Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes

Summary

Plant mitochondrial genomes are usually transmitted to the progeny from the maternal parent. However, cases of paternal transmission are known and are perhaps more common than once thought. This review will consider recent evidence, both direct and indirect, of paternal transmission (leakage) of the mitochondrial genome of seed plants, especially in natural populations, and how this can result in offspring that carry a mixture of maternally and paternally derived copies of the genome; a type of heteroplasmy. It will further consider how this heteroplasmy facilitates recombination between genetically distinct partners; a process that can enhance mitochondrial genotypic diversity. This will then form the basis for a discussion of five evolutionary questions that arise from these observations. Questions include how plant mitochondrial genome evolution can be placed on a sexual to asexual continuum, whether cytoplasmic male sterility (CMS) facilitates the evolution of paternal leakage, whether paternal leakage is more likely in populations undergoing admixture, how leakage influences patterns of gene flow, and whether heteroplasmy occurs in natural populations at a frequency greater than predicted by crossing experiments. It is proposed that each of these questions offers fertile ground for future research on a diversity of plant species.

I. Introduction

Plant mitochondrial genomes are often described as being highly conserved with regard to coding region sequence but quite fluid with regard to genome size and structure (Wolfe et al., 1987; Palmer & Herbon, 1988; Sloan, 2013). A fairly large proportion of the genome consists of noncoding DNA organized into repeats of varying size, and it is variation in noncoding content that contributes to the enormous range of mitochondrial genome sizes (200 kb to > 10 Mb) that has been observed among those seed plant species for which genome size has been documented (Sloan et al., 2012a). One major contributor to this structural fluidity is the frequent intra- and inter-molecular recombination events that are facilitated by these repeats (Kühn & Gualberto, 2012).

Additionally, and of particular relevance for this review, the mitochondrial genome is thought to be inherited maternally in most seed plants, though exceptions to the rule of maternal inheritance are known (Mogensen, 1996; Xu, 2005). Strict maternal inheritance of
the mitochondrial genome has several important consequences for its evolution. First, in seed plants maternal inheritance would mean that the direct fitness currency of the mitochondrial genome is derived from the quantity and quality of seeds produced; in contrast to that of the nuclear genome, whose fitness is a function of frequency of transmission through both seeds and pollen. Transmission solely through seed could also mean that nuclear and mitochondrial genomes could have very different dispersal patterns and population genetic structures, especially in plant species capable of long-distance pollen dispersal combined with limited seed dispersal.

Maternal inheritance of the mitochondrial genome also tends to limit within-individual (or within-cell) genetic diversity (heteroplasmy) and ultimately limits generation of mitochondrial genotypic diversity at the population level. First, and most obviously, it reduces mitochondrial heteroplasmy (or maintains homoplasy) because maternal inheritance precludes any mixing of maternal and paternal mitochondrial genomes, whereas bi-parental inheritance would result in offspring heteroplasmy to the degree that the maternally and paternally derived mitochondrial genomes differ. (Note that the term heteroplasm, as used here, refers to heterogeneity in genetic content, not genomic structure.) Homoplasy also reduces the ability of inter-molecular recombination to generate novel multilocus genotypes because creation of new genotypes by recombination requires that the two participating partners differ in genetic composition (just as recombination between two nuclear genes will only generate a new two-locus genotype if both loci are heterozygous). Indeed, it is for this reason that mitochondrial lineages have been described as evolving in a fashion similar to asexual organisms (e.g. Lynch & Blanchard, 1998).

While maternal inheritance of the mitochondrial genome does seem to be the usual case in seed plants, evidence for paternal transmission through pollen (leakage) continues to emerge from crossing studies conducted on a number of crop and model plant species. For example, documented paternal leakage ranges from crossing studies conducted on a number of crop and model plant transmission through pollen (leakage) continues to emerge from several single nucleotide polymorphisms (SNPs) within the cob sequence of Silene acaulis. Haplotype diversity determined by several single nucleotide polymorphisms (SNPs) within the cob sequence obeyed the four-gamete rule of Hudson & Kaplan (1985), which posits recombination to be a parsimonious explanation for the evolution of certain patterns of intra- or inter-genic haplotype diversity (see Fig. 1 for a schematic illustration). Note that recombination in this context generates this type of genotypic diversity when it occurs between two molecules that differ in sequence within the region of interest (i.e. within cells that are heteroplasmic, as might result from bi-parental inheritance). Similar evidence for recombination (and therefore indirect evidence for heteroplasmy and paternal leakage)

II. Indirect and direct evidence for paternal leakage and heteroplasmy

While observations of paternal transmission of plant mitochondrial genes have been made for more than 20 yr, most of these have resulted from controlled crosses, including interspecific crosses. Evaluation of the evolutionary significance of paternal transmission also requires an understanding of its frequency and consequences in natural populations.

Indirect evidence that paternal leakage may lead to heteroplasmy in natural populations came first from a study by Städler & Delph (2002) of sequence diversity of the mitochondrial gene cob in Silene acaulis. Haplotype diversity determined by several single nucleotide polymorphisms (SNPs) within the cob sequence obeyed the four-gamete rule of Hudson & Kaplan (1985), which posits recombination to be a parsimonious explanation for the evolution of certain patterns of intra- or inter-genic haplotype diversity (see Fig. 1 for a schematic illustration). Note that recombination in this context generates this type of genotypic diversity when it occurs between two molecules that differ in sequence within the region of interest (i.e. within cells that are heteroplasmic, as might result from bi-parental inheritance). Similar evidence for recombination (and therefore indirect evidence for heteroplasmy and paternal leakage)

Evolution of a fourth haplotype defined by A/C and G/T SNPs

CT originates by recurrent mutation

CT originates by recombination

![Fig. 1 Illustration of the ‘four-gamete rule’ of Hudson & Kaplan (1985) in which the fourth possible haplotype defined by two bi-allelic single nucleotide polymorphisms (SNPs) evolves by recombination between SNPs (right) or by recurrent mutation (homoplasy) without recombination (left). Note that SNPs can fall within the sequence of a given gene (allowing detection of intra-genic recombination) or individual SNPs can each fall within the sequence of a different gene (allowing detection of inter-genic recombination).](image-url)
was also obtained from analysis of the sequence diversity of mitochondrial genes sampled from natural populations of *Silene vulgaris* and *Silene nutans* (McCauley et al., 2005; Houlston & Olson, 2006; Touzet & Delph, 2009).

Indirect evidence for leakage and heteroplasmy also comes from a population genetic study of wild carrot (*Daucus carota* ssp. *carota*), a weedy relative of domesticated carrot. In that study, SNPs and indel variants were scored for three mitochondrial genes found in individuals sampled from a number of North American natural populations (Mandel et al., 2012). Evidence of recombination between genetically distinct mitochondrial genomes again came from observations of all four possible two-locus genotypes among the individuals scored for their *atp1* and *atp9* genotypes. As in the *Silene* studies, the interpretation was that creation of novel two-locus mitochondrial genotypes by recombination requires heteroplasmity, which in turn could arise when paternal leakage contributes to biparental inheritance of the mitochondrial genome (Mandel et al., 2012).

Direct evidence of mitochondrial paternal leakage and heteroplasmy in the genus *Silene* (specifically *S. vulgaris*) was first presented by McCauley et al. (2005). That study made use of the inheritance of SNPs in two mitochondrial genes that were assayed using a PCR/RFLP methodology (Ota et al., 2007). Experimental crosses were made between individuals that differed at a restriction site in either the *atp1* or *cox1* genes. Genotyping revealed maternal inheritance of the *atp1* or *cox1* SNPs in 96% of the offspring. However, a few offspring displayed restriction fragment patterns that resembled paternal genotypes or a restriction pattern that appeared to be an overlay of maternal and paternal genotypes. It was suggested that the results could be explained by occasional paternal leakage of the mitochondrial genome resulting in apparent paternal inheritance, or in heteroplasmy. Note that in this study, and other studies of heteroplasmy in *Silene* described in the following paragraphs, the DNA used in genotyping an individual was extracted from a small fragment of a single leaf (not single cells), and any assessment of heteroplasmy from such extractions might not reflect the level of heteroplasmy at lower, or higher, levels of organization of that individual.

The results of McCauley et al. (2005) also suggested that in some cases the relative representation of two variants in a heteroplasmic individual could be so uneven as to be undetectable by standard PCR/RFLP methods (so-called ‘cryptic heteroplasmy’). In a follow-up study of *S. vulgaris*, Welch et al. (2006) developed a quantitative PCR (q-PCR) method designed to detect heteroplasmy for *atp1* SNP markers even when the relative representation of two variants is very unequal. The method also allows quantification of their relative proportions. In that study, genotypic arrays were created by genotyping maternal leaf material collected from natural populations and offspring genotypes derived from individuals grown from seeds carried by the mothers at the time of the field leaf sample collection. Relevant results were that a few of the mothers were heteroplasmic, as were a number of the offspring, including some with homoplasmic mothers. A heteroplasmic offspring of a homoplasmic mother would be suggestive of bi-parental inheritance via paternal leakage, though paternal genotypes were unknown in these open-pollinated systems.

A study by Pearl et al. (2009) extended this approach to a much larger number of individuals sampled from a larger number of North American *S. vulgaris* populations. Again, DNA was extracted from leaf material sampled from mature individuals growing in natural populations, and from seedlings grown from seeds taken from a subset of those individuals. Overall 15.9% of the mature individuals sampled displayed heteroplasmy for either the *atp1* or *cox1* q-PCR marker (or both). However, in most cases the relative representation of the two variants was very uneven (Fig. 2). Approximately 5% of offspring were heteroplasmic. These included heteroplasmic offspring derived from homoplasmic mothers (suggestive of bi-parental inheritance following paternal leakage) and heteroplasmic offspring derived from heteroplasmic mothers (suggesting that heteroplasmy could be inherited once incorporated into a maternal line). Again, in the heteroplasmic offspring the relative representation of the two variants was typically uneven, with the maternal genotype predominant. In a few cases a homoplasmic mother produced an offspring that was homoplasmic for a different variant, suggesting pure paternal inheritance. In other cases a heteroplasmic mother produced offspring that were homoplasmic.

![Fig. 2](image_url) The proportion of *Silene vulgaris* adults with haplotype scores between 0.005 and 0.995. The haplotype score refers to the within-individual copy number of *atp1* haplotype B relative to the other *atp1* haplotypes (a), or the within-individual copy number of *cox1* haplotype 3 relative to the other *cox1* haplotypes (b). Heteroplasmy is defined arbitrarily as cases in which the frequency of the less common of two co-occurring haplotypes exceeds 0.005. (From Pearl et al., 2009, with permission from Oxford University Press.)
Transmission of a finite number of mitochondria from mother to daughter cells during cell division, and other processes associated with replication that can limit the effective number of mitochondrial genomes/cell, can result in a type of random sampling (see Birky, 2001 for more detail). These sampling events can be compounded across cell divisions and result in a process within individuals, somewhat akin to random genetic drift, in which all mitochondrial genetic diversity is lost within a cell line. Thus, any heteroplasmy that might be already present in an individual, arise by mutation, or arise by bi-parental inheritance, could be lost in cell lineages destined to produce egg cells, or during cell divisions following the fertilization of those eggs. Because the results of Pearl et al. (2009) were again derived from open-pollinated systems, none of the possible modes of inheritance could be demonstrated directly.

In order to study the inheritance of the mitochondrial genome and the origin and transmission of heteroplasmy more directly, Bentley et al. (2010) conducted a number of experimental glasshouse crosses between *S. vulgaris* individuals of known *atp1* q-PCR marker. Crosses between individuals that were homoplasmic for different markers would demonstrate paternal leakage in any heteroplasmic offspring or an offspring homoplasmic for the paternal genotype. Overall, c. 2.5% of the 1152 offspring assayed displayed evidence of paternal leakage. The majority of those individuals indicating leakage were heteroplasmic with the maternal genotype predominant, although in a few individuals the paternal genotype predominated (Fig. 3). Since the publication of Bentley et al. (2010) additional crosses diagnostic for the same *atp1* q-PCR marker have produced another 719 offspring genotypes (D. E. McCauley, unpublished data, summarized in Supporting Information Table S1). Again rare leakage was detected (3.7% of offspring), with any paternal contribution typically minor relative to the maternal representation. Combining both data sets (1871 total offspring genotypes) yields a leakage rate of 3.0%.

Other crosses conducted by Bentley et al. (2010) involved heteroplasmic mothers and pollen donors homoplasmic for the predominant maternal type. Assuming that paternal leakage is rare, a heteroplasmic offspring resulting from such a cross would have inherited that condition from the mother. Maternal heteroplasmy would be assumed to have been lost by vegetative segregation in any homoplasmic offspring produced by such a cross. The results of these crosses were that 19.6% of the offspring were heteroplasmic, demonstrating that heteroplasmy can be inherited, although it is more often lost as a result of vegetative segregation. In those offspring that did inherit heteroplasmy from the maternal line, the level of heteroplasmy (evenness of representation) declined, relative to that found in the mother, more often than it increased.

Recall that two-locus *atp1/cox1* PCR/RFLP genotyping of a number of individuals sampled from natural populations by McCauley et al. (2005) resulted in a distribution of genotypes indicative of cross-over events involving genomes that differed at the marker SNPs found in the two genes. McCauley & Ellis (2008) expanded this study by obtaining three-locus (*atp1, cox1* and *cob*) PCR/RFLP mitochondrial genotypes for a much larger sample of individuals from 22 North American *S. vulgaris* populations. By examining the diversity of all possible two-locus genotypes (*atp1/cob*, *atp1/cox1* and *cox1/cob*) they inferred recombination between each pair of these three genes. The population genetic consequences of this recombination were quantified by calculating pairwise linkage disequilibrium (LD) values. LD is a measure of the degree of nonrandom association between two genetic markers and is positive when a particular pair of markers co-occurs more often than the random expectation, and negative when the co-occurrence is at a frequency less than expected (Hedrick, 2011). LD can be defined at the intra- or inter-locus level and is broken up by recombination. When standardized, LD values would be expected to approach 1 (maximum positive statistical association, given the allele frequencies), or −1 (maximum negative statistical association) with no recombination, and approach a value of zero when recombination occurs frequently. Therefore, it is a good measure of the impact of paternal leakage, heteroplasmy, and recombination on the population genetics of mitochondrial genomes. McCauley & Ellis (2008) found absolute standardized LD values (symbolized |D|) ranging from 0.17 to 0.78 when the two most common alleles at each of two respective loci were considered. This suggests that recombination can enhance multilocus mitochondrial genotype diversity in natural populations, but is not common enough to completely eliminate associations between alleles that might arise through historical evolutionary processes.

That heteroplasmy resulting from bi-parental inheritance of mtDNA allows for recombination between divergent mitochondrial genomes was shown directly by Apitz et al. (2012).
Recombinant mitochondrial genotypes were detected in F1 offspring produced by crossing two species of *Pelargonium* displaying diagnostic differences in two mitochondrial genes.

### III. Five evolutionary issues raised by these observations

The results described in the previous section suggest that, in at least some populations of some species, paternal leakage of the mitochondrial genome occurs and results in heteroplasmcy, and that heteroplasmy, in turn, enables recombination to generate genotypic diversity not easily evolved by mutation alone. Further, the frequency and evenness of that heteroplasmy are controlled by the balance between gains by leakage and loss via vegetative segregation. These observations raise five evolutionary questions discussed in some detail in the following sections.

1. **How sexual are plant mitochondrial genomes?**

Mitochondrial genomes are often thought of as evolving in a fashion similar to asexual organisms and, if so, may be particularly vulnerable to mutation accumulation via Muller’s ratchet (e.g., Gabriel *et al.*, 1993; Lynch & Blanchard, 1998; Hooekstra, 2000; Loewe, 2006; Neiman & Taylor, 2009; see Muller, 1964 and Felsenstein, 1974 for general discussions of the ratchet process). The asexual-like nature of mitochondrial genomes derives from maternal inheritance. As already described, uni-parental inheritance would enforce homoplasy within individuals, and, in turn, limit the opportunity for recombination to occur between genetically distinct partners. That is, the mode of inheritance limits the opportunity for the creation of genotypic novelty via recombination, a key feature of sexual systems. The relevance of Muller’s ratchet to the evolution of plant mitochondrial genomes may be questioned if the rate of deleterious mutation within their coding regions is very low, as evidenced by general low rates of plant mtDNA nucleotide substitution (Wolfe *et al.*, 1987). However, the view that such coding regions are universally highly conserved has been modified recently, depending on which mitochondrial genes and which plant species are considered (Cho *et al.*, 2004; Barr *et al.*, 2007; Sloan *et al.*, 2008).

Recently, the view that mitochondrial genomes evolve as asexual lineages has been challenged as a result of a more complete understanding of mitochondrial transmission and population biology, especially the observation that paternal leakage and heteroplasmcy might be more common than once thought in both plants and animals (e.g., McVean, 2001; Kmiec *et al.*, 2006; McCauley & Olson, 2008; White *et al.*, 2008; Galtier *et al.*, 2009; Wolff *et al.*, 2013). The challenge to asexuality might be especially appropriate for plant mitochondrial genomes as it is well known that in plant mitochondrial genomes intra- and inter-molecular recombination is facilitated by the presence of large noncoding repeats and by frequent mitochondrial fusion (Mackenzie & McIntosh, 1999; Arimura *et al.*, 2004; Sheahan *et al.*, 2005; Kühn & Gualberto, 2012). The results discussed in section II indicate that recombination can sometimes occur between genetically distinct mitochondrial genomes. Given that plant mitochondrial genomes do have properties not associated with strictly asexual systems, how ‘sexual’ are they? From a population genetic perspective the degree of sexuality of an organellar genome might be measured by the degree to which genotypic novelty is generated by recombination, or by the magnitude of LD among mitochondrial genes. As different parts of the mitochondrial genome could vary in propensity for recombination as a result of an uneven distribution of the repeats and other aspects of genomic structure (Sloan, 2013), the degree of sexuality might differ depending on which portion of the mitochondrial genome is considered, just as the recombination rate varies across the nuclear genome.

Given the accumulated knowledge with regard to paternal leakage, heteroplasmcy, and the mitochondrial genomics of *S. vulgaris*, it would be useful to try to place the mitochondrial genome of that species on a continuum of sexuality that ranges from strictly asexual (no genotypic diversity generated by recombination; maximum LD among mitochondrial genes) to purely sexual (all mitochondrial genes in linkage equilibrium). This might then provide a benchmark for future analyses of the mitochondrial genomes of other species of plants. In many ways this question mirrors studies of the degree of clonality of bacteria and other microbes. Those studies also use patterns of LD in natural populations to infer the impact of recombination on genotypic diversity (e.g., Lenski, 1993; Maynard Smith *et al.*, 1993; Halkett *et al.*, 2005; Apitz *et al.*, 2012; Tibayrenc & Ayala, 2012).

While the results of McCauley & Ellis (2008) provide some insight into the question of sexuality in *S. vulgaris* mitochondrial genome evolution by considering LD among a trio of mitochondrial genes, a more complete answer would require sampling a larger number of genes located across the mitochondrial genome. To that end, 75 individuals from 22 North American *S. vulgaris* populations were recently genotyped for SNPs in nine mitochondrial loci (*atp1, atp6, atp9, ccmFc, ccmFh, cob, cox1, matRnad4*) and the resulting nine-locus genotypes (mitotypes) examined for evidence of intra- and inter-genic recombination (D. E. McCauley, unpublished). Sampled individuals fell into two categories. Some individuals were a subset of those studied by McCauley & Ellis (2008) but now genotyped for all nine genes, while others were sampled from natural populations in the summers of 2010 and 2011.

SNPs used as markers were identified as follows. Coding regions derived from two of the complete *S. vulgaris* mitochondrial genome sequences reported in Sloan *et al.* (2012b) were aligned and SNPs identified in the nine genes of interest. Primers were designed to target regions flanking these SNPs (Table S2) and used in PCR amplification of each of these regions from genomic DNA following the methods of McCauley & Ellis (2008). The resulting PCR products were subjected to Sanger sequencing by the Vanderbilt University Sequencing Facility. Sequences were aligned using the software SEQUENCER (Gene Codes Corp., Ann Arbor, MI, USA) and then trimmed so that the SNPs of interest could be scored within these sequences, along with several additional SNPs not present in the original pairwise genomic sequence comparisons (Table S3).

Altogether 37 SNPs were considered across the sequences of the nine genes studied (Table 1). Nineteen of these SNPs (51%) were
within the atp1 sequence, a result consistent with previous documentation of especially high variability of atp1 in S. vulgaris (e.g. Barr et al., 2007). The number of unique haplotypes per gene ranged from four in atp1 to two in cob, atp9 ccmFc and atp6 (Tables 1, S3). Considering all nine genes together, individuals carrying 23 different nine-locus haplotype (mitotype) combinations were observed (Table 1).

How does this nine-locus genotypic diversity compare to that expected with no recombination? It is possible to extend the four-gamete rule of Hudson & Kaplan (1985), which applies to two loci, each with two alleles (Fig. 1), to a more general case. If multilocus genotypes evolve by sequential mutations that result in a branching process with no recombination or recurrent mutation (homoplasia), then the number of multilocus genotypes possible for n loci, with x_i alleles per locus, equals \( \sum_i x_i - n + 1 \). For the data set at hand that would be \((24 - 9 + 1) = 16\) possible genotypic combinations. Thus, by this measure, the organization of the observed allelic diversity into 23 nine-locus mitochondrial genotypes is an indication of the impact of inter-genic recombination on genotypic diversity.

When haplotype combinations for individual pairs of the nine genes were considered, evidence for inter-genic recombination using the four-gamete rule of Hudson & Kaplan (1985) was found for 21 of the 36 possible pairs (Table 2). This test could also be used to examine intra-genic recombination within the six genes containing multiple SNPs. Of these, only atp1 showed evidence for intraingenic recombination. The population genetic consequences of recombination were further quantified by calculating the absolute value of the standardized LD (|D|) among all pairs of genes. Ignoring intra-genic recombination, and even then pairwise LD among the 19 atp1 SNPs was near its maximum.

Individual loci differed considerably in their respective degree of recombination with the rest of the genome. Three loci (atp1, cox1 and nad4) each showed evidence of recombination with seven of the eight other loci (i.e. pairwise ID' \( < 1.00 \)), whereas atp6 only recombined with atp9 by this measure. This yielded average locus-specific values of pairwise ID' as low as 0.72 (for nad4 and atp9) and as high as 0.99 (for atp6).

This method was also used to calculate ID' within the six loci pairing, to 0.91 for the 15 pairs of loci showing no evidence of recombination using the four-gamete test (Table 2). How might this information be used to quantify the degree of sexuality of the S. vulgaris mitochondrial genome? In an ideal fully sexual panmictic system, we might expect complete linkage equilibrium among all pairs of genes. Ignoring intra-genic recombination (which is apparently rare in our system), the SNPs found among the nine mitochondrial loci studied can be combined into \( \prod_{k=1}^{9} k! = 5184 \) possible nine-locus genotypes, where \( k! \) represents the number of haplotype blocks at locus \( i \). The expected frequency of each combination would then be \( \prod_{j=1}^{n} p_{ij} \), where \( p_{ij} \) represents the

---

**Table 1** Single nucleotide polymorphisms (SNPs) in nine mitochondrial genes that, considered together, define 23 unique haplotypes (mitotypes), along with the number (n) of each haplotype encountered in a sample of 75 *Silene vulgaris* individuals taken from 22 North American populations (the ordering of the genes is arbitrary)

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>atp1</th>
<th>cox1</th>
<th>cob</th>
<th>atp9</th>
<th>ccmFc</th>
<th>nad4</th>
<th>atp6</th>
<th>ccmFn</th>
<th>matR</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CATCGGATCGCCTTGATGAA</td>
<td>ACCAC</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>ATC</td>
<td>CA</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>TACTAGACACCCTGATGAA</td>
<td>ACCAC</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>AA</td>
<td>CGC</td>
<td>ATC</td>
<td>CA</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>TACTAGACACCCTGATGAA</td>
<td>ACCAC</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>TACTAGACACCCTGATGAA</td>
<td>ACCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>TA</td>
<td>CGC</td>
<td>AGA</td>
<td>CC</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>AGA</td>
<td>CA</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>TA</td>
<td>CGC</td>
<td>AGA</td>
<td>CC</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>TACTAGACACCCTGATGAA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>TATTGACACATCCGAGA</td>
<td>ACCAC</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>ATC</td>
<td>CA</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>TATTGACACATCCGAGA</td>
<td>ACCAC</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>AA</td>
<td>CGC</td>
<td>ATC</td>
<td>CA</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>TATTGACACATCCGAGA</td>
<td>ACCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>AGA</td>
<td>CA</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>AC</td>
<td>CGC</td>
<td>ATC</td>
<td>CA</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>AGA</td>
<td>AA</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AA</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>AGA</td>
<td>AA</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>TATTGACACATCCGAGA</td>
<td>CTCAC</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>CTC</td>
<td>CC</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>TGTCGATCGAGCGTACGGT</td>
<td>ACTCA</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>AC</td>
<td>Ttg</td>
<td>AGA</td>
<td>AA</td>
<td>15</td>
</tr>
<tr>
<td>21</td>
<td>TGTCGATCGAGCGTACGGT</td>
<td>ACTCA</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>AGA</td>
<td>AA</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>TGTCGATCGAGCGTACGGT</td>
<td>ACTCA</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>Ttg</td>
<td>AGA</td>
<td>AA</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>TGTCGATCGAGCGTACGGT</td>
<td>ACTCA</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>AC</td>
<td>CGC</td>
<td>AGA</td>
<td>AA</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 2 Absolute values of the standardized linkage disequilibrium |D|’ computed for all 36 possible pairings of nine mitochondrial genes (off diagonal) and within the six of those genes (diagonal) in which the presence of more than one marker single nucleotide polymorphism (SNP) allowed for calculation of |D|’

<table>
<thead>
<tr>
<th></th>
<th>atp1</th>
<th>cox1</th>
<th>cob</th>
<th>ccmFc</th>
<th>nad4</th>
<th>atp6</th>
<th>ccmFn</th>
<th>matR</th>
</tr>
</thead>
<tbody>
<tr>
<td>atp1</td>
<td>0.99</td>
<td>0.65</td>
<td>0.81</td>
<td>0.71</td>
<td>0.80</td>
<td>0.67</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>cox1</td>
<td></td>
<td>1.00</td>
<td>0.83</td>
<td>0.61</td>
<td>0.83</td>
<td>0.66</td>
<td>1.00</td>
<td>0.66</td>
</tr>
<tr>
<td>cob</td>
<td></td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>0.78</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>atp9</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.61</td>
<td>0.91</td>
<td>0.39</td>
<td>0.91</td>
</tr>
<tr>
<td>ccmFc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>nad4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>atp6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ccmFn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>matR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

In each entry |D|’ was calculated separately for all possible combinations of alleles and then averaged according to a weighting scheme described in the text.

frequency of haplotype block j found in locus i. This expectation could be calculated for the data at hand by estimating pij from the 75 individuals genotyped. When this is done, the expected frequency of the genotype combining the most common allele at each of the nine loci is 0.006, with the vast majority of the 5184 possible nine-locus genotypes having an expected frequency < 0.005 (calculations not shown). Perhaps not surprisingly, of the 23 observed multilocus genotypes, five were among the ten expected to be most common by this measure.

How, then, does the observed genotypic diversity seen in our data compare to the maximum expected with complete linkage equilibrium within the mitochondrial genome? The expected genotypic diversity is 1 – |D|’ = 1 – |D|’ = 0.998 when the 5184 expected genotypic frequencies are used in the calculation. This can be compared to the observed genotypic diversity, calculated as 1 – Σ(qj^2), where qj represents the observed frequency of each of the 23 observed genotypes listed in Table 1. This yields a value of 0.905. Thus, the data set includes 0.905/0.998 = 90.6% of the total genotypic diversity theoretically possible if all nine mitochondrial loci were in linkage equilibrium with one another, as might be expected in a hypothetical panmictic sexual system.

Two things are clear from this analysis. First, recombination within the mitochondrial genome of S. vulgaris generates much, but not all, of the genotypic diversity expected for a fully ‘sexual’ system. Secondly, the impact of that recombination on genetic diversity is not evenly distributed among the genes used in this study. For example, it appears that the atp6 gene has scarcely recombined with the rest of the genome. As already mentioned, enhancement of genotypic diversity via recombination requires that recombination be between pairs of genes that are each heteroplasmic within the same cell. In plant mitochondrial genomes, the likelihood that recombination occurs between a given pair of genes would seem to depend on their positions relative to one another. The results of Sloan et al. (2012b) suggest, however, that the structure of the S. vulgaris mitochondrial genome is sufficiently fluid that the relative positions of a given pair of genes can vary greatly from individual to individual. This would make it difficult to evaluate the influence (if any) of structural linkage on the frequency of recombination. However, the distance between marker SNPs might explain the observed paucity of intra-genic recombination, as members of a pair of marker SNPs within a gene must be consistently in much greater proximity to one another than those on different genes.

Given that recombination within the mitochondrial genome of S. vulgaris appears to generate more genotypic diversity than might be expected in an otherwise equivalent asexual system, one might consider the possible consequences of that recombination for the evolution of the mitochondrial genome. As already stated, much of the original motivation for thinking about the population genetics of mitochondrial genomes as essentially asexual derived from the fact that Muller’s ratchet operates more efficiently in the absence of recombination. However, even a small amount of recombination can counter this (e.g. Charlesworth et al., 1993 and others; reviewed in Neiman & Taylor, 2009) and so the impact of Muller’s ratchet on the accumulation of deleterious mutations within the mitochondrial genome of S. vulgaris may be limited when compared with a truly asexual genome. Neiman & Taylor (2009) point out that a lack of recombination also facilitates the operation of Muller’s ratchet by lowering effective population size (N_e), an effect that could also be ameliorated by the partially ‘sexual’ nature of the mitochondrial genome of S. vulgaris.

Population genetic studies of plant mitochondrial genomes (and especially measures of LD among mitochondrial genes) are limited in number and so it is not yet clear how the degree of sexuality found for the mitochondrial genome of S. vulgaris compares with that in other species of plants. Once this is known it might be interesting to compare those species with regard to levels of mutation accumulation. It would be particularly important to do so in those species for which the rate of nucleotide substitution in mtDNA coding regions would generate rates of advantageous or deleterious mutation sufficient for recombination within the genome to be an important influence on its evolution.

2. Paternal leakage and CMS

The apparent near ubiquity of uni-parental, usually maternal, inheritance of mtDNA in nature has led to consideration of what evolutionary pressures might select for this mode of inheritance. Discussion has centered on the fact that the mitochondrial and nuclear genomes replicate largely independently of one another, a circumstance that could result in inter-genomic evolutionary conflict (Cosmides & Tooby, 1981). Burt & Trivers (2006) suggest that uni-parental inheritance of mtDNA is in the evolutionary interests of the nuclear genome in that it limits within-individual mtDNA diversity and thus the opportunity for within-individual selection. It is this within-individual selection that can lead to a ‘selfish’ mitochondrial genome. Thus, regulation of the mode of inheritance is thought to be a nuclear trait and arise by one of several mechanisms. First, replication of mtDNA in the sperm or pollen cell could be reduced just prior to fertilization, limiting the paternal mtDNA copy number that could be introduced to the egg. Secondly, any paternal contribution could be actively or passively
eliminated subsequent to fertilization, leaving only the maternal contribution in the developing embryo (Nagata, 2010; Wang et al., 2010; Matsushima et al., 2011).

One particularly interesting feature of some plant mitochondrial genomes is the existence of CMS factors that interfere with anther and pollen development and render otherwise hermaphroditic individuals functionally female (Schnable & Wise, 1998). When maternally inherited, these CMS factors can spread in populations if they increase the quantity or quality of seeds produced by the individuals that carry them, given that seeds, but not pollen, are the direct currency of fitness for maternally inherited plant genes. Indeed, it is the advantage through seed production created by CMS factors that is often thought to drive the evolution of gynodioecious plant species in which female and hermaphroditic individuals coexist (Frank, 1989; Shykoff et al., 2003; Dufay & Billard, 2012). Such systems often have a form of cytonuclear sex determination in which the phenotypic effect of CMS can be modified by suitable nuclear restorers (McCauley & Bailey, 2009; Touzet, 2012).

Burt & Trivers (2006) have proposed in a verbal argument that the presence of CMS factors in gynodioecious plant systems can actually select for paternal transmission (leakage) of their mitochondrial genomes. This is because it is to the evolutionary advantage of a nuclear genome to be paired with a fertile cytotype, as those nuclear genes could then be transmitted via pollen. In systems with CMS and cytonuclear sex determination, a pollen donor is more likely to carry a fertile cytotype than a pollen recipient, as the donor must be a hermaphrodite while the recipient could be either hermaphroditic or female. Hence, leakage enhances the probability of producing a male-fertile phenotype. A formal mathematical model of this process does not yet exist.

The possibility that gynodioecious species with cytonuclear sex determination are more likely to display paternal leakage of the mitochondrial genome has interesting potential consequences for the evolution and maintenance of that breeding system. When mitochondrial genomes can be transmitted through pollen, the currency of fitness of those genomes differs from the ‘seeds only’ currency incorporated into models of the evolution of gynodioecy that assume strict maternal inheritance of mitochondrial genes. As pointed out by Wade & McCauley (2005), this gives fertile cytotypes an advantage when rare, because of the high per capita availability of female mates. Under those conditions, transmission of male-fertile mitochondrial genomes through pollen produced by hermaphrodites can offset the added seed productivity often associated with females carrying CMS. Because the advantage is frequency-dependent it can stabilize the sex ratio of gynodioecious systems.

If CMS does select for paternal leakage, and if this stabilizes gynodioecious systems, one might predict cases of leakage to be over-represented in gynodioecious species with cytonuclear sex determination. In fact, recall that several of the species for which evidence of paternal leakage is strongest (e.g. S. acaulis, S. nutans, S. vulgaris and D. carota) have at least some gynodioecious populations with cytonuclear sex determination. However, there could be a type of ascertainment bias given the suggestion that gynodioecious species tend to have higher levels of mitochondrial gene sequence diversity of the sort required for construction of the markers needed to detect leakage (Touzet, 2012). Clearly, rigorous tests for mitochondrial leakage would have to be conducted across a wide array of species, representing diverse breeding systems, before any statistical evidence for an evolutionary association between CMS and leakage could be established. In addition, it would be worthwhile to develop a formal mathematical model of the argument by Burt and Trivers that CMS should select for leakage in order to provide a clear conceptual framework for the empirical studies. This is especially desirable given empirical observations that paternal leakage can result in heteroplasmy, in addition to pure paternal inheritance.

3. What determines among-population variation in rates of mitochondrial leakage?

Bentley et al. (2010) noted that in their crosses plants originating from different populations had different leakage rates. This heterogeneity was found to be statistically significant (ranging from zero leakage among the offspring of pollen donors whose ancestry traced to some populations, to > 10% among offspring sired by pollen donors traced to others). When these same crosses were grouped according to the population of origin of the pollen recipient, rather than the pollen donor, population-specific leakage rates were nearly as heterogeneous in magnitude, although that heterogeneity was not statistically significant.

Combining the data of Bentley et al. (2010) with that derived from the unpublished experimental crosses described in Table S1 allows a similar analysis of the resulting larger data set, but with greater statistical power. Population-specific leakage rates were calculated only for those populations contributing > 100 offspring to this larger data set (Fig. 4). These rates were then tested for heterogeneity using a G-test of independence (Sokal & Rohlf, 2012). Statistically significant heterogeneity in rates of leakage was
detected among both origins of pollen donors \((G = 21.47; \text{df} = 7, \ P = 0.003)\) and origins of pollen recipients \((G = 20.25; \text{df} = 7, \ P = 0.005)\). However, Fig. 4 shows that population-specific rates for pollen donors and pollen recipients are not correlated.

One might wonder what factor or factors influence observed among-population variation in the propensity for mitochondrial leakage in *S. vulgaris*. One possibility is that there is genetic variation for traits contributing to paternal leakage, and that this is structured geographically. Could this be related to the history of those populations? The literature suggests a general association, in both plants and animals, between mtDNA and/or cpDNA leakage and hybridization, perhaps resulting from a breakdown in the mechanisms that usually enforce uni-parental inheritance (Rokas *et al.*, 2003; Bazz *et al.*, 2005; Xu, 2005). For some specific recent examples, see Akeyanova *et al.* (2005), Jaramillo-Correa & Bousquet (2005), Fontaine *et al.* (2007), Trusty *et al.* (2007), Hoarau *et al.* (2009) and Weihe *et al.* (2009). Mechanisms could include properties of the pollen or sperm determined before fertilization and/or exclusion or dilution of any paternal contribution after fertilization (Birky, 1995, 2001). Given the apparent association between hybridization and leakage, one could speculate that at the intra-specific level a higher incidence of leakage could be found in populations displaying a high degree of admixture, if the consequences of admixture for enforcement of uni-parental inheritance resemble those of interspecific hybridization.

It would be interesting to see whether mitochondrial paternal leakage is more common in species in which admixture of genetically differentiated variants is pervasive. The latter might be the case, for example, in those invasive plant species in which multiple introductions result in admixture of geographic variants found in the native range. Could this be the case in *S. vulgaris*, which is native to Europe and invasive in North America? A recent study by Keller & Taylor (2010) showed that North American populations of *S. vulgaris* display varying degrees of admixture of diverse European ancestry as a result of the colonization process. Perhaps those North American populations with a greater occurrence of leakage (Fig. 4) tend to be those with a greater degree of nuclear or cytonuclear admixture.

4. Paternal leakage, gene flow and population structure

The population genetic structure of a species can be defined as the distribution of genetic variance within and among its populations, at some spatial scale. It is often quantified by Wright’s *F* sub population statistic (Wright, 1978), whose value can range from zero to one. In a simple two-allele case an *F* sub population is equal to zero means that all genetic variance is within populations (i.e. all populations of equal allele frequency). *F* sub population equal to one indicates that all genetic variance is distributed among populations (i.e. all populations fixed for one allele or the other). For selectively neutral genes, *F* sub population is a function of rates of gene flow combined with effective population size and history.

It has been noted that in plants the population structure \((F_{st})\) of mitochondrial (and chloroplast) gene variants is often much greater than that of nuclear genetic markers (Petit *et al.*, 2005). This is thought to be due to two reasons. First, if a cytoplasmic genome has strictly maternal inheritance then its gene flow results only from seed movement, whereas nuclear genes can potentially move both in seeds and in pollen. Secondly, homoplasmic cytoplasmic genomes are essentially haploid, resulting in a lower effective population size than that of the diploid nuclear genome of the same individuals, especially in species with separate sexes (Takahata & Maruyama, 1981; Birky *et al.*, 1983). Petit *et al.* (1993) point out, however, that even a low level of paternal leakage of a cytoplasmic genome would be expected to reduce its population genetic structure relative to that expected with strict maternal inheritance because it would allow some gene flow by pollen. This raises the question of whether the mitochondrial genomes of species in which paternal leakage has been implicated have particularly low levels of population structure.

In *S. vulgaris* this appears not to be the case. The structure of local populations as measured by the *F* sub population of mtDNA variants is nearly three times greater than the population structure measured for the same populations by nuclear allozyme markers (McCaulay, 1998; Olson & McCauley, 2002). However, the *F* sub population of 0.34 reported for the *D. carota* mitochondrial genome by Mandel *et al.* (2012) is substantially lower than the average *F* sub population (0.64) reported for plant cytoplasmic genomes by Petit *et al.* (2005) in their literature survey of many species.

It would appear from theory that paternal leakage of mitochondrial genes has the potential to have a large impact on their population structure in those species in which it occurs. Whether that is indeed the case remains an open question, given the paucity of species for which the presence or absence of paternal transmission of mitochondrial genes is known definitively, along with the genetic structure of their natural populations. One group of species of particular interest in this regard might be those with CMS-based gynodioecy, as the relative population structures of CMS variants and their associated nuclear restorers are thought to influence spatial sex-ratio variation and the evolution of the sex ratio (e.g. Bailey & Delph, 2007; McCaulay & Bailey, 2009). If CMS indeed selects for mitochondrial paternal leakage (see question 2), the propensity for CMS variants to be more highly spatially structured than their restorers might not be as great as a broader comparison of mtDNA and nuclear DNA population structures based largely on non-CMS species might suggest.

5. Does the rate of leakage observed in experimental crosses predict the level of heteroplasmy found in nature?

Ignoring rare input from new mutations, it would seem that the standing crop of heteroplasmic individuals that occurs in a natural population would represent a mixture of individuals made heteroplasmic by paternal leakage during their respective fertilization events and individuals who inherited heteroplasm from heteroplasmic mothers (Pearl *et al.*, 2009). The proportion of heteroplasmic individuals would depend, then, on the balance between the rate at which heteroplasmy is gained as a result of recent paternal leakage, and the rate at which it is lost as a result of vegetative segregation, the so-called ‘leakage equilibrium’ of McCaulay & Olson (2008). That equilibrium point could be predicted crudely with the crossing data at hand and compared to the frequency of heteroplasmy observed in
natural populations of *S. vulgaris*. If the prediction is not supported by the observation, perhaps forces other than vegetative segregation maintain heteroplasmy, or drive it from populations, once it is established by leakage.

Let \( H_0 \) be the frequency of heteroplasmy in this generation and \( H_1 \) be the frequency of heteroplasmy in the next. Let the term \( X \) symbolize the probability of paternal leakage leading to heteroplasmy and \( Y \) be the probability that any heteroplasmy in a mother will be transmitted to an offspring by maternal inheritance. In that case

\[
H_1 = Y H_0 + (1 - H_0) X + H_0 (1 - Y) X
\]

Eqn 1

At equilibrium, \( H_{eq} = H_0 = H_1 \) and simplifying yields

\[
H_{eq} = X / (1 - Y + XY)
\]

Eqn 2

From the combined results of the experimental crosses presented in Bentley *et al.* (2010) and the more recent unpublished data set, \( X = 0.027 \) (recognizing that six cases of paternal leakage resulted in pure paternal inheritance rather than heteroplasmy) and \( Y = 0.196 \). Substituting these values in Eqn 2 yields \( H_{eq} = 0.033 \), or that at equilibrium c. 3.3% of plants should be heteroplasmic.

Note the large difference between the proportion of heteroplasmic individuals found in natural populations by Pearl *et al.* (2009) (c. 15%) and the proportion predicted by direct observation of leakage and sorting in the crosses (c. 3.3%). The real difference may, in fact, be considerably greater since Pearl *et al.* (2009) point out that when the pollen donor is unknown in natural populations there are likely to be a considerable number of fertilization events in which there is leakage from a pollen donor who carries the same marker genotype as the pollen recipient. In such cases any leakage or resulting heteroplasmy could not be detected using those specific markers, in contrast to controlled crosses in which the two parents are known to differ in marker genotype.

Several explanations for the difference between the observed and predicted levels of heteroplasmy come to mind. First, the rates of gain and loss of heteroplasmy measured by the crosses are each calculated with some degree of statistical uncertainty that could propagate to even greater uncertainty in the prediction of the equilibrium level from Eqns 1 and 2. Secondly, the mechanics of artificial pollination and natural insect pollination might result in differing leakage rates. Thirdly, recall that offspring heteroplasmy was assayed at the seeding stage in the crosses, as opposed to samples from mature plants in natural populations. Further, heteroplasmic and homoplasmic plants were assumed to have equal fitness in Eqns 1 and 2. It could be that heteroplasm within individuals, or heteroplasmic individuals within populations, are maintained or amplified in some selective way owing to unknown fitness consequences of heteroplasmy. Certainly, to date the fitness consequences of heteroplasm in natural populations are understudied. Investigations of the relationship between heteroplasmy and fitness would need to ask whether the magnitude of heteroplasmy is a factor (it is treated as an all or none trait in the calculations above) and whether the manner in which it is distributed spatially within individuals need be considered.

IV. Conclusions

It appears that in plants maternal inheritance of the mitochondrial genome is the norm, with paternal leakage likely to be a relatively rare event in many of those species in which it does occur. When paternal leakage does occur, however, it may be important for the evolution of the mitochondrial genome both because the resulting heteroplasmy could facilitate creation of genotypic diversity via recombination, and because dispersal through pollen could alter population structure. However, the existence of paternal leakage might be a somewhat open question even in many of the species for which crossing data are at hand because of the fact that leakage rates on the order of a few per cent require quite large sample sizes for detection (Milligan, 1992). Further, because mitochondrial genes are generally highly conserved it could prove difficult to find the diagnostic marker polymorphisms needed to distinguish parents in crosses using species not yet studied. Indirect indicators of leakage via detection of the products of intra- or inter-genic recombination might be more easily applicable to a larger array of species, but still the issue of availability of necessary markers remains. However, recent advances in next-generation sequencing and genomics should facilitate the development of such markers (Rokas & Abbot, 2009) and allow the methods described in this review to be applied to a larger variety of plant species.

With an increase in the number of plant species in which the inheritance of the mitochondrial genome has been categorized, it will be possible for the community of plant evolutionary biologists to test three of the hypothesis outlined in this review – that leakage should be more common in those species whose natural populations harbor CMS, that it should be more common in those invasive species whose recent history promotes admixture, and that the population structure of mtDNA should be lower in those plant species in which mitochondrial transmission through pollen occurs.

Acknowledgements

I thank Michelle McCauley, Calvin Patimeeteporn, Sara Samoray, Samantha Skains, and Patricia Soria for laboratory assistance while gathering the unpublished data reported in this paper. That work was supported by Award DEB-1050331 from the US National Science Foundation. The manuscript benefited from the constructive comments of Lynda Delph, Dan Sloan, Mike Wade and two anonymous reviewers.

References


Sheahan MB, McCurdy DW, Rose RJ. 2005. Mitochondria as a connected population: ensuring continuity of the mitochondrial genome during plant


Supporting Information

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

**Table S1** Results of unpublished crosses used to assay paternal leakage of the mitochondrial gene *atp1* in *Silene vulgaris*

**Table S2** Primers used for PCR amplification of nine mitochondrial genes in *Silene vulgaris*

**Table S3** Sequences of portions of nine mitochondrial genes in which the SNPs described in Table 1 were found

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.