Rare gene capture in predominantly androgenetic species

Shannon M. Hedtke*,†, Matthias Gaubrecht‡, and David M. Hillis*†

*Section of Integrative Biology and Center for Computational Biology and Bioinformatics, University of Texas at Austin, Austin, TX 78712; and ‡Museum für Naturkunde Berlin, Leibniz Institute for Research in Evolution and Biodiversity at the Humboldt University Berlin, D-10115 Berlin, Germany

Contributed by David M. Hillis, April 27, 2011 (sent for review February 28, 2011)

The long-term persistence of completely asexual species is unexpected. Although asexuality has short-term evolutionary advantages, a lack of genetic recombination leads to the accumulation over time of deleterious mutations. The loss of individual fitness as a result of accumulated deleterious mutations is expected to lead to reduced population fitness and possible lineage extinction. Persistent lineages of asexual, all-female clones (parthenogenetic and gynogenetic species) avoid the negative effects of asexual reproduction through the production of rare males, or otherwise exhibit some degree of genetic recombination. Another form of asexuality, known as androgenesis, results in offspring that are clones of the male parent. Several species of the Asian clam genus Corbicula reproduce via androgenesis. We compared gene trees of mitochondrial and nuclear loci from multiple sexual and androgenetic species across the global distribution of Corbicula to test the hypothesis of long-term donality of the androgenetic species. Our results indicate that low levels of genetic capture of maternal nuclear DNA from other species occur within otherwise androgenetic lineages of Corbicula. The rare capture of genetic material from other species may allow androgenetic lineages of Corbicula to mitigate the effects of deleterious mutation accumulation and increase potentially adaptive variation. Models comparing the relative advantages and disadvantages of sexual and asexual reproduction should consider the possibility of rare genetic recombination, because such events seem to be nearly ubiquitous among otherwise asexual species.

Muller's ratchet | Meselson effect | Hill-Robertson effects | phylogenetics

Although asexual reproduction is widely distributed across eukaryotes, obligate asexuality is comparatively rare, and obligately asexual species tend to be of recent origin (1–4). Because of this observed rarity, there are presumably costs to obligate asexuality that allow sexual lineages to outcompete their asexual relatives, even though sexual reproduction itself carries costs [particularly the “two-fold cost of sex”: the cost of producing males or the cost of meiosis (2, 3)]. The literature on genetic (or mutational) costs to asexuality is extensive, but the main ideas can be broadly summarized as (i) asexuals can only rely on mutation to generate adaptive variation (e.g., refs. 1 and 5); and (ii) deleterious mutations accumulate faster in asexuals than sexuals (e.g., refs. 6–9). When more than one asexual lineage is observed in a closely related group of organisms, some process must either allow for diversification despite the costs, or else repeatedly generate new asexual lineages. Understanding mechanisms that drive asexual diversity can reveal much about the costs and benefits of asexual reproduction.

Androgenesis is a form of asexual reproduction in which offspring are clones of the father (10, 11). Several species of Asian clams in the genus Corbicula produce paternal clones through androgenesis (12–14). Androgenetic Corbicula are hermaphrodites and fertilize both their own eggs and eggs from closely related congeners to make paternal clones that do not incorporate maternal nuclear chromosomes (15, 16). Could the interaction between gametes of closely related Corbicula species allow rare sex between lineages, changing our expectations for the persistence of androgenetic species, and thus the time available to generate asexual diversity? Or is the observed diversity a result of recent, replicate origins [as in ostracods, stick insects, and water fleas (17–19)]?

Many asexual lineages have mechanisms for rare genetic capture (20–24). In parthenogenetic lineages, a male may be produced after several generations of female clonal reproduction, and sex could occur between the parthenogens and the rare male (25). In several pathogenic fungi that seem to be asexual, DNA sequencing revealed the presence of functional genes for mating and meiosis, suggesting that cryptic or unobserved sex could be occurring (26). In one type of asexual reproduction, gynogenesis, females require sperm from males of closely related sexual species to activate embryogenesis, but males do not regularly contribute genes to offspring. For example, even though behaviors associated with sexual reproduction are inherited from a sexual ancestor, Amazon mollies (Poecilia formosa) are considered asexual because daughters are usually genetically identical to their mothers. However, gametic interactions between sexuals and asexuels lead to rare genetic capture: microchromosomes are passed on from host males, and these microchromosomes contain genes with effects on phenotype (27, 28). Some of these microchromosomes are inherited by subsequent generations of female clones (27, 29). Rarely, entire chromosomal haploid sperm are integrated with the diploid maternal genome, causing triploid, gynogenetic offspring (30), although triploidy is unstable and individuals can end up with both diploid and triploid somatic cells (31). Thus, gynogenetic lineages that gain genetic material from males may persist for longer periods of time than otherwise expected (32).

The ancient asexual bdelloid rotifers are often considered an exception to the rule of persistent asexuality. No males have been observed in either contemporary or fossil populations going back at least 35 million years (33), and genomes of many bdelloids contain highly divergent alleles (34)—some to the point where two alleles may have different functions (35). However, during cycles of desiccation, the genome is fragmented. When the rotifer is rehydrated, DNA from the environment may be incorporated into the genome during chromosome repair, resulting in genetic recombination (36). Between morphologically and genetically distinct species, some nearly identical alleles are shared, even though these species seem to have been reproducing asexually for many millions of years (37).

In androgenetic Corbicula, genetic recombination could occur when gametes interact during reproduction. After fertilization, maternal nuclear chromosomes are extruded from the egg cell;
only paternal nuclear chromosomes from the sperm remain in the zygote (12–14). Mitochondria, however, are maternally inherited, so offspring possess the maternal mitochondrial genome and the paternal nuclear genome. Phylogenies built from mitochondrial vs. nuclear markers can be compared to identify between-species genetic recombination (10). Well-supported phylogenetic incongruence between mitochondrial and nuclear gene trees would be a signal of egg-capture (and thus mitochondrial genome-capture) in the genus (10, 15, 16). Nuclear gene-capture by one (paternal) species would be indicated if nuclear alleles found within a single individual were highly divergent and also had different evolutionary histories (as seen after sexual hybridization between species). We used these assumptions to compare a mitochondrial phylogeny of *Corbicula* (from ref. 16) to two phylogenies built using single-copy nuclear introns. We find evidence that multiple androgenetic lineages have captured mitochondrial and nuclear DNA from other species and suggest that rare sex could generate genetic diversity and allow androgenetic lineages to persist for longer than expected.

**Results**

We estimated phylogenies for two single-locus nuclear markers: the third intron of the α-amylase gene (*amy*; Fig. 1) and a putative intron of the α subunit of adenosine triphosphate synthase (*atps-α*; Fig. 2). When sexual individuals have more than one allele at a locus, those alleles are closely related to each other. Some asexual individuals, however, contain very divergent alleles at each locus. Hypothesis testing indicates that, in both nuclear trees, alleles within a single androgenetic individual are not always monophyletic; even when other androgenetic alleles are permitted to be nested within that individual’s clade. Neither are asexual lineages as a whole monophyletic (Figs. 1 and 2 and Table S1). However, in both gene trees, there is one allele that appears in most of the asexual populations.

Three phylogenies, *amy*, *atps-α*, and the mitochondrial cytochrome oxidase I (*cox-I*) data from Hedtke et al. (16), were tested to find whether observed differences in topology were a result of statistically significant differences in phylogenetic signal among the data sets (Table S2). For the six taxa sequenced in all three data sets, we tested support for each bipartition found in any of the three trees by comparing the posterior probabilities between data sets. To test whether there was a single taxon responsible for gene tree incongruence, we deleted each taxon and determined posterior probabilities for each reduced backbone tree. To test whether there was a single taxon responsible for gene tree incongruence among the data sets. The clades with posterior probabilities (>0.5 in any data set are found in Table S1. These results indicate well-supported incongruence in relationships of taxa among gene trees.

**Discussion**

Androgenetic clams in the genus *Corbicula* are morphologically and genetically distinct (e.g., 11, 13, 37–42). Several hypotheses could explain this observed androgenetic species diversity, given that androgenetic species do not share a recent common ancestor on a mitochondrial gene tree (15, 16, 41) (Fig. 3). First, androgenesis might have multiple origins as a result of repeated loss-of-function mutations or hybridization events between sexual species. Second, androgenesis might have a single origin; divergent mitochondrial lineages within asexuals might be the result of egg capture by an androgenetic sperm with complete maternal nuclear chromosome extrusion. Third, androgenesis might have a single origin, and egg capture by androgenetic species might have been rarely accompanied not only by capture of maternal mitochondrial, but also nuclear, DNA. Last, combinations of these processes may have acted in this system to form morphologically and genetically distinct androgenetic species.

**Potential Sources of Phylogenetic Incongruence.** Each of the three individual gene trees has statistically supported incongruent relationships between species (Figs. 1 and 2 and Table S1). Gene tree incongruence can be caused by systematic error in the analysis. For example, one gene may be prone to long-branch attraction, whereas another is not because of differences in the rate of evolution (43). Of the three gene trees considered, the only obvious topology for which long-branch attraction might have been a problem was *cox-I* (16). When the outgroups to *Corbicula* (*Neocorbicula* and *Polymesoda*) and to freshwater *Corbicula* (*Corbicula cf japonica*) were removed from analysis,
relationships within freshwater *Corbicula* did not change significantly. This suggests that incongruence between mitochondrial and nuclear gene trees was not due to long-branch attraction between any freshwater taxon and the long-branched outgroups.

There are biological processes that can lead to topological incongruence between gene trees. Given the short branch lengths inferred between taxa, freshwater *Corbicula* seem to have recently radiated, with rapid morphological evolution (41), making incomplete lineage-sorting between sexual taxa a plausible source of incongruence. The analogous process in asexuals would be retention of ancestral polymorphism (via heterozygosity of the newly asexual clams) and would result in divergent alleles within individuals. However, because some identical alleles are held in common across different species of asexual clam, for ancestral polymorphism to adequately explain the observed topologies, androgenesis would need to have arisen multiple times within an ancestral population, and such would need to be a mechanism that would slow evolution of the shared allele but not of other alleles at that locus. Gene duplication and subsequent loss also cannot be rejected as possible processes that could confuse topological inference, although to explain the observed topologies, duplication would have occurred early in the history of the genus, and different copies independently lost in almost all species.

Topological incongruence between gene trees can be caused by hybridization and capture of mitochondria. Given known mechanisms of androgenesis in *Corbicula* (12–14), they are highly likely to have caused discordance between mitochondrial and nuclear gene trees. These gene tree incongruences, rather than obscuring origins of androgenetic species diversity, can, instead, reveal processes important to the evolution of androgenetic *Corbicula*.

**Origins of Androgenesis.** There are multiple morphologically distinct species of *Corbicula* identified as androgenetic through cytological examination of fertilization (12–14) or the presence of genetic invariance, polyploidy, and biflagellate sperm (15, 38–42). Our phylogenetic trees do not support a simple scenario of repeated, multiple origins of androgenesis, which would require congruence between nuclear and mitochondrial gene trees (Fig. 3A). Nuclear alleles across androgenetic taxa are not monophyletic, which counters expectations under a single origin and subsequent diversification of androgenetic taxa (Fig. 3B). Instead, the observed phylogenetic pattern is what we expect under a relatively recent origin of androgenesis, followed by postorigin nuclear hybridization as asexual taxa spread and came into contact with different sexual species (Fig. 3C).

Most androgenetic individuals share a single allele in common (e.g., “ABDE” in Fig. 4A) or have alleles found in clades of closely related sequences (Fig. 4B). This suggests that at least a portion of the genome of all androgenetic species shares a common history that excludes sexual taxa. Because these alleles are identical or very similar, the shared history seems to be relatively recent; alleles have not had time to diverge since their common ancestor. This suggests a relatively recent origin of androgenesis in the genus, likely from an ancestor of *Corbicula sandai*, the sexual species found most closely related to androgenetic taxa across phylogenies.

Some individual asexuals contain a highly divergent allele (e.g., “D” in Fig. 4A or “E” in Fig. 4B). The “Meselson effect,” in which alleles within an asexual lineage diverge as they are...
Evolutionary Consequences of Androgenesis. When obligate androgenesis arises, under most conditions androgenetic individuals are expected to have a reproductive advantage over sexual members of the same population, and obligate androgenesis will spread to fixation (50). In addition to the direct reproductive benefits of egg capture, androgenesis could benefit from infrequent chromosomal rescue. If harmful mutations are recessive, they could be masked by nuclear gene capture or polyploidization. Alternatively, recombination with the maternal genome, gene conversion, or postcapture gene loss could replace alleles in the paternal genome. This would allow usually clonal androgenetic species to slow the rate of accumulation of deleterious mutations and potentially introduce beneficial allelic variants.

Persistent asexual lineages are expected to become extinct over evolutionary time, as they accumulate deleterious mutations, cannot free beneficial alleles from a poor genetic background, and must rely on mutation (not recombination) to introduce new beneficial alleles. Therefore, the persistence of asexual lineages is made more probable through rare genetic capture from divergent lineages. Even a limited amount of recombination could ameliorate the negative effects associated with asexual reproduction (e.g., refs. 24 and 51). Thus, any gene capture that occurs during androgenetic reproduction in Corbicula reduces the extinction risk of these otherwise asexual species.

Primarily asexual lineages have increasingly been shown to have mechanisms for rare acquisition of genetic material from other lineages (e.g., refs. 20–27 and 36). Capture of complete or partial maternal genomes in androgenetic Corbicula seems to be an additional mechanism of rare genetic recombination in otherwise asexual lineages. Therefore, modeling asexual reproduction as the complete lack of genetic recombination no longer seems appropriate in considering the evolution and persistence of asexual lineages. To understand the distribution and diversity of sexual vs. asexual species, we should turn to models that explore not whether sex is advantageous under a given set of conditions, but rather, how much sex is sufficient to counteract the deleterious effects of mutation accumulation (e.g., refs. 3 and 52–54).

Materials and Methods

Sequencing. DNA was extracted from frozen tissue specimens of Corbicula from North and South America and the Netherlands using the Viogene DNA Blood and Tissue Genomic extraction kits (Viogene Biotek). DNA was extracted from ethanol-preserved samples of Corbicula australis, Corbicula cf elongata, Corbicula fluminea, Corbicula fluminalis, Corbicula loehensis, Corbicula matannensis, C. molkitana, Corbicula tobiae, Corbicula leana, and Corbicula madagascariensis using a standard phenol-chloroform protocol modified by suggestions in ref. 55. A total of 19 specimens were used (Table S2). Primers for PCR amplification were developed using an α-amylase gene sequence from Corbicula (GenBank accession no. AF468816). Primers for atps-α were designed from sequence obtained using universal primers designed to amplify specific introns of the atps-α gene (56), but we could not confirm the identity of sequences amplified by these primers, because there are no existing sequences for these introns on GenBank for any Corbicula or their close relatives. SI Materials and Methods provides details on primer development, sequences, and amplification conditions.

Binning of Alleles. Cloning and sequencing introduce noise into a data set because Taq polymerase makes errors in replication (57, 58), and these errors can be sampled by cloning. Multiple sequences from one individual may not represent genetic alleles but rather PCR error. To reduce noise arising from PCR error, clones were binned such that a separate allele was called if there were more than three base pair differences between that sequence and another group of sequences. The consensus sequence for each bin was the most common base pair sampled across sequences in the bin, or coded using International Union of Pure and Applied Chemistry ambiguity rules if two or more base pairs were equally represented between sequences at a given site. The number of alleles inferred for each marker was consistent with the ploidy expected (triploid for most androgenetic clams, diploid for sexual species; Table S2).

Hedtke et al.

PNAS | June 7, 2011 | vol. 108 | no. 23 | 9523
Alignment. Alignments of sequence data were constructed using MacClade (G). Each alignment was trimmed at each end to exclude regions of missing data. For each aligned data set, we ran analyses in which all indels were included or excluded (any: 560 bp with indels included, 379 bp with indels excluded; atps-i: 347 bp with indels included, 263 bp with indels excluded). Excluding indels did not make a significant difference in phylogenetic inference, and we present results for analyses in which indels were excluded. We did not find evidence for recombining sequences (detailed in SI Materials and Methods).

Phylogenetic Analyses. We performed phylogenetic analyses on each data set using maximum likelihood as implemented in GARLI v0.96 (60) (SI Materials and Methods) and Bayesian phylogenetic analysis using MrBayes v.3.1.2 (61) (SI Materials and Methods). The outgroup to freshwater Corbicula, Corbicula japonica, did not amplify or was too divergent from the freshwater ingroup for successful gene alignment. Topologies are therefore unrooted for these markers (although midpoint rooting is used for the convenience of visualization). Details of hypothesis testing are presented in the SI Materials and Methods.

ACKNOWLEDGMENTS. We thank R. Linder and the members of the D.M.H., Bull, and Cannella laboratory groups at the University of Texas at Austin for advice on project development and constructive criticism of the manuscript; S. Otto, B. Normark, and D. Mark Welch for providing additional helpful comments; D. Ofhoiigel, C. Iltarte, and A. de Vaate for contributing tissue samples used in this study; R. Heineman, N. Wilson, and B. Ramesh for assisting with laboratory work; and Thomas von Rintelen for help during field work, which was supported by Deutsche Forschungsgemeinschaft Grant GL 297/7 (to M.G.). This work was funded by Doctoral Dissertation Improvement Grant DEB-0513426 from the National Science Foundation (to S.M.H.). S.M.H. was also supported by a National Science Foundation Graduate Fellowship.