Horizontal Transfer of Entire Genomes via Mitochondrial Fusion in the Angiosperm *Amborella*

Danny W. Rice[1], Andrew J. Alverson[1,a], Aaron O. Richardson[1], Gregory J. Young[1,b], M. Virginia Sanchez-Puerta[1,c], Jérôme Munzinger[2,d], Kerrie Barry[3], Jeffrey L. Boore[3,e], Yan Zhang[4], Claude W. dePamphilis[4], Eric B. Knox[1], Jeffrey D. Palmer[1,f]

[1]Department of Biology, Indiana University, Bloomington, IN 47405, USA
[2]IRD UMR AMAP, Laboratoire de Botanique et d'Ecologie Végétale Appliquées, Nouméa, New Caledonia.
[3]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA
[4]Department of Biology, Penn State University, University Park, PA 16802, USA

[a]Present Address: Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA
[b]Present Address: DuPont Pioneer, Wilmington, Delaware 19880, USA
[c]Present Address: CONICET & Universidad Nacional de Cuyo, Mendoza, Argentina
[d]Present address: IRD UMR AMAP, TAA51/PS2, 34398 Montpellier Cedex 5, France
[e]Present Address: Genome Project Solutions, Hercules, CA 94547, USA
[f]Author for correspondence: jpalmer@indiana.edu

Abstract:
We report the complete mitochondrial genome sequence of the flowering plant *Amborella trichopoda*. This enormous, 3.9 Mb genome contains six genome equivalents of foreign mitochondrial DNA, acquired from green algae, mosses, and other angiosperms. Many of these horizontal transfers were large, including acquisition of entire mitochondrial genomes from three green algae and one moss. We propose a fusion-compatibility model to explain these findings, with *Amborella* capturing whole mitochondria from diverse eukaryotes, followed by mitochondrial fusion (limited mechanistically to green plant mitochondria), and then genome recombination. *Amborella*'s epiphyte load, propensity to produce suckers from wounds, and low rate of mitochondrial DNA loss probably all contribute to the high level of foreign DNA in its mitochondrial genome.

One-sentence summary: A mitochondrial genome with six genome equivalents of foreign mitochondrial DNA, including four whole-genome transfers, has rich implications for mechanisms of horizontal transfer in mitochondria.

Main text: Many of the fundamental properties of eukaryotes arose from horizontal evolution on a grand scale, i.e., the endosymbiotic origin of the mitochondrion and plastid from bacterial progenitors (*1*). Since their birth, however, mitochondrial and plastid genomes seem to have been little affected by horizontal gene transfer (HGT). The most notable exception involves land plants, especially flowering plants (angiosperms), in which HGT is surprisingly common in the mitochondrial genome but unknown in plastids (*2-10*).

To gain insight into the causes and consequences of HGT in mitochondrial DNA (mtDNA), we sequenced the mitochondrial genome of *Amborella trichopoda* because PCR-based sampling had shown it to be rich in foreign genes (*4*). This large shrub is endemic to rain forests of New Caledonia and is probably sister to all other angiosperms, a divergence dating back about 200 million years ago (*11,12*).

Overall genome properties. The *Amborella* mitochondrial genome assembled as five autonomous, circular-mapping chromosomes of lengths 3179, 244, 187, 137, and 119 kb, giving a total genome size of 3,866,039 bp (Figs. 1A, S1-4) (*13*). The five chromosomes are distinct in sequence, but similar in base composition (45-47% G+C), stoichiometry, and HGT properties (Figs. 1A, S2, S4). Stoichiometry was assessed by sequencing coverage and Southern blot analysis of 32 individuals from three populations (fig. S5) (*13*).

As described in the next three sections, *Amborella* mtDNA possesses an extraordinary assemblage of foreign sequences, corresponding to about six genome equivalents of mtDNA acquired from mosses, angiosperms, and green algae. Multi-gene HGT has been described in two other lineages of plant mtDNA (*8,10*), but not on a scale approaching *Amborella*. The *Amborella* mitochondrial genome also contains a large amount (138 kb) of plastid DNA (ptDNA) (Figs. 1A, S2; table S1).

Multichromosomal mitochondrial genomes in plants were only recently discovered (*14,15*) and mostly involve large (>1 Mb) genomes, with *Silene* genomes of 6.7 and 11.3 Mb dwarfing *Amborella* in size and chromosome number (*15*). These three mitochondrial genomes are the largest completely assembled organelle genomes, larger than many bacterial genomes and even some nuclear genomes. However, the processes responsible for their expansion differ in that *Silene* genomes possess no readily discernible foreign mtDNA and relatively little ptDNA (*15)*.

HGT from mosses. *Amborella* mtDNA contains four regions, of lengths 48, 40, 9, and 4 kb, acquired from moss mtDNA (Figs. 1A, S2). With one exception, the 41 protein and rRNA genes from these four regions were placed phylogenetically, almost always strongly, as sister to the moss *Physcomitrella* (Figs. 2A-D, S8 and S9). Gene order in the four regions (Figs. 3, S6) is highly similar to both *Physcomitrella* and *Anomodon* (mosses which are themselves identical in gene order and content) (*16*) and extremely different from angiosperms. The moss-like regions in *Amborella* also harbor the same 27 introns and largely the same set of intergenic sequences as moss mtDNAs (Fig. S6) (*13*).

The four moss regions contain one and only one copy of 61 of the 65 genes present in sequenced moss mtDNAs (Figs. 3, S6) (*13*). Taking into account six inferred deletions and duplications larger than 100 bp, the 101.8 kb of moss DNA in *Amborella* reconstructs to a hypothetical donor genome of 106.0 kb, compared to the 104.2- and 105.3-kb genomes in *Physcomitrella* and *Anomodon*, respectively. We infer, therefore, that *Amborella* captured an entire mitochondrial genome (*13*) from a moss with nearly identical mtDNA architecture to those of *Physcomitrella* and *Anomodon.* This foreign genome subsequently rearranged into four pieces, with a few gene-order changes and 11 gene losses, truncations, and/or partial duplications, all of which are associated with rearrangement breakpoints (Figs. 1A, 3, S2, S6, table S2).

HGT from green algae. The *Amborella* mitochondrial genome contains an average of three green-algal-derived copies of each protein and rRNA gene commonly found in green algal mtDNAs (Figs. 1A, 2A-D, S2, S4, S8, S10, S11; table S3). Many of these genes are clustered in two large tracts of lengths 83 and 61 kb. The 83-kb tract (B1+A2 in Fig. 1A) contains two copies of a 10-gene cluster (each marked by 10 red arrows in the top comparison of Fig. 4), with all 10 "duplicates" highly divergent from each other. The 61-kb tract (B2+A1 in Fig. 1A) lacks these 10 genes and instead contains highly divergent duplicates of two genes that are absent from the 83-kb tract. A single hypothesized recombination event between these two tracts (Figs. 1A, 4) accounts for the above duplications, with the initial, 92- and 52-kb regions each containing a nearly complete set of green - algal mitochondrial genes and no extra copies (fig. S11). We conclude that the 83-kb and 61-kb tracts arose by acquisition of whole mitochondrial genomes (designated the A and B genomes) from two green algae, followed by a single recombination between them and a few gene losses (*13*). Additionally, the two inferred donor genomes are phylogenetically distinct: Whenever *Amborella* has three or more green- algal copies of a given gene, the A-genome copy is separated by a relatively long branch from a well-supported clade containing the other green algal copies (Figs. 2A-D, S8).

Furthermore, the two regions assigned to the A genome have a lower non-coding G+C composition (39%) than the two B-genome regions (47%) (table S4).

Most of the remaining green algal mtDNA in *Amborella*, comprising tracts of lengths 49, 18, 16, and 2 kb (Figs. 1A, S2), also appears, on the basis of synteny and genome reconstruction (Figs. 4, S11), to be derived by whole genome transfer (from donor C). Seven of the eight remaining, mostly short tracts of green algal mtDNA (Fig. 1A) can tentatively be reconstructed as resulting from the transfer and/or retention of about one-third of a genome from a fourth green -algal donor (donor D); alternatively, the D regions may result from multiple HGT events. Although the B-D genomes are relatively similar in sequence (Figs. 2A-D and S8), their many differences in gene order (Figs. 1A, 4) and intron content (e.g., *cox1* has two introns in the D genome but none in B) rule out the possibility that they result from only one or two transfers followed by large-scale duplication within *Amborella*. We therefore conclude that *Amborella* acquired its ~3.3 genome equivalents of green algal mtDNA (Fig. 1A) via at least four transfers, including three whole-genome transfers.

The multiple copies of each green-algal gene present in *Amborella* almost always ally, usually strongly, with the trebouxiophyte *Coccomyxa* (Figs. 2A-D, S8). Likewise, gene order within the A-C genomes is most similar to that of *Coccomyxa* (fig. S7). The B, C, and D copies of each gene invariably form a strongly supported clade (Figs. 2A-D, S8, S10), with the B+C genomes sister to the A genome in gene-loss phylogeny (fig. S12). Thus, *Amborella* probably acquired its green algal mtDNA from the *Coccomyxa* subgroup of trebouxiophytes. Because members of this subgroup often live as lichen photobionts, and lichens commonly grow on *Amborella* (Fig. 5), its algal genomes may have been acquired from lichens.

HGT from angiosperms. *Amborella* mtDNA contains 150 angiosperm-like copies (full or partial) of the 49 protein and rRNA genes likely present in the ancestral angiosperm mitochondrial genome (fig. S13) (*17*) [see (*13*) for how *trans*-spliced genes are counted (table S5)]. We designated 82 of these copies as foreign, 63 as native, and 5 as uncertain (table S3). Angiosperm-specific phylogenetic analyses provided strong support for 26 (32%) of the foreign assignments and 16 (25%) of the natives (figs. S14-S16, table S6). These analyses were consistent, but with lower support, with an additional 20 foreign and 22 native assignments. These lower values reflect the generally poor resolution in many of the trees (fig. S14), which is a consequence of low substitution rates in most angiosperm mtDNAs (*18*).

Four other lines of evidence were used to distinguish foreign from native angiosperm genes and intergenic DNA. First, the extent of C-to-U RNA editing, which is much higher in *Amborella* than in all examined eudicots and monocots (table S7) (*13*), provided evidence for native vs. foreign origin for many of the 150 angiosperm genes in *Amborella* mtDNA (*13*). Second, six genes were exceptionally divergent relative to all other genes analyzed phylogenetically (fig. S15), suggesting that they came from angiosperms with much higher mtDNA substitution rates than *Amborella* (fig. S17) (*13,18*). Third, levels of sequence identity to other angiosperm mtDNAs were measured on a genome-wide basis to define native as well as angiosperm-HGT regions (*13*). Finally, native (or angiosperm-HGT) sequences defined by the above four criteria and located within 5 kb of each other were combined into continuous native (or angiosperm-HGT) tracts (*13*).

These analyses identified 753 kb of DNA as having been acquired from other angiosperms (Figs. 1A, S2, S4). This DNA contains an average of 2.0 copies of the 32 protein and rRNA genes that are virtually always present in angiosperm mtDNA (table S3) (*17*) and thus corresponds to roughly two genome equivalents of foreign angiosperm mtDNA. Most (86%) of the 753 kb is intergenic, consistent with the high proportion of intergenic mtDNA in angiosperms (*11,13*). About half of the 753 kb shares ≥90% sequence identity with one or more sequenced angiosperm mitochondrial genomes (fig. S4). This far surpasses the level of highly conserved mtDNA in other angiosperms (fig. S18) (*13*). The 753-kb estimate is probably conservative owing to the limited number of angiosperm mtDNAs available for comparison (*13*).

Angiosperm donors. One class of plastid-derived DNA played a key role in donor identification. Phylogenetic analysis shows that most of the 138 kb of ptDNA present in *Amborella* mtDNA was acquired via intracellular gene transfer (IGT), i.e., from the *Amborella* plastid genome (Figs. 2E-H, S19). Analysis of the remaining 10 kb of ptDNA, which probably entered *Amborella* via foreign mitochondria, identified donors with much greater specificity than did the mitochondrial gene analyses (*13*). Four of the HGT plastid regions identified Fagales, Oxalidales, or the predominantly parasitic Santalales as donor, while a fifth pointed to Magnoliidae (Figs. 2E-H, S18). A Santalalean origin is also supported by four of the five mitochondrial genes for which multiple Santalales have been sampled (figs. S14-nad1b, S20). The exceptionally high and specific similarity of two featureless regions to *Ricinus communis* or *Bambusa oldhamii* (Figs. 1B, S21) identified transfers from these lineages. Finally, the exceptionally high divergence that diagnosed six angiosperm-like genes as foreign also suggests that they came from additional donors, with high mitochondrial substitution rates.

Because some angiosperm HGT tracts in *Amborella* mtDNA are of mixed phylogenetic origin (Fig. 1) (*13*), some of its foreign DNA may be the product of serial, angiosperm-to-angiosperm-to-angiosperm HGT (*13*). In particular, the *rbcL* gene of santalalean origin (Fig. 2E) resides only 3 kb from the *Bambusa*-derived sequence on the same 27-kb foreign tract (Fig. 1B). Because all four genes of meaningful length on this tract evidently came from core eudicots (fig. S14), and because parasitic plants are especially active in mitochondrial HGT (*5,7-10*), this tract probably came from a santalalean donor that had previously acquired *Bambusa* DNA via HGT (*13*). The presence of santalalean DNA in six, mostly long HGT tracts (Fig. 1A) suggests that a large portion of the foreign angiosperm DNA in *Amborella* came from Santalales. Indeed, RNA-editing data indicate that the 27-kb tract of putative santalalean origin may actually be part of a much larger (>105 kb) HGT tract (*13*).

A graveyard of foreign genes. The 197 foreign mitochondrial protein genes in *Amborella* are predominantly pseudogenes, with only 50 (25%) of them having full-length, intact open reading frames (tables S2, S8). The intact genes are predominantly short (figs. S22, S23), suggesting that many of these have remained intact by chance, i.e., are pseudogenes that have yet to sustain an obvious pseudogene mutation. Consistent with this, many of these intact genes are not expressed properly.

On the basis of phylogenetic, RNA editing, and/or linkage evidence (table S9) (*13*), *Amborella* mtDNA is hypothesized to contain a functional, native copy of all but one (*rpl10*) of the 49 mitochondrial protein and rRNA genes inferred to be present in the ancestral angiosperm (fig. S13) (*17*). cDNA sequencing of 44 of the 48 native genes showed that, with one apparent exception, they are all transcribed and properly RNA edited (table S10) (*13*). In contrast, no transcripts were detected for many genes of foreign origin, and 13 of 14 transcribed genes of foreign angiosperm origin (eight of them intact) were poorly edited, suggesting that they are pseudogenes (table S10) (*13, 19*).

The strongest candidates for functional replacement of native genes are tRNA genes. Several native tRNA genes are missing from *Amborella* mtDNA (fig. S13). These, and even some of the native tRNA genes still present (*20*), may have been functionally replaced by some of its dozens of intact foreign tRNA genes (figs. S2, S4) (*13*). This would not be surprising, because cognate tRNAs of diverse origin (plastid, nuclear, bacterial) often replace native tRNAs in plant mitochondrial translation (*6,11,20,21*). Moreover, even a modest number of tRNA gene replacements could have led to the fixation, via genetic hitchhiking, of a considerable portion of the foreign mtDNA in *Amborella*.

In summary, the great majority of the foreign mitochondrial genes in *Amborella* are unlikely to be functional. Given its six genomes-worth of foreign mitochondrial genes, *Amborella* mtDNA serves as a striking example of neutral evolution.

Ancient transfers, remarkably intact. Our ability to date the many mitochondrial HGTs in *Amborella* is limited. However, the extensive pseudogene decay of its foreign DNA (tables S2, S8) –

in conjunction with low mitochondrial substitution rates in angiosperms (including *Amborella*; fig. S17) (*18*) and low rates of pseudogene decay (*19*) – suggests that most transfers are probably millions of years old (*13*).

Angiosperm mitochondrial genomes typically experience high rates of DNA gain, loss, and rearrangement (*13,17*). *Amborella* mtDNA seems, however, less prone to lose and rearrange DNA. Relative to their many pseudogene mutations, the four moss and green-algal whole-genome transfers are surprisingly intact with respect to overall sequence content and arrangement. Only 11% of the protein-coding sequence content inferred to be present at the time of these four transfers has been deleted, mostly due to a few single- or multi-gene deletions (Figs. 4, S6; tables S2, S8) (*13*). The green algal A and B genomes are both intact syntenically except for a single, mutual recombination event, while the C and moss genomes have each been fragmented into just four segments (Figs. 1, 3, 4). In typical angiosperm mtDNAs, comparably old and large tracts of largely nonfunctional DNA would be expected to have mostly been lost by now, and what remained to be more highly rearranged (*13,17*).

Mitochondrial fusion drives and limits mitochondrial HGT. Two mechanisms have been proposed to account for the relatively high frequency of HGT in land plant mitochondria and its absence from plastids of land plants, including *Amborella* (*6,8,9*). First, plant mitochondria are transformation competent (*22*), whereas no such evidence has been reported for plastids. Second, plant mitochondria regularly fuse *in vivo*, whereas plastids do not (*23,24*). Three aspects of the horizontally acquired DNA in *Amborella* argue that its entry into the mitochondrion was driven principally, if not entirely, by mitochondrial fusion − i.e., this DNA entered predominantly in large pieces, including whole genomes (*13*), is limited to other mitochondrial genomes (*13*), and is limited to green algae and land plants.

Why are the many *Amborella* donors limited to green plants, as opposed to, for instance, fungi, given their pervasive interactions with plants as mycorrhizal partners, endophytes, epiphytes, and pathogens? We propose that this reflects a phylogenetically-deep incompatibility in the mechanism of mitochondrial fusion. The mechanism of mitochondrial fusion in fungi and animals is fundamentally the same, involving a core machinery of dynamin-related GTPases that are absent from green plants (*25-27*). This absence, combined with evidence for differences in the physiological requirements for fusion, has prompted speculation that mitochondrial fusion occurs by a different mechanism in angiosperms than in animals and fungi (*24, 27, 28*). Our data provide evolutionary support for this hypothesis and also lead us to propose that mitochondrial fusion occurs in a fundamentally similar manner across land plants and green algae (Fig. 6). This model explains why, despite presumably broad phylogenetic exposure to foreign mitochondria, the vast majority of HGT in the mitochondrion of *Amborella* – and other plants (*2-10,13*) – is restricted to other plant mitochondria.

Capture of foreign mitochondria. Biological vectors large enough to mobilize entire mitochondria, such as pollen (*9,29*), insects, and fungi, could account for some of the mitochondrial HGT in Amborella (bacteria and viruses are presumably too small to transfer an entire mitochondrion). However, in light of its ecology and development, processes involving direct contact between *Amborella* and potential donors probably predominate. *Amborella* grows in montane rainforests, often covered by a diversity of epiphytes, mostly bryophytes (including mosses) and lichens (a potential source of its green algal genomes), and sometimes even other angiosperms (Fig. 5). *Amborella* is often wounded and responds by producing abundant suckers (Figs. 5A-B). Wounding can break cells belonging to both *Amborella* and the organisms growing on and within it. We postulate that some of the broken *Amborella* cells are healed and incorporated into a new meristem – a new germline arising thanks to the totipotency of plant cells. Indeed, plant meristems often form in direct response to wounding and may be especially active in “massive mitochondrial fusion” (*24*). Given the ease of both mitochondrial membrane fusion and mitochondrial genome recombination, those healed cells that have taken up a mitochondrion from another green plant could well incorporate a portion of the foreign mitochondrial genome. A fraction of these transfers could then become fixed.

The wounding-HGT model applies not only to plants that live on *Amborella*, but also parasites. The Santalales – probably the major source of foreign angiosperm mtDNA in *Amborella* – are also the major group of parasitic plants in New Caledonia and the largest group of parasitic angiosperms worldwide (*30,31*).

Concluding Remarks. The *Amborella* mitochondrial genome has both captured other mitochondrial genomes whole and also retained them in remarkably intact form for ages. Its assemblage of foreign mtDNA probably reflects a range of factors – ecological, developmental, and molecular – that promote the capture of foreign mtDNA and retard its loss and rearrangement. This genome highlights the potential scale of neutral evolution and is thus relevant to current debates on the issue of “junk DNA” in nuclear genomes (*32*). The greatest significance of this genome is mechanistic: It provides compelling support for mitochondrial fusion as the key that unlocks mitochondrial HGT and for fusion incompatibility as a major barrier to phylogenetically unconstrained mitochondrial “sex”.

References and Notes

1. J. Archibald, Origin of eukaryotic cells: 40 years on. *Symbiosis*, **54**, 69-86 (2011).
2. U. Bergthorsson, K .L. Adams, B. Thomason, J. D. Palmer, Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* **424**, 197-201 (2003).
3. H. Won, S. S. Renner, Horizontal gene transfer from flowering plants to *Gnetum*. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10824-10829 (2003).
4. U. Bergthorsson, U., A. O. Richardson, G. J. Young, L. R. Goertzen, J. D. Palmer, Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17747-17752 (2004).
5. C. C. Davis, K. J. Wurdack, Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. *Science* **305**, 676–678 (2004).
6. D. W. Rice, J. D. Palmer, An exceptional horizontal gene transfer in plastids: Gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. *BMC Biology* **4**, 31 doi:10.1186/1741-7007-4-31 (2006).
7. T. J. Barkman et al., Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evol Biol* **7**, 248 (2007).
8. J. P. Mower *et al.*, Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. *BMC Biol.* **8**, e150 (2010).
9. J. P. Mower, K. Jain K, N. J. Hepburn, The role of horizontal transfer in shaping the plant mitochondrial genome. In L. Maréchal-Drouard, Ed., Advances in Botanical Research Volume 63: Mitochondrial Genome Evolution. Elsevier, pp. 41-69 (2012).
10. Z. Xi *et al.*, Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLOS Genetics* **9**, e1003265 (2013).
11. S. A. Smith, J. M. Beaulieb, M. J. Donoghue, An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5897-5902 (2010).
12. D. E. Soltis *et al.,* Angiosperm phylogeny: 17 genes, 640 taxa. *Amer. J. Bot.* 98, 704-730 (2011).
13. Supplementary materials for this article are available on *Science* Online.

14. A. J. Alverson, D. W. Rice, S. Dickinson, K. Barry, J. D. Palmer, Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber (*Cucumis sativus*). *Plant Cell* **23**, 2499-2513 (2011).
15. D. B. Sloan *et al.*, Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* **10**, e1001241 (2012).
16. Y. Liu, J.-Y. Xue, B. Wang, L. Li, Y.-L. Qiu, The mitochondrial genomes of the early land plants *Treubia lacunose* and *Anomodon rugelii*: Dynamic and conservative evolution. *PLoS ONE* **6**, e25836, (2011).
17. J. P. Mower, D. B., Sloan, A J. Alverson, Plant mitochondrial genome diversity: the genomics revolution. In J. F. Wendel, J. Greilhuber, J. Dolezel, I. J. Leitch, Eds., Plant genome diversity Vol. 1. Plant genomes, their residents, and their evolutionary dynamics. Vienna: Springer, pp. 123-144 (2012).
18. J. P. Mower, P. Touzet, J. S. Gummow, L. F. Delph, J. D. Palmer, Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evol. Biol.* **7**, e135 (2007).
19. H. C. Ong, J. D. Palmer, Pervasive survival of expressed mitochondrial *rps14* pseudogenes in grasses and their relatives for 80 million years following three functional transfers to the nucleus. *BMC Evol. Biol.* **6**, e55 (2006).

20. K. Kitazaki et al., A horizontally transferred tRNA^{Cys} gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. *Plant J.* **68**, 262-272 (2011).
21. L. Maréchal-Drouard, *et al*. Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. *Nucleic Acids Res.* **18**, 3689-3696 (1990).
22. D. Mileshina, M. Koulintchenko, Y. Konstantinov, A. Dietrich, Transfection of plant mitochondria and in organello gene integration. *Nucleic Acids Res.* **39**, e115 (2011
23. S. Arimura, J. Yamamoto, G. P. Aida, M. Nakazono, N. Tsutsumi, Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7805-7808 (2004).
24. M. B. Sheahan, D. W. McCurdy, R. J. Rose, Mitochondria as a connected population: ensuring continuity of the mitochondrial genome during plant cell dedifferentiation through massive mitochondrial fusion. *Plant J* **44**, 744-755 (2005).
25. B. Westermann, Mitochondrial fusion and fission in cell life and death. *Nature Rev. Mol. Cell Biol.* **11**, 872-884 (2010).
26. T. Kuroiwa *et al.*, Structure, function and evolution of the mitochondrial division apparatus. *Biochim. Biophys. Acta* **1763**, 510-521 (2006).
27. H. Gao, T. L. Sage, K. W. Osteryoung, FZL, an FZO-like protein in plants, is a determinant of thylakoid and chloroplast morphology. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6759-6764 (2006).
28. K. Wakamatsu, M. Fujimoto, M. Nakazono, S. Arimura, N. Tsutsumi, Fusion of mitochondria in tobacco suspension cultured cells is dependent on the cellular ATP level but not on actin polymerization. *Plant Cell Rep*. **29**, 1139-1145 (2010).
29. P.-A. Christin *et al.*, Adaptive evolution of C4 photosynthesis through recurrent lateral gene transfer. *Curr. Biol.* **22**, 445-449 (2012).
30. D. L. Nickrent, V. Malécot, Romina. Vidal-Russell, J. P. Der, A revised classification of Santalales. *Taxon* **59**, 538-558 (2010).
31. P. Morat *et al.*, The taxonomic database « FLORICAL » and characteristics of the indigenous flora of New Caledonia. *Adansonia sér. 3,* **34**, 177-219 (2012).
32. W. F. Doolittle, Is junk DNA bunk? A critique of ENCODE. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5294-5300 (2013).
33. V. Poncet *et al.,* Phylogeography and niche modelling of the relict plant *Amborella trichopoda* (Amborellaceae) reveal multiple Pleistocene refugia in New Caledonia. *Mol. Ecol.*, in press (2013?).
34. *Amborella* Genome Project, The *Amborella* genome and the evolution of flowering plants. *Science*, in press? (2013?).
35. A. Zuccolo *et al.*, A physical map for the *Amborella trichopoda* genome sheds light on the evolution of angiosperm genome structure. *Genome Biology* **12**, eR48, http://genomebiology.com/2011/12/5/R48 (2011).
36. J. J. Doyle, J. L. Doyle, A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15 (1987).
37. R. Kolodner, K. K. Tewari, Physicochemical characterization of mitochondrial DNA from pea leaves. *Proc Natl Acad Sci U.S.A.* **69**, 1830-1834 (1972).
38. M. Bahloul, G. Burkard, An improved method for the isolation of total RNA from spruce tissues. *Plant Molec. Biol. Reporter* **11**, 212-215 (1993).
39. X. Huang, A. Madan, CAP3: A DNA sequence assembly program. *Genome Res.* **9**, 868-877 (1999).
40. D. R. Smith *et al.*, The GC-rich mitochondrial and plastid genomes of the green alga Coccomyxa give insight into the evolution of organelle DNA nucleotide landscape. *PLoS ONE* **6**, e23624, (2011).

41. Rodríguez-Moreno et al., Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics* **11**, e424 (2011).
42. B. L. Ward, R. S. Anderson, A. J. Bendich, The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). Cell 26, 793-803 (1981).
43. Liu *et al*., The Complete Mitochondrial Genome of *Gossypium hirsutum* and Evolutionary Analysis of Higher Plant Mitochondrial Genomes. PLoS One. 8, e69476. 10.1371/journal.pone.0069476 (2013).
44. A. O. Richardson, D. W. Rice, G. J. Young, A. J. Alverson, J. D. Palmer, The "fossilized" mitochondrial genome of *Liriodendron tulipifera*: Ancestral gene content and order, ancestral editing sites, and extremely low mutation rate. *BMC Biol.* 11, e29 (2013).
45. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).
46. J. P. Mower, Modeling sites of RNA editing as a fifth nucleotide state reveals progressive loss of edited sites from angiosperm mitochondria. *Molec. Biol. Evol*. **25**, 52-61 (2008).
47. D. C. Shields, K. H. Wolfe, Accelerated evolution of sites undergoing mRNA editing in plant mitochondria and chloroplasts. Molec. Biol. Evol. 14, 344-349 (1997).
48. H. Shimodaira, M. Hasegawa, Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114-1116 (1999).
49. W. Martin, Molecular evolution - Lateral gene transfer and other possibilities. *Heredity* **94**, 565-566 (2005).
50. V. V. Goremykin, F. Salamini, R. Velasco, R. Viola, Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. *Molecular Biology and Evolution* **26**, 99-110 (2009).
51. Y. Cho, J. P. Mower, J. D. Palmer, Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proc. Natl. Acad. Sci. USA* **101**, 17741-17746 (2004).
52. C. L. Parkinson et al., Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. *BMC Evol. Biol.* **5**, 73 (2005).
53. D. B. Sloan, B. Oxelman, A. Rautenberg, and D. R. Taylor. A phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe *Sileneae* (Caryophyllaceae). *BMC Evolutionary Biology* **9**, 260 (2009)
54. J. O. Allen et al., Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* **177**, 1173-1192 (2007).
55. D. B. Sloan, K. Müller, D. E. McCauley, D. R. Taylor DR, H. Štorchová H, Intraspecific variation in mitochondrial genome sequence, structure, and gene content in *Silene vulgaris*, an angiosperm with pervasive cytoplasmic male sterility. *New Phytologist*. **196**, 1228-1239 (2012).
56. A. Darracq et al., Structural and content diversity of mitochondrial genome in beet: A Comparative Genomic Analysis. *Genome Biol. Evol.* **3**, 723-736 (2011).
57. S. W. Clifton et al., Sequence and comparative analysis of the maize NB mitochondrial Genome. *Plant Physiology* **136**, 3486-3503 (2004).
58. G. Drouin, H. Daoud, J. Xia, Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol*. **49**, 827-831 (2008).
59. J. P. Mower and J. D. Palmer, Patterns of partial RNA editing in mitochondrial genes of *Beta vulgaris*. *Mol. Gen. Genomics* 276, 285-296 (2004).
60. J. D. Palmer, L. A. Herbon, Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol* **28**, 87-97 1988).
61. P. Grandcolas *et al*., New Caledonia: a very old Darwinian island? *Phil. Trans. R. Soc. B* **363**, 3309-3317 (2008).

62. V. V. Goremykin, P. J. Lockhart, R. Viola, R. Velasco, The mitochondrial genome of *Malus domestica* and the import-driven hypothesis of mitochondrial genome expansion in seed plants. *Plant J.* **71**, 615-626 (2012).
63. R. Velasco *et al.*, The genome of the domesticated apple (*Malus* × *domestica* Borkh.). *Nature Genet.* **42**, 833–839 (2010).
64. Notsu *et al.*, The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol. Genet. Genomics* **268**, 434-445 (2002).
65. A. J. Alverson *et al.*, Insights into the evolution of plant mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol. Biol. Evol.* **27**, 1436-1448 (2010).
66. C. Férandon *et al.*, The *Agaricus bisporus cox1* gene: The longest mitochondrial gene and the largest reservoir of mitochondrial group I introns. *PLoS ONE* **5**, e14048 (2010).
67. L.-F. Yin *et al.*, Frequent gain and loss of introns in fungal cytochrome b genes. *PLoS ONE* **7**, e49096 (2012).
68. Y. Cho, Y.L., P. Kuhlman, and J. D. Palmer, Explosive invasion of plant mitochondria by a group I intron. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14244-14249 (1998).
69. M. V. Sanchez-Puerta, Y. Cho, J. P. Mower, A. J. Alverson, and J. D. Palmer, Frequent, phylogenetically local horizontal transfer of the *cox1* group I intron in flowering plant mitochondria. *Mol. Biol. Evol.* **25**, 1762-1777 (2008).
70. V. V. Goremykin, K. I. Hirsch-Ernst, S. Wolfl, F. H. Hellwig, Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Mol. Biol. Evol.* **20**, 1499-1505 (2003).
71. Zhang *et al.*, The complete chloroplast and mitochondrial genome sequences of *Boea hygrometrica*: Insights into the evolution of plant organellar genomes. *PLoS One*, **7**, e30531 (2012).
72. Iorizzo *et al.*, De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biol.* **12,** 61 (2012).
73. Y. Fang *et al.*, A complete sequence and transcriptomic analyses of date palm (Phoenix dactylifera L.) mitochondrial genome. *PLoS One* **7**, e37164 (2012).
74. K. L. Adams, Y.-L. Qiu, M. Stoutemyer, J. D. Palmer, Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer during angiosperm evolution. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9905-9912 (2002).
75. Y.-L. Qiu, J. D. Palmer, Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J. Mol. Evol.* 59, 80-89 (2004).
76. H. Bock, A. Brennicke, W. Schuster W, *rps3* and *rpl16* genes do not overlap in Oenothera mitochondria: GTG as a potential translation initiation codon in plant mitochondria. *Plant Mol Biol.* **24**, 811–818 (1994).

Acknowledgments: We thank E. Dalin, J. Gummow, and Pete Lowry for assistance; R. Wing and the Arizona Genomics Institute for *Amborella* BAC sequences; the North and South environmental services of New Caledonia for collecting permits; M. Moore, P. Soltis, and D. Soltis for two unpublished plastid-genome sequences; and those individuals (see table S11) who supplied the photographs for figures. This work was supported by NIH-RO1-GM-76012 (J.D.P., E.B.K), the DOE-JGI Community Sequencing Program under Contract DE-AC02-05CH11231 (J.D.P, E.B.K, J.L.B), NSF-GRF-112955 (A.O.R.), NSF-DBI-0638595 (C.W.D), and the METACyt Initiative of Indiana University, funded by the Lilly Endowment. The data reported in this paper are deposited in GenBank under accessions KF754799-KF754803 and KF798319-KF798355.

Fig. 1. Foreign DNA in the *Amborella* mitochondrial genome. (**A**) Map of its five chromosomes shown linearized and abutted (see arrows). Numbers give unified genome coordinates in kb. Shown are regions of inferred organelle origin whose ancestry was assignable (see key, mt = mitochondrial; pt=plastid). Full-height boxes indicate genes. Half-height boxes indicate “tracts” (see text) of native and angiosperm-HGT DNA. Labeled black lines indicate horizontally transferred mitochondrial genomes (Figs. 3, 4) or partial genomes. M1-M4 mark a moss-derived genome. A1, A2, B1, B2, and C1-C4 mark three green-algal-derived genomes. D marks the seven fragments of a partial genome from a fourth green-algal donor. Oxalidales, Santalales, Fagales, and *Ricinus* mark angiosperm tracts whose donors were identified to at least order. The pie chart depicts the roughly eight genome equivalents of organelle DNA present in *Amborella* mtDNA. Genome equivalents: mt moss, 1.0; mt green algal, 3.2; mt angiosperm, 2.0; mt native, 1.0; pt IGT, 0.8. See (*13*) and table S11 for all plant images, including their relevance. (**B**) Detailed view of three 150-kb regions of *Amborella* mtDNA. Histograms show the “angiosperm score” (*13*). Triangles indicate intergenic regions of species-specific identity to *Ricinus* and *Bambusa* (fig. S21). Gene names are given only for well-supported cases of angiosperm HGT.

Fig. 2. Maximum likelihood evidence for HGT in *Amborella* mtDNA. (**A-D**) Mitochondrial gene trees of land plants and green algae reveal diverse donors in *Amborella* mtDNA. Colors are as in Fig. 1. See fig. S8 for outgroups. Bootstrap values ≥50% are shown. Scale bars correspond to 0.1 (**A-D**) or 0.01 (**E-H**) substitutions/site. Bold branches are reduced in length by 50%. (**E-H)** Plastid gene trees of angiosperms showing strong support for HGT to the level of order: light blue, Santalales (**E-F**); brown, Oxalidales (**G**); violet, Fagales (**H**). *Amborella* labels: “Amb plastid”, gene in *Amborella* plastid; “Amb IGT”; gene in mitochondrion via IGT; red “Amb”, gene in mitochondrion via HGT. Outgroups are not shown, but see fig. S19 for more taxon-rich analyses, including outgroups. “rps7” denotes the *rps7-rps12-trnV-rrnS* cluster.

Fig. 3. A nearly full-length moss mitochondrial genome in *Amborella* mtDNA. Colored boxes and arrows indicate the position and relative orientation, respectively, of the seven blocks of synteny between the mitochondrial genome of the moss *Anomodon* (top) and the four moss-derived regions in *Amborella* mtDNA (M1-M4; Fig. 1A). Selected genes are shown; see figs. S2 and S6 for all genes.

Fig. 4. Pairwise comparisons of the green-algal B-genome donor to *Amborella* with the A- and C-genome donors. Brackets on the A-C genomes indicate their fragmentation in *Amborella* (Fig. 1A). Blocks of two or more genes with identical order in a comparison are colored the same, regardless of gene orientation. Open boxes mark genes present in both genomes but not part of a syntenic block. Bullets mark genes present in only one genome.

Fig. 5. Ecological setting of HGT in *Amborella*. (**A, B**) Prostrate branches of *Amborella* with suckers (green arrows) and epiphytes, including mosses, liverworts, ferns, and angiosperms. (**C–F**) *Amborella* leaves and branches covered predominantly with lichens (**C, F**), leafy liverworts (**D**), and mosses (**E**). See table S11 for photo credits.

Fig. 6. Evolutionary model of mitochondrial fusion compatibilities. Green and orange indicate different mechanisms of mitochondrial fusion (*24, 25, 27, 28*), due to either highly divergent evolution from a common ancestral mechanism or independent origins of fusion. See table S11 for photo credits.