On the Origins of Species: Does Evolution Repeat Itself in Polyploid Populations of Independent Origin?

D.E. SOLTIS,^{1,2} R.J.A. BUGGS,¹ W.B. BARBAZUK,^{1,2} P.S. SCHNABLE,³ AND P.S. SOLTIS^{2,4}

¹Department of Biology, University of Florida, Gainesville, Florida 32611; ²Genetics Institute, University of Florida, Gainesville, Florida 32610; ³Center for Plant Genomics, Iowa State University, Ames, Iowa 50011; ⁴Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611

Correspondence: dsoltis@botany.ufl.edu

Multiple origins of the same polyploid species pose the question: Does evolution repeat itself in these independently formed lineages? *Tragopogon* is a unique evolutionary model for the study of recent and recurrent allopolyploidy. The allotetraploids *T. mirus* (*T. dubius* × *T. porrifolius*) and *T. miscellus* (*T. dubius* × *T. pratensis*) formed repeatedly following the introduction of three diploids to the United States. Concerted evolution has consistently occurred in the same direction (resulting in loss of *T. dubius* rDNA copies). Both allotetraploids exhibit homeolog loss, with the same genes consistently showing loss, and homeologs of *T. dubius* preferentially lost in both allotetraploids. We have also documented repeated patterns of tissue-specific silencing in multiple populations of *T. miscellus*. Hence, some aspects of genome evolution may be “hardwired,” although the general pattern of loss is stochastic within any given population. On the basis of the study of F_1 hybrids and synthetics, duplicate gene loss and silencing do not occur immediately following hybridization or polyploidization, but gradually and haphazardly. Genomic approaches permit analysis of hundreds of loci to assess the frequency of homeolog loss and changes in gene expression. This methodology is particularly promising for groups such as *Tragopogon* for which limited genetic and genomic resources are available.

In *On the origin of species*, Darwin (1859) presented a mechanism by which evolution could occur via natural selection. Our understanding of evolution and speciation has obviously improved dramatically in the past 150 years. Of the numerous new insights, perhaps one of the more surprising discoveries is the relatively recent finding that the same species can actually form multiple times. Specifically, in the case of polyploid organisms, the same polyploid species may form not once, but repeatedly in multiple locations where the diploid parents come into contact and hybridize. In fact, the use of molecular techniques has shown that most polyploid species have probably formed more than once—that multiple origins of the same polyploid species is the norm, not the exception, in polyploidy organisms (see, e.g., Soltis and Soltis 1993, 1999). Given the importance and prevalence of polyploidy in some plant groups, particularly ferns and angiosperms, the recurrent formation of the same polyploid species becomes a major evolutionary factor. Numerous examples of recurrent polyploidization have now been proposed for polyploid animals and plants (for review, see Soltis and Soltis 1993, 1999). But, the “repeatability” of polyploid speciation may be best seen on a broad geographic scale in the arctic, where diploid progenitor species come into contact over and over again on a circumpolar scale, hybridizing and subsequently generating the same polyploids again and again (Brochmann et al. 2004; Grundt et al. 2006).

Recurrent formation of the same polyploid species poses intriguing evolutionary questions, a major one being: Does evolution repeat itself in polyploid lineages of independent

origin? On a grand scale, evolutionary biologists have pondered this question. Gould (1994), for example, suggested that if the tape of evolution of life on Earth could be replayed, it would play out differently each time—“history involves too much chaos,” and too many chance events are involved for the evolutionary process to be repetitive. However, is this true on a finer scale? What about at the level of species and during shorter time frames? That is, are some genetic features of the polyploidization process “hardwired” so that the same genomic/genetic changes will recur in polyploid populations of independent formation? Recent research has suggested that at deep levels across broad clades of life, preservation of duplicated gene copies following genome duplication is far from random, with specific functional categories preferentially retained and reduplicated in subsequent polyploidizations (Seoighe and Gehring 2004; Chapman et al. 2006). Independent whole-genome duplications in the ancestors of *Arabidopsis*, *Oryza* (rice), *Saccharomyces* (yeast), and *Tetraodon* (pufferfish) appear to have been followed by convergent fates of many gene families (Paterson et al. 2006). Collectively, these observations indicate that on a broad scale, there *may* exist certain “principles” that govern the fates of gene and genome duplications. On the basis of these data, perhaps the tape of evolution would replay in the very same or similar way each time at the level of independently formed polyploid lines. Conversely, perhaps stochasticity has a major role, resulting in little repeatability or predictability across populations of independent formation. As one more alternative, perhaps the end result is some place between these two extremes.

With this brief introduction, it is apparent that polyploid plants of independent origin provide an unusual opportunity to address a fundamental question: Does evolution repeat itself? A particularly useful plant system for studying the early phases of polyploidization and addressing this fundamental question is provided by members of the genus *Tragopogon* (goatsbeard) (Asteraceae; sunflower family). As reviewed below, *Tragopogon* is a unique evolutionary model for the study of recent and recurrent polyploidy, providing a superb system for investigating the repeatability of the evolutionary process.

THE *TRAGOPOGON* SYSTEM: A BRIEF HISTORY

Although polyploidy has long been recognized as prevalent in plants (see, e.g., Müntzing 1936; Darlington 1937; Clausen et al. 1945; Stebbins 1947, 1950; Löve and Löve 1949; Grant 1981), genomic data have now revealed that it is an even more significant force than previously proposed (see, e.g., Blanc et al. 2000, 2003; Paterson et al. 2000; Vision et al. 2000; Simillion et al. 2002; Bowers et al. 2003; Blanc and Wolfe 2004; Schlueter et al. 2004; Cui et al. 2006). The question being asked is no longer “what proportion of angiosperms are polyploid?” but “how many episodes of polyploidy characterize any given lineage?”

Despite enormous progress in our understanding of many aspects of polyploidy (see, e.g., Wendel 2000; Tate et al. 2004; Wendel and Doyle 2004; Doyle et al. 2008), the early stages of polyploid evolution remain poorly understood, particularly in natural populations. Although polyploidy is ubiquitous in plants, only a few polyploid species are known to have arisen recently, i.e., within just the past 150–200 years: *Cardamine schulzii* (Urbanska et al. 1997), *Spartina anglica* (Huskings 1931; for review, see Ainouche et al. 2004), *Senecio cambrensis* (Rosser 1955) and *Senecio eboracensis* (Abbott and Lowe 2004), and two species of *Tragopogon*, *T. mirus* and *T. miscellus* (Ownbey 1950; for review, see Soltis et al. 2004). All of these new polyploids provide the opportunity to examine the early stages of polyploidization; all but *C. schulzii* have garnered considerable recent attention (see, e.g., Ainouche et al. 2004; Hegarty et al. 2005, 2006). Of these taxa, *Tragopogon* provides the best system for the study of recent and recurring polyploidy in natural populations.

Tragopogon consists of ~150 species native to Eurasia, most of which are diploid ($2n = 12$). Three of these diploids, *T. dubius*, *T. pratensis*, and *T. porrifolius*, were introduced into North America as it was settled by Europeans. *T. dubius* and *T. pratensis* were likely introduced accidentally, but *T. porrifolius* (salsify) has an edible root, was planted, and escaped. In the Palouse region of eastern Washington and adjacent Idaho, the three diploids came into close contact, which rarely happens in Europe where they are ecologically and in part geographically isolated. Hybridization occurred and two new allopolyploid species with $2n = 24$ ultimately formed: *T. mirus* (*T. porrifolius* × *T. dubius*) and *T. miscellus* (*T. dubius* × *T. pratensis*) (Fig. 1). The new allotetraploids have not formed in Europe, but they are native to the New

Figure 1. Summary of parentage of tetraploid *Tragopogon* species comparing what we have produced synthetically (“man”) and what has occurred in nature (“wild”). The diploid parents (with $2n = 12$) are at the corners of the triangle; polyploids ($2n = 24$) are in between the corners. Synthetic polyploids are to the outside of the triangle; those polyploids forming naturally are to the inside of the triangle. In nature, *T. miscellus* has formed reciprocally, and *T. mirus* has formed only in one direction (with *T. porrifolius* as the maternal parent). However, we have made reciprocal synthetic lines of both and have also made reciprocal polyploids of *T. pratensis* × *T. dubius* (“*T. floridana*”); this polyploid has not formed in nature. Note that populations of *T. miscellus* of reciprocal origin differ in morphology. Those with *T. pratensis* as the maternal parent have short ligules, and those with *T. dubius* as the maternal parent have long ligules. (Photographs contributed by A. Doust and V. Symonds; plate courtesy of J. Tate.)

World, although their parents are aliens. Hybrids formed (and still form) between *T. pratensis* and *T. porrifolius*, but a polyploid has never been detected.

Ownbey first collected *T. mirus* and *T. miscellus* in 1949 and named these new species in 1950. Given his expertise as a systematist, it is likely that he discovered these new species not long after they first formed. Herbarium records indicate that the three diploid parents did not all occur in the Palouse region before 1928. Hence, the polyploids are probably not more than 80 years old. Given that these plants are biennials, the timescale involved since the formation of these two new species is fewer than 40 generations.

MULTIPLE ORIGINS AND A NORTH AMERICAN SUCCESS STORY

The parentage of both polyploids (*T. mirus* and *T. miscellus*) has now been confirmed using multiple approaches (for review, see Soltis et al. 2004). Ownbey (1950) referred to the small populations of the new polyploids as “small and precarious” but indicated that they appeared to be competing successfully with their parents and that it would be important to “follow the two polyploids over time.”

Novak et al. (1991) conducted a survey to determine how common the two polyploids had become since Ownbey's discovery. One or both polyploids were found in most towns of the Palouse with populations ranging from small (<100 individuals) to many thousands of individuals. *T. miscellus* had become one of the most common weeds in and around Spokane, Washington, as well as in parts of Moscow, Idaho, and Spangle, Washington. Populations of *T. mirus* and *T. miscellus* often formed dense stands and were, in fact, displacing their parents, particularly *T. pratensis* and *T. porrifolius*.

Genetic markers revealed how the two *Tragopogon* polyploids had so quickly occupied towns across eastern Washington and Idaho. Both polyploids had formed repeatedly. A diverse array of approaches (for review, see Soltis et al. 2004) indicated as many as 21 distinct lineages of separate origin of *T. miscellus* and perhaps 11 lineages of *T. mirus* (Soltis and Soltis 2000; Soltis et al. 2004). Ownbey had in fact suggested multiple formations based on morphology and cytology (Ownbey and McCollum 1953, 1954). Our recent use of microsatellite markers indicates even more extensive multiple formations than initially thought. In several cases, distinct populations of *T. mirus* and *T. miscellus* from the same small town, separated by less than several kilometers, have formed independently (VV Symonds et al., in prep.).

On a larger geographic scale, both polyploids have also formed in Flagstaff, Arizona; *T. miscellus* has formed in Gardiner, Montana (now apparently extinct; DE Soltis, unpubl.) and Sheridan, Wyoming (for review, see Soltis et al. 2004).

SYNTHETIC LINES: MAN VS. WILD

Adding to the utility of the *Tragopogon* system as an evolutionary model is the recent production of multiple synthetic lines of both *T. mirus* and *T. miscellus* (Tate et al. 2009). These lines provide the added opportunity of examining multiple synthetic lines of both polyploids, following polyploidization from its inception. Comparison of the synthetics to natural populations of separate origin adds another important dimension to the "does evolution repeat itself?" question.

In nature, all formations of *T. mirus* have *T. porrifolius* as the maternal parent and *T. dubius* as the paternal parent, but we have synthesized *T. mirus* reciprocally (both combinations). *T. miscellus* has formed reciprocally in nature (resulting in a dramatic change in floral head morphology; Fig. 1), and Tate et al. (2009) also created synthetic lines with *T. dubius* as both the maternal and the paternal parent. Tate et al. (2009) also produced what has not formed in nature, polyploids between *T. pratensis* and *T. porrifolius* (Fig. 1). All of these synthetic lines are now in the second generation and offer the unique opportunity for comparative genetic/genomic study of repeated formations of both natural and synthetic polyploids.

So, with this background to the *Tragopogon* polyploid evolutionary model system, does genome evolution repeat itself in natural populations of separate origin of *T. miscellus* and *T. mirus* and also in multiple synthetic lines

of the two species? We examine the data now available in the sections below.

rDNA LOCI AND CONCERTED EVOLUTION

Concerted evolution, which results in the homogenization of gene sequences to one type, is a common feature of ribosomal RNA (rRNA) genes (see, e.g., Zimmer et al. 1980). In F_1 hybrids and in some allopolyploids, the rDNA types of both parent diploids are present. Both parental arrays may be present in an allopolyploid, as in some allopolyploids of *Glycine* (Doyle and Beachy 1985; Rauscher et al. 2004), *Triticum* (Appels and Dvorak 1982), *Paenonia* (Sang et al. 1995), *Krigia* (Kim and Jansen 1994), *Brassica napus* (Bennett and Smith 1991), and *Arabidopsis suecica* (O'Kane et al. 1996). But in some allopolyploids, only one parental type is present, with homogenization to one parental type having occurred as reported in species of *Gossypium* L. (Wendel et al. 1995), *Nicotiana* L. (Volkov et al. 1999; Lim et al. 2000; Kovarik et al. 2004), *Cardamine* L. (Franzke and Mummenhoff 1999), *Triticum* L. (Flavell and O'Dell 1976), *Glycine* (Rauscher et al. 2004), and *Senecio* L. (Abbott and Lowe 2004).

In *T. mirus* and *T. miscellus*, concerted evolution is ongoing, but incomplete; i.e., we have essentially caught it in the act (Kovarik et al. 2005). F_1 hybrids have equal contributions of the diploid parents, as do the new synthetic polyploids and the earliest natural populations of *T. mirus* and *T. miscellus* (based on DNA from herbarium specimens). But in all modern day natural populations examined representing distinct origins, the rDNA type of *T. dubius* is consistently in very low abundance, with either the *T. pratensis* rDNA type (in *T. miscellus*) or *T. porrifolius* rDNA type (in *T. mirus*) in much greater abundance. Thus, concerted evolution has consistently occurred in these new polyploid lines of separate origin, and it has repeatedly operated "against" *T. dubius*, homogenizing those copies in the direction of the other parent. This is readily seen on Southern blots (Fig. 2). Thus, in the case of the rDNA cistron, molecular evolution of rDNA *does appear* to have repeated itself in *Tragopogon*. Surprisingly, despite being the least abundant in terms of rDNA gene copy number, *T. dubius* is by far the most abundant transcript in natural polyploidy populations (Matyasek et al. 2007).

HOMEOLOG LOSS AND GENE SILENCING

Tragopogon is an evolutionary model, but not a genetic model organism; hence, genetic resources are not available. As a result, genetic and genomic changes in the newly formed tetraploids have been so far examined using a "one gene at a time" approach (Fig. 3). This is slow tedious work that has required ~5 years to examine ~30 genes in multiple populations of both *T. mirus* and *T. miscellus* (Tate et al. 2006; Buggs et al. 2009; J Koh et al., in prep.). The genes analyzed to date were chosen based on several different approaches. Amplified-fragment-length polymorphism (AFLP)-cDNA display was initially used to screen plants of *T. miscellus* and *T. mirus* and parental diploids to look

Figure 2. Southern blot hybridization of the ITS region in *Tragopogon* polyploids and parents. (Top) Southern blot hybridization of genomic DNAs showing variability in the ITS1 content among populations of separate origin of *T. mirus*. The DNAs were digested with restriction enzymes yielding diagnostic fragments for the diploid parents: *T. porrifolius* (po) and *T. dubius* (do). (Bottom) Southern blot hybridization of genomic DNAs showing variability in the ITS1 content among populations of separate origin of *T. miscellus*. Same enzymes and digestion conditions as in top panel.

for promising candidate genes, i.e., fragments that did not show additivity in the allopolyploids as would be expected (Tate et al. 2006; J Koh et al., in prep.). However, additional genes were chosen for survey because they were orthologous to genes that were singletons in *Arabidopsis* (Buggs et al. 2009; J Koh et al., in prep.); the fate of such genes seemed to be of particular interest in new polyploids. Are these “singleton” genes rapidly returned to single-copy status in new *Tragopogon* polyploids?

The results of these gene surveys are presented in detail elsewhere, and we only summarize the major features of those studies here (Tate et al. 2006; Buggs et al. 2009; J

Koh et al., in prep.). Some important generalizations have emerged from these studies across polyploid populations of separate origin of both *T. mirus* and *T. miscellus*. Most of the changes observed in populations of both young polyploids are homeolog-loss events; these far outnumber gene-silencing events in these plants. Furthermore, most of the homeolog losses in both polyploids have involved *T. dubius*, the diploid parent shared by both *T. mirus* and *T. miscellus*.

It is also noteworthy that the same suite of genes consistently shows additivity of the parental gene copies (no loss or silencing) in polyploid populations of separate

Figure 3. Example of use of genomic CAPS (cleaved amplified polymorphic sequence) markers to examine homeolog loss (see Tate et al. 2006; Buggs et al. 2009). The stained 4% metaphor agarose gel shows homeolog loss in two *T. miscellus* individuals (lanes 13 and 16) from the Oakesdale population for gene D1 (an LRR [leucine-rich repeat] protein kinase). The DNA fragments visualized are products of a restriction enzyme digest on polymerase chain reaction (PCR)-amplified fragments of the gene in individual plants. The first eight lanes of the gel show *T. pratensis* individuals (Soltis and Soltis collection number 2672, individuals 1, 2, 3, 6, 7, 8, 9, and 10, respectively); lanes 9 to 16 show *T. miscellus* individuals (Soltis and Soltis collection number 2671, individuals 2, 3, 4, 5, 7, 8, 10, and 11, respectively); lanes 17 to 23 show *T. dubius* individuals (Soltis and Soltis collection number 2670, individuals 1–7, respectively); lane 24 shows a negative PCR control; and lane 25 shows HyperLadder IV (Biolone, Taunton, Massachusetts) markers from 100–500 bp. (This example modified from Buggs et al. 2009.)

origin, whereas some of the genes analyzed consistently show some evidence of loss across at least some of the populations surveyed. Significantly, at deep levels across eukaryotes, Paterson et al. (2006) showed that genes with some PFam domains may be consistently returned to singleton status following genome-wide duplication, whereas other genes consistently are retained as duplicates across highly divergent lineages. Thus, our observations for *Tragopogon* polyploids are also in agreement with the hypothesis that some underlying “principles” to polyploidization may exist at the genetic/biochemical level.

But there is an element of randomness operating as well in these young polyploids. Although homeolog loss is present in the polyploid populations, the process is ongoing and appears to be stochastic within individual populations. In no population examined has silencing or loss been complete, i.e., observed in all individuals of a population. Furthermore, these losses and gene-silencing events were not detected in F_1 hybrids or early-generation synthetic lines (S_1). Hence, loss of homeologs and gene silencing are not immediate consequences of hybridization or polyploidization in *Tragopogon*, but they appear to occur somewhat gradually and haphazardly for certain genes following polyploidy.

TISSUE-SPECIFIC SILENCING

Several outcomes exist for genes following duplication: (1) Both members of a duplicate gene pair may retain their original function; (2) one copy of a duplicate gene pair may retain the original function, but the other copy may become lost or silenced (Lynch and Conery 2000; Adams et al. 2003; Adams 2007; Sterck et al. 2007); (3) duplicate genes may partition the original gene function (subfunctionalization), with one copy active, for example, in one tissue and the other copy active in another tissue (Lynch and Conery 2000; Lynch and Force 2000); and (4) one copy may retain the original function while the other develops a new function (neofunctionalization) (Drea et al. 2006; Teshima and Innan 2008).

We recently examined (RJA Buggs et al., in press) tissue-specific silencing in 10 individuals from two reciprocally formed natural populations of *T. miscellus* (Asteraceae). Using cleaved amplified polymorphic sequence analysis of 18 homeologs, we found homeolog silencing in 14 genes in at least one individual and, of these, eight genes showed tissue-specific silencing. Patterns of tissue-specific homeolog silencing varied a great deal among individuals, but there was an occasional repetition of pattern, such as repeated silencing in corolla tissue. Patterns of silencing were not determined by the direction of the parental cross. By comparison with synthetic allopolyploids, F_1 hybrids, and parental diploids, we showed that most cases of homeolog silencing have arisen in early generations after whole-genome duplication. Semiquantitative analysis showed greater variance of expression between individuals and tissues in the natural populations compared to the synthetics and hybrids for seven genes. Some silencing was due to genomic loss of

duplicate genes. Our data suggest that more homeolog expression changes occur in early generations of new polyploids than do the processes of hybridization and whole-genome duplication and thus contribute to duplicated gene evolution.

TRAGOPOGON GOES GENOMIC

As noted, *Tragopogon* is not a genetic model system. Hence, our investigations to date have been limited to a gene-by-gene approach. Recent advances in high-throughput sequencing technology provide a rapid and cost-effective means to generate sequence data. This and other genomic approaches offer the opportunity to accelerate dramatically our ability to survey gene loss and expression changes in nonmodel species such as *Tragopogon*. Essentially, instead of building a wall one brick at a time, with genomic methods, we can pour it or build it all at once.

We recently (RJA Buggs et al., in press) undertook a hybrid next-generation sequencing approach to identifying single-nucleotide polymorphism (SNP) markers for homologous genes in *Tragopogon miscellus*. This general approach (Fig. 4) makes it possible to rapidly build genetic resources for many “nonmodel” plant systems such as *Tragopogon*. We (RJA Buggs et al., in press) generated reference expressed sequence tags (ESTs) from the transcriptome of *T. dubius* using 454 FLX sequencing. We then generated and aligned Illumina reads from the diploids *T. pratensis* and *T. dubius* to this reference. The resulting alignments generated 7782 SNPs within 2885 contigs between *T. dubius* and *T. pratensis* at high stringency. We then examined these SNPs in a pilot transcriptome profile for *T. miscellus* using Illumina sequence reads. Of the 7782 SNPs, 2064 (27%) appeared to show equal homeolog expression in *T. miscellus*, 671 (9%) showed differential expression in *T. miscellus*, and 254 (3%) showed potential homeolog loss in *T. miscellus*. Most of the potential homeolog losses were of the *T. dubius* homeolog (164/254) with a minority the *T. pratensis* homeolog (90/254), in agreement with results from our gene-by-gene approach. Sequenom analyses confirmed that in a sample of 27 of the SNPs showing potential gene loss from the transcriptome profile, 23 (85%) were cases of genomic homeolog loss.

MECHANISM OF GENE LOSS

Using GISH (genomic in situ hybridization) and FISH (fluorescence in situ hybridization), we have detected surprising chromosomal variation in natural populations of *T. mirus* and *T. miscellus* including inversions and intergenomic translocations, as well as fertile plants of both polyploids having three copies of one chromosome, but one copy of another (reciprocal trisomy/monosomy) (Lim et al. 2008). These rearrangements provide one possible mechanism for homeolog loss in these plants. Additional chromosomal studies are needed; it is unknown, for example, if similar chromosomal changes are present in polyploid populations of separate formation.

Figure 4. General flow chart illustrating methods that can be used to rapidly develop genomic resources and markers for the *Tragopogon* system.

COMPARING *TRAGOPOGON* TO WELL-STUDIED SYSTEMS

Genome evolution has been well studied in model polyploid systems (e.g., *Gossypium*, *Triticum*, *Brassica*, and *Arabidopsis*). Although much more work on *Tragopogon* is needed, at this stage, polyploid evolution in *T. mirus* and *T. miscellus* exhibits important similarities and differences to the better-studied genetic models. *Tragopogon* is noteworthy in that initial studies show that homeolog losses far outnumber true changes in duplicate gene expression (Tate et al. 2006; Buggs et al. 2009; T Koh et al., unpubl.). For example, of the initial 23 genes analyzed in *T. miscellus*, 15 showed homeolog loss in one or more plants from nature, and only eight showed true expression changes; results for *T. mirus* are comparable. In contrast, in synthetic wheat (*Triticum*) and synthetic *Arabidopsis thaliana* and *A. suecica* polyploids, expression changes dominate (Kashkush et

al. 2002; Madlung et al. 2005; Wang et al. 2006), although losses do occur in wheat. Genome evolution in *Tragopogon* may be most similar to *Brassica napus*, in which most of the apparent expression changes observed in later generations were due to losses, most likely resulting from genomic rearrangements (Song et al. 1995; Gaeta et al. 2007).

In *Tragopogon*, *T. dubius* appears to be the “loser” diploid genome in both *T. mirus* and *T. miscellus*. Preferential loss of expression, or homeologs, has also been observed in some of the model polyploid systems. For example, across 50 synthetic lines of *B. napus*, genetic changes are equally distributed between the parental diploid genomes (Gaeta et al. 2007). The system is dynamic—some lines become more “*oleracea*-like” and others more “*rapa*-like” in terms of losses and corresponding expression differences. In *A. suecica* allopolyploids, silencing of homeologs from one parent (*A. thaliana*) was observed more frequently than silencing of homeologs

from the other parent, *A. arenosa* (Wang et al. 2006). In synthetic polyploids in *Triticale*, the contribution of the rye genome (*Secale cereale*) is preferentially silenced (Ma et al. 2004; Ma and Gustafson 2006). In the allohexaploid *Senecio cambrensis*, gene expression is more similar to that of the tetraploid parent (*S. vulgaris*) than to the diploid parent *S. squalidus* (Hegarty et al. 2006). In cotton, tissue-specific subfunctionalization was observed; there is no global bias in the expression of transcripts from one parental diploid genome over the other (Adams et al. 2003; Adams and Wendel 2004). In both *T. miscellus* and *T. mirus*, homeologs of one diploid genome (*T. dubius*) are more often lost or not expressed; i.e., *T. dubius* is often the “loser genome” based on the set of genes surveyed to date. It remains to be seen whether these patterns are maintained over the entire genome; i.e., are homeologs of the diploid parent (*T. dubius*) shared by both *T. mirus* and *T. miscellus* (Fig. 1) preferentially lost or silenced?

SUMMARY/FUTURE PROSPECTS

Most of what we know about the genetic and genomic consequences of polyploidy has been derived from the study of crops, synthetic polyploids, and models. To understand how polyploidy shapes genome evolution and gene function in nature, we must extend from a few models and synthetics to naturally occurring polyploids. *Tragopogon* provides the unique opportunity to investigate the genetic and genomic changes that occur across a continuum from F_1 hybrids, synthetic allopolyploids, independently formed natural populations of *T. mirus* and *T. miscellus* that are 60–80 years postformation, to older Eurasian polyploids. With this system, we can assess the relative frequency of gene loss and gene expression changes in natural populations of the two allotetraploid species and in *synthetic* lines of the two allotetraploids and determine whether *evolution repeats itself*.

To date, we have used a gene-by-gene approach to examine gene loss and expression changes. The data so far garnered for ~30 genes indicate that some aspects of genomic evolution may be hardwired—that some features are repeated in polyploid populations of separate origin; that there may be underlying “rules” that govern some genomic interactions in polyploids. But we need to extend from a few genes to a better representation of the genome. We are now using a genomics approach; rather than slowly building a wall of genetic data for *Tragopogon* one brick at a time, we are using a genomics approach to build a wall all at once, querying thousands of genes. The proposed methodology has not been previously applied to “nonmodel” species, and our initial data indicate that this general approach is particularly promising for species such as *Tragopogon* for which limited genetic and genomic resources are available.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grants MCB 034637, DEB 0919254, and DEB 0919348.

REFERENCES

- Abbott RJ, Lowe AJ. 2004. Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biol J Linn Soc* **82**: 467–474.
- Adams KL. 2007. Evolution of duplicate gene expression in polyploid and hybrid plants. *J Hered* **98**: 136–141.
- Adams KL, Wendel JF. 2004. Exploring the genomic mysteries of polyploidy in cotton. *Biol J Linn Soc* **82**: 573–581.
- Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci* **100**: 4649–4654.
- Ainouche ML, Salmon A, Baume A, Yannic G. 2004. Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytol* **161**: 165–172.
- Appels R, Dvorak J. 1982. The wheat ribosomal DNA spacer region: Its structure and variation in populations and among species. *Theor Appl Genet* **63**: 337–348.
- Bennett RI, Smith AG. 1991. Use of a genomic clone for ribosomal RNA from *Brassica oleracea* in RFLP analysis of *Brassica* species. *Plant Mol Biol* **16**: 685–688.
- Blanc G, Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* **16**: 1667–1678.
- Blanc G, Barakat A, Guyot R, Cooke R, Delseny M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12**: 1093–1101.
- Blanc G, Hokamp K, Wolfe KH. 2003. A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res* **13**: 137–144.
- Bowers JE, Chapman BA, Rong JK, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**: 433–438.
- Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen A-C, Elven R. 2004. Polyploidy in arctic plants. *Biol J Linn Soc* **82**: 521–536.
- Buggs J, Chamala S, Wu W, Gao L, May GD, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB. 2009. Characterization of duplicate gene evolution in the recent natural allopolyploid *Tragopogon miscellus* by next-generation sequencing and Sequenom iPLEX genotyping. *Mol Ecol* (in press).
- Buggs J, Doust AN, Tate JA, Koh J, Soltis K, Feltus FA, Paterson AH, Soltis PS, Soltis DE. 2009. Gene loss and silencing in *Tragopogon miscellus* (Asteraceae): Comparison of natural and synthetic allotetraploids. *Heredity* **24**: 1–9.
- Chapman BA, Bowers JE, Feltus FA, Paterson AH. 2006. Buffering of crucial functions by paleologous duplicated genes may contribute cyclicity to angiosperm genome duplication. *Proc Natl Acad Sci* **103**: 2730–2735.
- Clausen J, Keck DD, Hiesey WM. 1945. Experimental studies on the nature of species. II. Plant evolution through amphiploidy and autopolloidy, with examples from the Madiinae. *Carnegie Inst Wash Publ*, no. 564.
- 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 Res* **16**: 738–749.
- Darlington CD. 1937. *Recent advances in cytology*. Blakiston, Philadelphia.
- Darwin C. 1859. *On the origin of species by means of natural selection*, 1st ed. Murray, London.
- Doyle JJ, Beachy RN. 1985. Ribosomal gene variation in soybean (*Glycine max*) and its relatives. *Theor Appl Genet* **70**: 369–376.
- Doyle JJ, Fligel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annu Rev Genet* **42**: 443–461.
- Drea SC, Lao NT, Wolfe KH, Kavanagh TA. 2006. Gene duplication, exon gain and neofunctionalization of OEP16-related genes in land plants. *Plant J* **46**: 723–735.
- Flavell RB, O'Dell M. 1976. Ribosomal RNA genes in homeologous chromosomes of groups 5 and 6 in hexaploid wheat. *Heredity* **37**: 377–385.

- Franzke A, Mummenhoff K. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae): Conversion of nuclear ribosomal ITS sequences in statu nascendi. *Theor Appl Genet* **98**: 831–834.
- 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.
- Gould SJ. 1994. The evolution of life on earth. *Sci Am* **271**: 85–86.
- Grant V. 1981 *Plant speciation*. Columbia University Press, New York.
- Grundt HH, Kjolner S, Borgen L, Rieseberg LH, Brochmann C. 2006. High biological species diversity in the arctic flora. *Proc Natl Acad Sci* **103**: 972–975.
- Hegarty MJ, Jones JM, Wilson ID, Barker GL, Coghill JA, Sanchez-Barcaldo P, Liu G, Buggs RJA, Abbott RJ, Edwards KJ, Hiscock SJ. 2005. Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation. *Mol Ecol* **14**: 2493–2510.
- 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. *Curr Biol* **16**: 1652–1659.
- Huskins CL. 1931. The origin of *Spartina townsendii*. *Genetica* **12**: 531–538.
- Kashkush K, Feldman M, Levy AA. 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* **160**: 1651–1659.
- Kim KJ, Jansen RK. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*): Additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Syst Evol* **190**: 157–159.
- Kovarik A, Matyasek R, Lim KY, Skalická K, Koukalová B, Knapp S, Chase M, Leitch AR. 2004. Concerted evolution of 18–5.8–26S rDNA repeats in *Nicotiana* allotetraploids. *Biol J Linn Soc* **82**: 615–625.
- Kovarik A, Pires JC, Leitch AR, Lim KY, Sherwood A, Matyasek R, Rocca J, Soltis DE, Soltis PS. 2005. Rapid concerted evolution in two allopolyploids of recent and recurrent origin. *Genetics* **169**: 931–944.
- Lim KY, Matyasek R, Lichtenstein CP, Leitch AR. 2000. Molecular cytogenetic analyses and phylogenetic studies in the *Nicotiana* section *Tomentosae*. *Chromosoma* **109**: 245–258.
- Lim KY, Soltis DE, Soltis PS, Tate J, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR. 2008. Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). *PLoS One* **3**: e3353.
- Löve A, Löve D. 1949. The geobotanical significance of polyploidy. I. Polyploidy and latitude. *Port Acta Biol Ser A* (special vol) **2**: 273–352.
- Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151–1155.
- Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**: 459–473.
- Ma XF, Gustafson JP. 2006. Timing and rate of genome variation in *Triticale* following allopolyploidization. *Genome* **49**: 950–958.
- Ma XF, Fang P, Gustafson JP. 2004. Polyploidization-induced genome variation in *Triticale*. *Genome* **47**: 839–848.
- Madlung A, Tyagi AP, Watson B, Jiang HM, Kagochi T, Doerge RW, Martienssen R, Comai L. 2005. Genomic changes in synthetic *Arabidopsis* polyploids. *Plant J* **41**: 221–230.
- Matyasek R, Tate JA, Lim YK, Srubarova H, Koh J, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 2007. Concerted evolution of rDNA in recently formed *Tropogon* allotetraploids is typically associated with an inverse correlation between gene copy number and expression. *Genetics* **176**: 2509–2519.
- Müntzing A. 1936. The evolutionary significance of autopolyploidy. *Hereditas* **21**: 263–378.
- Novak SJ, Soltis DE, Soltis PS. 1991. Ownbey *Tragopogons*: 40 Years later. *Am J Bot* **78**: 1586–1600.
- O’Kane SL, Schaal BA, Al-Shehbaz IA. 1996. The origins of *Arabidopsis suecica* as indicated by nuclear rDNA sequences. *Syst Bot* **21**: 559–566.
- Ownbey M. 1950. Natural hybridization and amphiploidy in the genus *Tragopogon*. *Am J Bot* **37**: 487–499.
- Ownbey M, McCollum GD. 1953. Cytoplasmic inheritance and reciprocal amphiploidy in *Tragopogon*. *Amer J Bot* **40**: 788–796.
- Ownbey M, McCollum GD. 1954. The chromosomes of *Tragopogon*. *Rhodora* **56**: 7–21.
- Paterson AH, Bowers JE, Burow MD, Draye X, Elisk CG, Jiang C-X, Katsar CS, Lan T-H, Lin Y-R, Ming R, Wright RJ. 2000. Comparative genomics of plant chromosomes. *Plant Cell* **12**: 1523–1539.
- Paterson AH, Chapman BA, Kissinger JC, Bowers JE, Feltus FA, Estill JC. 2006. Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*. *Trends Genet* **22**: 597–602.
- Rauscher JT, Doyle JJ, Brown AHD. 2004. Multiple origins and rDNA internal transcribed homoeolog evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. *Genetics* **166**: 987–998.
- Rosser EM. 1955. A new British species of *Senecio*. *Watsonia* **3**: 228–232.
- Sang T, Crawford DJ, Stuessy TF. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. *Proc Natl Acad Sci* **92**: 6813–6817.
- Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC. 2004. Mining EST databases to resolve evolutionary events in major crop species. *Genome* **47**: 868–876.
- Seoighe C, Gehring C. 2004. Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends Genet* **20**: 461–464.
- Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y. 2002. The hidden duplication past of *Arabidopsis thaliana*. *Proc Natl Acad Sci* **99**: 13627–13632.
- Soltis DE, Soltis PS. 1993. Molecular data facilitate a reevaluation of traditional tenets of polyploid evolution. *Crit Rev Plant Sci* **12**: 243–273.
- Soltis DE, Soltis PS. 1999. Polyploidy: Recurrent formation and genome evolution. *Trends Ecol Evol* **14**: 348–352.
- Soltis PS, Soltis DE. 2000. The role of genetic and genomic changes in the success of polyploids. *Proc Natl Acad Sci* **97**: 7051–7057.
- Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate JA, Mavrodiev EV. 2004. Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): Cytogenetic, genomic, and genetic comparisons. *Biol J Linn Soc* **82**: 485–501.
- Song K, Lu P, Tang K, Osborn TC. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploidy evolution. *Proc Natl Acad Sci* **92**: 7719–7723.
- Stebbins GL. 1947. Types of polyploids: Their classification and significance. *Adv Genet* **1**: 403–429.
- Stebbins GL. 1950. *Variation and evolution in plants*. Columbia University Press, New York.
- Sterck L, Rombauts S, Vandepoele K, Rouze P, Van de Peer Y. 2007. How many genes are there in plants (... and why are they there)? *Curr Opin Plant Biol* **10**: 199–203.
- Tate JA, Soltis PS, Soltis DE. 2004. Polyploidy in plants. In *The evolution of the genome* (ed. TR Gregory), pp. 372–426. Academic, New York.
- Tate JA, Ni ZF, Scheen AC, Koh J, Gilbert CA, Lefkowitz D, Chen ZJ, Soltis PS, Soltis DE. 2006. Evolution and expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* **173**: 1599–1611.
- Tate JA, Symonds VV, Doust AN, Buggs RJA, Mavrodiev EV, Soltis PS, and Soltis DE. 2009. Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae): 60 Years after Ownbey’s discovery. *Am J Bot* **96**: 979–988.

- Teshima KM, Innan H. 2008. Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics* **178**: 1385–1398.
- Urbanska KM, Hurka H, Landolt E, Neuffer B, Mummenhoff K. 1997. Hybridization and evolution in *Cardamine* (Brassicaceae) at Urnerboden, Central Switzerland: Biosystematic and molecular evidence. *Plant Syst Evol* **204**: 233–256.
- Vision TJ, Brown DG, Tanksley SD. 2000. The origins of genomic duplications in *Arabidopsis*. *Science* **290**: 2114–2117.
- Volkov RA, Borisjuk NV, Panchuk II, Schweizer D, Hemleben V. 1999. Elimination and rearrangement of parental rDNA in the allotetraploid *Nicotiana tabacum*. *Mol Biol Evol* **16**: 311–320.
- Wang J, Tian L, Lee HS, Wei NE, Jiang H, Watson B, Madlung A, Osborn TC, Doerge RW, Comai L, Chen ZJ. 2006. Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**: 507–517.
- Wendel JF. 2000. Genome evolution in polyploids. *Plant Mol Biol* **42**: 225–249.
- Wendel JF, Doyle JJ. 2004. Polyploidy and evolution in plants. In *Plant diversity and evolution* (ed. R. Henry), pp. 97–117. CABI, Cambridge.
- Wendel JF, Schnabel A, Seelanan T. 1995. Bidirectional inter-locus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci* **92**: 280–284.
- Zimmer EA, Martin SL, Beverley SM, Kan YW, Wilson AC. 1980. Rapid duplication and loss of genes coding for the α chains of hemoglobin. *Proc Natl Acad Sci* **77**: 2158–2162.

Cold Spring Harbor Symposia on Quantitative Biology

On the Origins of Species: Does Evolution Repeat Itself in Polyploid Populations of Independent Origin?

D.E. Soltis, R.J.A. Buggs, W.B. Barbazuk, et al.

Cold Spring Harb Symp Quant Biol 2009 74: 215-223 originally published online August 17, 2009
Access the most recent version at doi:[10.1101/sqb.2009.74.007](https://doi.org/10.1101/sqb.2009.74.007)

References

This article cites 70 articles, 18 of which can be accessed free at:
<http://symposium.cshlp.org/content/74/215.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the
top right corner of the article or [click here](#).

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
<http://symposium.cshlp.org/subscriptions>