation. Ecological divergence may be acquired by positive heterosis (hybrid vigour; Lippman & Zamir, 2007) and transgressive segregation (Gross & Rieseberg, 2005). Numerous experimental crosses have suggested that the optimal degree of genetic divergence for maximal expression of positive heterosis occurs within a range of divergence that is narrow enough for cytological irregularities not to be apparent (Moll *et al.*, 1965; Cox & Murphy, 1990). Reproductive isolation can also be achieved by sorting pre-existing chromosomal rearrangements that differentiate the parental species, resulting in the formation of a novel recombinant genotype that is homozygous for these rearrangements ('recombinational' speciation, Grant, 1981, p. 250). Stronger genetic isolation from progenitors is most likely when a barrier differentiating the parents is genetically and/or chromosomally extensive and complex (Rieseberg, 2000) However, the parental species must have genomes similar enough for pairing and recombination to occur.

There are at least two important implications of our results. First, there may be a fitness and fertility valley between homoploid hybrids and allopolyploids with increasing genomic divergence of progenitors (see also Darlington, 1937). A second adaptive valley may be represented by intermediates between typical autopolyploids and allopolyploids, which will suffer the effects of mixed multivalent and bivalent formation in meiosis. Greater support for these hypotheses requires further experimental study.

Acknowledgements

We thank D. Albach, R. Bateman, E. Conti, A. Davis, Y. Guo, G. Mansion, A. Mast, E. E. Schilling and R. E. Timme for providing aligned DNA sequence matrices, M. Winkler for comments on appropriate statistical analysis, I. Leitch and A. Leitch for insightful discussions, and J. Wendel and one anonymous reviewer for useful criticisms. This research was funded by an Erwin Schrödinger fellowship (Austrian Science Fund, FWF; project J26406-B03) and an Intra-European Marie Curie fellowship (*DactGene*, MEIF-CT-2007-040494) to O.P.

References

Allan GJ, Clark C, Rieseberg LH. 1997. Distribution of parental DNA markers in *Encelia virginensis* (Asteraceae: Heliantheae), a diploid species of putative hybrid origin. *Plant Systematics and Evolution* **205**: 205–221.

Allen GA. 2001. Hybrid speciation in *Erythronium* (Liliaceae): a new allotetraploid species from Washington State. *Systematic Botany* **26**: 263–272.

Allen GA, Soltis DE, Soltis PS. 2003. Phylogeny and biogeography of *Erythronium* (Liliaceae) inferred from chloroplast *matK* and nuclear rDNA ITS sequences. *Systematic Botany* **28**: 512–523.

Andersson E. 1949. *Introgressive hybridization*. New York, NY, USA: Wiley.

Arnold M. 1997. *Natural hybridization and evolution*. Oxford, UK: Oxford University Press.

Barton NH. 2001. The role of hybridization in evolution. *Molecular Ecology* **10**: 551–568.

Bomblies K, Weigel D. 2007. Hybrid necrosis: autoimmunity as a common barrier to gene flow in plants. *Nature Review in Genetics* **8**: 382–393.

Bottini MCJ, De Bustos A, Sanso AM, Jouve N, Poggior L. 2007. Relationships in Patagonian species of *Berberis* (Berberidaceae) based on the characterization of rDNA internal transcribed spacer sequences. *Botanical Journal of the Linnean Society* **153**: 321–328.

Brochmann C, Borgen L, Stabbetorp OE. 2000. Multiple diploid hybrid speciation of the Canary Island endemic *Argyranthemum sundingii* (Asteraceae). *Plant Systematics and Evolution* **220**: 77–92

Buerkle CA, Rieseberg LH. 2008. The rate of genome stabilization in homoploid hybrid species. *Evolution* **62**: 266–275.

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.

Chapman MA, Burke JM. 2007. Genetic divergence and hybrid speciation. *Evolution* **61**: 1773–1780.

Chase CD. 2007. Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. *Trends in Genetics* **23**: 81–90.

Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, Joseph J, Savolainen V, Parokonny AS. 2003. Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). *Annals of Botany* **92**: 107–127.

Chat J, Jáuregui B, Petit RJ, Nadot S. 2004. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. *American Journal of Botany* **91**: 736–747

Chen ZJ. 2007. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review in Plant Biology* **58**: 377–406.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the CLUSTAL series of programs. *Nucleic Acids Research* **31**: 3497–3500.

Clarkson JJ, Knapp S, Garcia VF, Olmstead RG, Leitch AR, Chase MW. 2004. Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. *Molecular Phylogenetics and Evolution* **33**: 75–90.

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 Institution of Washington Publications* **564**: 1–174.

Comai L. 2005. The advantages and disadvantages of being polyploid. *Nature Review in Genetics* **6**: 836–846.

Cox TS, Murphy JP. 1990. The effect of parental divergence on F_2 heterosis in winter wheat crosses. *Theoretical Applied Genetics* **79**: 241–250.

Coyne JA, Orr HA. 2004. Polyploidy and hybrid speciation. In: Coyne JA, Orr HA, eds. *Speciation*. Sunderland, MA, USA: Sinauer Associates, 321–351.

Cozzolino S, D'Emerico S, Widmer A. 2004. Evidence for reproductive isolate selection in Mediterranean orchids: karyotype differences compensate for the lack of pollinator specificity. *Proceedings of the Royal Society of London Series B. Biological Sciences* **271**: S259–S262.

Cronn RC, Wendel JF. 2004. Cryptic trysts, genomic mergers and plant speciation. *New Phytologist* **161**: 133–142.

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A *et al.* 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Research* **16**: 738–749.

Darlington CD. 1937. *Recent advances in cytology*, 2nd edn. Philadelphia, PA, USA: Blakiston.

De Bodt S, Maere S, van de Peer Y. 2005. Genome duplication and the origin of angiosperms. *Trends in Ecology and Evolution* **20**: 591–597.

Edmands S. 2002. Does parental divergence predict reproductive compatibility? *Trends in Ecology and Evolution* **17**: 520–527.

Ellstrand NC, Whitkus R, Rieseberg LH. 1996. Distribution of spontaneous plant hybrids. *Proceedings of the National Academy of Sciences, USA* **93**: 5090–5093.

Fehlberg SD, Ranker TA. 2007. Phylogeny and biogeography of *Encelia* (Asteraceae) in the Sonoran and Peninsular deserts based on multiple DNA sequences. *Systematic Botany* **32**: 692–699.

Fjellheim S, Jørgensen MH, Kjos M, Borgen L. 2009. A molecular study of hybridization and homoploid hybrid speciation in *Argyranthemum* (Asteraceae) on Tenerife, the Canary Islands. *Botanical Journal of the Linnean Society* **159**: 19–31.

Flagel L, Udall J, Nettleton D, Wendel J. 2008. Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. *BMC Biology* **6**: 16 doi: 10.1186/1741-7007-6-16.

Ford VS, Gottlieb LD. 2002. Single mutations silence *PgiC2* genes in two very recent allotetraploid species of *Clarkia*. *Evolution* **56**: 699–707.

Ge S, Sang T, Lu B-R, Hong D-Y. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences, USA* **96**: 14400–14405.

Grant V. 1981. *Plant speciation*, 2nd edn. New York, NY, USA: Columbia University Press.

Gross BL, Rieseberg LH. 2005. The ecological genetics of homoploid hybrid speciation. *Journal of Heredity* **96**: 241–252.

Guggisberg A, Bretagnolle F, Mansion G. 2006. Allopolyploid origin of the Mediterranean endemic, *Centaurium bianoris* (Gentianaceae), inferred by molecular markers. *Systematic Botany* **31**: 368–379.

Guo Y-P, Ehrendorfer F, Samuel R. 2004. Phylogeny and systematics of *Achillea* (Asteraceae–Anthemideae) inferred from nrITS and plastid *trnL-F* DNA sequences. *Taxon* **53**: 657–672.

Guo Y-P, Saukel J, Mittermayr R, Ehrendorfer F. 2005. Amplified fragment length polymorphism (AFLP) analyses demonstrate genetic divergence, hybridization, and multiple polyploidization in the evolution of *Achillea* (Asteraceae-Anthemideae). *New Phytologist* **166**: 273–290.

Guo Y-P, Vogl C, van Loo M, Ehrendorfer F. 2006. Hybrid origin and differentiation of two tetraploid *Achillea* species in East Asia: molecular, morphological and ecogeographical evidence. *Molecular Ecology* **15**: 133–144.

Hedrén M. 2001. Systematics of the *Dactylorhiza euxina/incarnata/maculata* polyploidy complex (Orchidaceae) in Turkey: evidence from allozyme data. *Plant Systematics and Evolution* **229**: 23–44.

Hegarty MJ, Barker GL, Wilson ID, Abbott RJ, Edwards KJ, Hiscock SJ. 2006. Transcriptome shock after interspecific hybridization in *Senecio* is ameliorated by genome duplication. *Current Biology* **16**: 1652–1659.

Howarth DG, Baum DA. 2005. Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (Goodeniaceae) in the Hawaiian islands. *Evolution* **59**: 948–961.

Howarth DG, Gustafsson MFG, Baum DA, Motley TJ. 2003. Phylogenetics of the genus *Scaevola* (Goodeniaceae): implications for dispersal patterns across the pacific basin and colonization of the Hawaiian islands. *American Journal of Botany* **90**: 915–923.

Hughes CE, Govindarajulu R, Robertson A, Filer DL, Harris SA, Bailey CD. 2007. Serendipitous backyard hybridization and the origin of crops. *Proceedings of the National Academy of Sciences, USA* **104**: 14389–14394.

Husband BC. 2000. Constraints on polyploid evolution: a test of the minority cytotype exclusion principle. *Proceedings of the Royal Society of London Series B. Biological Sciences* **267**: 217–223.

Jung E, Tate PL, Horn R, Kochert G, Moore K, Abbott AG. 2003. The phylogenetic relationship of possible progenitors of the cultivated peanut. *Journal of Heredity* **94**: 334–340.

Kim Y-D, Kim S-H, Landrum LR. 2004. Taxonomic and phytogeographic implications from ITS phylogeny in *Berberis* (Berberidaceae). *Journal of Plant Research* **117**: 175–182.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**: 111–120.

Koch M, Al-Shehbaz I-A. 2002. Molecular data indicate complex intra- and intercontinental differentiation of American *Draba* (Brassicaceae). *Annals of the Missouri Botanical Garden* **89**: 88–109.

Kuzoff RK, Soltis DE, Hufford L, Soltis PS. 1999. Phylogenetic relationships within *Lithophragma* (Saxifragaceae): hybridization, allopolyploidy, and ovary diversification. *Systematic Botany* **24**: 598–615.

Lee J, Baldwin BG, Gottlieb LD. 2002. Phylogeny of *Stephanomeria* and related genera (Compositae–Lactuceae) based on analysis of 18S-26S nuclear rDNA ITS and ETS sequences. *American Journal of Botany* **89**: 160–168.

Leitch AR, Leitch IJ. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* **320**: 481–483.

Levin RA, Wagner WL, Hoch PC, Hahn WJ, Rodriguez A, Baum DA, Katinas L, Zimmer EA, Sytsma KJ. 2004. Paraphyly in tribe Onagreae: insights into phylogenetic relationships of Onagraceae based on nuclear and chloroplast sequence data. *Systematic Botany* **29**: 147–164.

Li J, Huang H, Sang T. 2002. Molecular phylogeny and infrageneric classification of *Actinidia* (Actinidiaceae). *Systematic Botany* **27**: 408–415.

Lim KY, Matyasek R, Kovarik A, Leitch AR. 2004. Genome evolution in allotetraploid *Nicotiana*. *Biological Journal of the Linnean Society* **82**: 599–606.

Lippman ZB, Zamir D. 2007. Heterosis: revisiting the magic. *Trends in Genetics* **23**: 60–66.

Liu Q, Brubaker CL, Green AG, Marshall DR, Sharp PJ, Singh SP. 2001. Evolution of the *FAD2-1* fatty acid desaturase 5′ UTR intron and the molecular systematics of *Gossypium* (Malvaceae). *American Journal of Botany* **88**: 92–102.

Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society of London B* **363**: 3009–3021.

Maki M, Murata J. 2001. Allozyme analysis of the hybrid origin of *Arisaema ehimense* (Araceae). *Heredity* **86**: 87–93.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* **20**: 229–237.

Mallet J. 2007. Hybrid speciation. *Nature* **446**: 279–283.

Mansion G, Zeltner L, Bretagnolle F. 2005. Phylogenetic patterns and polyploid evolution within the Mediterranean genus *Centaurium* (Gentianaceae – Chironieae). *Taxon* **54**: 931–950.

Maurin O, Davis AP, Chester M, Mvungi EF, Jaufeerally-Fakim Y, Fay MF. 2007. Towards a phylogeny for *Coffea* (Rubiaceae): identifying well-supported lineages based on nuclear and plastid DNA sequences. *Annals of Botany* **100**: 1565–1583.

Mavrodiev EV, Tancig M, Sherwood AM, Gitzendanner MA, Rocca J, Soltis PS, Soltis DE. 2005. Phylogeny of *Tragopogon* L. (Asteraceae) based on internal and external transcribed spacer sequence data. *International Journal of Plant Sciences* **166**: 117–133.

Meyers LA, Levin DA. 2006. On the abundance of polyploidy in flowering plants. *Evolution* **60**: 1198–11206.

Moll RH, Lonnquist JH, Wlez Fortuno J, Johnson EC. 1965. The relationship of heterosis and genetic divergence in maize. *Genetics* **52**: 139–144.

Orr HA. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**: 1805–1813.

Otto SP. 2007. The evolutionary consequences of polyploidy. *Cell* **131**: 452–462.

Otto SP, Whitton J. 2000. Polyploid incidence and evolution. *Annual Review in Genetics* **34**: 401–437.

Pan J, Zhang D, Sang T. 2007. Molecular phylogenetic evidence for the origin of a diploid hybrid of *Paeonia* (Paeoniaceae). *American Journal of Botany* **94**: 400–408.

Paun O, Fay MF, Soltis DE, Chase MW. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon* **56**: 649–656.

Paun O, Stuessy TF, Hörandl E. 2006. The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist* **171**: 223–236.

Pillon Y, Fay MF, Hedrén M, Bateman RM, Devey DS, Shipunov AB, van

der Bank M, Chase MW. 2007. Evolution and temporal diversification of western European polyploid species complexes in *Dactylorhiza* (Orchidaceae). *Taxon* **56**: 1185–1208.

Pillon Y, Fay MF, Shipunov AB, Chase MW. 2006. Species diversity versus phylogenetic diversity: a practical study in the taxonomically difficult genus *Dactylorhiza* (Orchidaceae). *Biological Conservation* **129**: 4–13.

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

Ramsey J, Schemske DW. 2002. Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics* **33**: 589–639.

Renner SS, Zhang L-B, Murata J. 2004. A chloroplast phylogeny of *Arisaema* (Araceae) illustrates tertiary floristic links between Asia, North America, and East Africa. *American Journal of Botany* **91**: 881–888.

Rieseberg LH. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**: 359–389.

Rieseberg LH. 2000. Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution* **54**: 859–865.

Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* **16**: 351–358.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**: 1211–1216.

Rieseberg LH, Sinervo B, Linder CR, Ungerer M, Arias DM. 1996. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* **272**: 741–745.

Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**: 1120–1136.

Sang T, Pan J, Zhang D, Ferguson D, Wang C, Paun K-Y, Hong D-Y. 2004. Origins of polyploids: an example from peonies (*Paeonia*) and a model for angiosperms. *Biological Journal of the Linnean Society* **82**: 561–571.

Schilling EE, Linder CR, Noyes RD, Rieseberg LH. 1998. Phylogenetic relationships in *Helianthus* (Asteraceae) based on nuclear ribosomal DNA internal transcribed spacer region sequence data. *Systematic Botany* **23**: 177–187.

Seehausen O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* **19**: 198–207.

Seelanan T, Brubaker CL, Stewart J, Craven LA, Wendel JF. 1999. Molecular systematics of Australian *Gossypium* section *Grandicalyx* (Malvaceae). *Systematic Botany* **24**: 183–208.

Seelanan T, Schnabel A, Wendel JF. 1997. Congruence and consensus in the cotton tribe (Malvaceae). *Systematic Botany* **22**: 259–290.

Senchina DS, Alvarez I, Cronn RC, Liu B, Rong J, Noyes RD, Paterson AH, Wing RA, Wilkins TA, Wendel JF. 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Molecular Biology and Evolution* **20**: 633–643.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, DePamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* **96**: 336–348.

Song K, Lu P, Tang K, Osborn CT. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences, USA* **92**: 7719–7723.

Stebbins GL. 1950. *Variation and evolution in plants*. New York , NY, USA: Columbia University Press.

Sun K, Chen X, Ma R, Li C, Wang Q, Ge S. 2002. Molecular phylogenetics of *Hippophae* L. (Elaeagnaceae) based on the internal transcribed spacer (ITS) sequences of nrDNA. *Plant Systematics and Evolution* **235**: 121–134.

Swofford DL. 2003. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*, version 4.0b10. Sunderland, MA, USA: Sinauer Associates.

Thompson JD, Lumaret R. 1992. The evolutionary dynamics of polyploid plants: origins, establishment and persistence. *Trends in Ecology and Evolution* **7**: 302–307.

Timme RE, Simpson BB, Linder CR. 2007. High-resolution phylogeny for *Helianthus* (Asteraceae) using the 18S-26S ribosomal DNA external transcribed spacer. *American Journal of Botany* **94**: 1837–1852.

van Tuyl JM, De Vries JN, Bino RJ, Kwakkenbos TAM. 1989. Identification of 2n-pollen producing interspecific hybrids of *Lilium* using flow-cytometry. *Cytologia* **54**: 737–745.

Ungerer MC, Baird SJE, Pan J, Rieseberg LH. 1998. Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Sciences, USA* **95**: 11757–11762.

Vander Stappen J, Gama Lopez S, Davila P, Volckaert G. 2002. Molecular evidence for the hybrid origin of a new endemic species of *Stylosanthes* Sw. (Fabaceae) from the Mexican Yucatán Peninsula. *Botanical Journal of the Linnean Society* **140**: 1–13.

Wang A, Schluetz F, Liu J. 2008. Molecular evidence for double maternal origins of the diploid hybrid *Hippophae goniocarpa* (Elaeagnaceae). *Botanical Journal of the Linnean Society* **156**: 111–118.

Weiss-Schneeweiss H, Schneeweiss GM, Stuessy TF, Mabuchi T, Park J-M, Jang C-G, Sun B-Y. 2007. Chromosomal stasis in diploids contrasts with genome restructuring in auto- and allopolyploid taxa of *Hepatica* (Ranunculaceae). *New Phytologist* **174**: 669–682.

Wendel JF, Cronn RC. 2003. Polyploidy and the evolutionary history of cotton. *Advances in Agronomy* **78**: 139–186.

White MJD. 1978. *Modes of speciation*. San Francisco, CA, USA: WH Freeman & Co.

Widmer A, Baltisberger M. 1999. Molecular evidence for allopolyploid speciation and a single origin of the narrow endemic *Draba ladina* (Brassicaceae). *American Journal of Botany* **86**: 1282–1289.

Widmer A, Lexer C, Cozzolino S. 2009. Evolution of reproductive isolation in plants. *Heredity* **102**: 31–38.

Winge Ø. 1917. The chromosomes. Their numbers and general importance. *Comptes-rendus des travaux de la laboratoire Carlsberg* **13**: 131–275.

Wolfe A, Randle CP. 2001. Relationships within and among species of the holoparasitic genus *Hyobanche* (Orobanchaceae) inferred from ISSR banding patterns and nucleotide sequences. *Systematic Botany* **26**: 120–130.

Wolfe AD, Randle CP, Datwyler SL, Morawetz JL, Arguedas N, Diaz J. 2006. Phylogeny, taxonomic affinities, and biogeography of *Penstemon* (Plantaginaceae) based on ITS and cpDNA sequence data. *American Journal of Botany* **93**: 1699–1713.

Wolfe AD, Xiang Q-Y, Kephart SK. 1998. Diploid hybrid speciation in *Penstemon* (Scrophulariaceae). *Proceedings of the National Academy of Sciences, USA* **95**: 5112–5115.

Wood TE, Rieseberg LH. 2005. The frequency of polyploid speciation in land plants: a phylogenetic approach. Abstract 476 In: *Botany 2005: Learning From Plants*, Scientific meeting, August 13–17, 2005. Austin, Texas. Available at http://www.2005.botanyconference.org/engine/search/index.php?func=detail&aid=476.

Wu C-I. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**: 851–865.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Details on the data included in the analyses and results for each individual hybrid case

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.