



Both nuclear and cytoplasmic polymorphisms are involved in genetic conflicts over male fertility in the gynodioecious snail, *Physa acuta*

Fanny Laugier, Nathanaëlle Saclier, Kévin Béthune, Axelle Braun, Lara Konecny-Dupré, Tristan Lefébure, Émilien Luquet, Sandrine Plénet, Jonathan Romiguier, Patrice David

► To cite this version:

Fanny Laugier, Nathanaëlle Saclier, Kévin Béthune, Axelle Braun, Lara Konecny-Dupré, et al.. Both nuclear and cytoplasmic polymorphisms are involved in genetic conflicts over male fertility in the gynodioecious snail, *Physa acuta*. *Evolution - International Journal of Organic Evolution*, 2024, 78, pp.1227-1236. 10.1093/evolut/qpae053 . hal-04707802

HAL Id: hal-04707802

<https://hal.science/hal-04707802v1>

Submitted on 24 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Final version published as :

Laugier, F., Saclier, N., Béthune, K., Braun, A., Konecny, L., Lefébure, T., Luquet, E., Plénet, S., Romiguier, J., David, P. 2024 Both nuclear and cytoplasmic polymorphisms are involved in genetic conflicts over male fertility in the gynodioecious snail, *Physa acuta*. *Evolution*, 78(7) : 1227-1236 <https://doi.org/10.1093/evolut/qpae053>

Both nuclear and cytoplasmic polymorphisms are involved in genetic conflicts over male fertility in the gynodioecious snail, *Physa acuta*

Fanny Laugier¹, Nathanaëlle Saclier², Kévin Béthune¹, Axelle Braun³, Lara Konecny³, Tristan Lefébure³, Emilien Luquet³, Sandrine Plénet³, Jonathan Romiguier², Patrice David¹

¹CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France

²ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

³Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, F-69622 Villeurbanne, France

Corresponding author: CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France. Email: fanny.laugier@cefe.cnrs.fr

Abstract: Gynodioecy, the coexistence of hermaphrodites with females, often reflects conflicts between cytoplasmic male sterility (CMS) genes and nuclear genes restoring male fertility. CMS is frequent in plants, and has been recently discovered in one animal: the freshwater snail *Physa acuta*. In this system, CMS was linked to a single divergent mitochondrial genome (D), devoid of apparent nuclear restoration. Our study uncovers a second, novel CMS-associated mitogenome (K) in *Physa acuta*, demonstrating an extraordinary acceleration of molecular evolution throughout the entire K mitochondrial genome, akin to the previously observed pattern in D. This suggests a pervasive occurrence of accelerated evolution in both CMS-associated lineages. Through a 17-generation introgression experiment, we further show that nuclear polymorphisms in K-mitogenome individuals contribute to the restoration of male function in natural populations. Our results underscore shared characteristics in gynodioecy between plants and animals, emphasizing the presence of multiple CMS mitotypes and cyto-nuclear conflicts. This reaffirms the pivotal role of mitochondria in influencing male function and in generating genomic conflicts that impact reproductive processes in animals.

Key words: gynodioecy, cytoplasmic male sterility, restoration, arms-race, mitochondrial genomes, *Physa acuta*

Introduction

Gynodioecy is a sexual polymorphism in which hermaphrodites coexist with male-sterile individuals, that are therefore functionally female (Saumitou-Laprade et al. 1994, Touzet 2012). This coexistence led Darwin, in 1877, to question the Linnean categories of sexual systems in plants and to become interested in the evolution of modes of reproduction. Botanists extensively described gynodioecy (reviewed in Barrett, 2010), recognizing its prevalence in angiosperms (approximately 20% of families, Dufay et al. 2014, Caruso et al. 2016). Regarded as a potential transition from hermaphroditism to dioecy (Barrett 2002), the evolutionary dynamics of gynodioecy have been thoroughly explored in plants, combining theoretical and empirical perspectives (Shykoff et al. 2003, Delph et al. 2007, McCauley & Bailey 2009). Gynodioecy in plants usually reflects the presence of mitochondrial Cytoplasmic Male Sterility (CMS) genes that suppress the male function in individuals that would otherwise be hermaphroditic (Chase 2007). Because of maternal inheritance, such genes can increase in frequency as soon as they slightly enhance female fitness, irrespective of their effects on male fitness (Lewis 1941, Lloyd 1974). However this suppression adversely affects nuclear genes, as half of a hermaphrodite's fitness relies on the male function. Consequently, some nuclear genes, known as restorers of male fertility, evolve to counteract CMS by converting individuals with male-sterilizing cytoplasms back into hermaphrodites (Cosmides and Tooby 1981).

Given the prevalence of the mito-nuclear conflict in plants, and the shared mechanisms of mitochondrial and nuclear inheritance, the question arises whether animals follow a similar evolutionary path. Although about one-third of non-insect animal species are hermaphrodite (Jarne & Auld 2006), gynodioecy is very rare (Weeks 2012). This may be explained by the compact and structurally stable nature of animal mitochondrial genomes, that may limit the emergence of male-sterile variants (Boore 1999, Galtier 2011). In addition some gynodioecious animals may remain undetected, as many hermaphroditic animals have concealed sperm-producing organs, making male-sterile individuals challenging to identify.

To date, only one animal with confirmed cytoplasmic male sterility (CMS) has been identified: *Physa acuta*, a freshwater snail (David et al. 2022). This snail provides a unique opportunity to assess how universal concepts developed in plants are. The initial study however revealed two major differences from known gynodioecious plants. First, only one CMS-associated mitochondrial type (D) was found, with extreme molecular divergence from normal cytotypes (N) at all mitochondrial genes. Second, there was no evidence for nuclear restoration, as nearly all captured D individuals exhibited zero male fitness (David et al. 2022).

In contrast, multiple CMS genes often coexist in natural populations of gynodioecious plants (e.g. in *Beta vulgaris* Cuguen et al. 1994, *Thymus vulgaris* Belhassen et al. 1991, *Silene vulgaris* Charlesworth and Laporte 1998) as well as in cultivated crops (28 CMS genes identified in 13 species, Kim & Zhang 2018). In some cases, no fertile (i.e. non-CMS or maintainer) cytoplasms can be found (e.g. *Plantago lanceolata*, de Haan et al. 1997). The different CMS types within a species have specific nuclear restorers, usually genes encoding for pentatricopeptide repeat (PPR)-containing proteins that act by suppressing the expression of mitochondrial genes that cause sterility (Dahan & Mireau 2013). For example, maize requires the genes Rf1 and Rf2 for cms-T, Rf3 for cms-S and Rf4 for cms-C (Snyder and Duvick 1969, Kamps and Chase 1997, Sisco 1991). The diversity of CMS types is explained by (i) the high rate of mitochondrial DNA rearrangements facilitating the production of new CMS mutants (Adhikari et al. 2019), (ii) frequency-dependent selection that maintains intermediate frequencies of CMS types and corresponding Rf alleles due to the pleiotropic costs of restoration on fitness (Gouyon et al. 1991, Bailey 2002, Del Castillo & Trujillo 2009, Caruso & Case 2013).

The presence of a single male sterility mitotype in *P. acuta* is unexpected, given that the high mutation rate thought to be responsible for the divergence of the D mitotype could lead to multiple CMS types. It remains unknown whether CMS can occur in this system without high molecular divergence, or whether sterility and divergence are interdependent. No evidence for restorers was found in *P. acuta* – although some D-mitotype individuals were male-fertile, their frequency was low (8%) and it is uncertain whether they were caused by nuclear alleles, *versus* an incomplete penetrance of the male sterility phenotype. If this species does indeed have a single male-sterile mitotype without restoration, the limits to CMS expansion might be very different from those usually acting in plants. To investigate this question, we explored mitochondrial diversity in populations around Lyon (France), where D individuals were discovered by David et al. (2022), looking for potential new mitotypes and evidence of restoration.

Here, we report the discovery of a new mitochondrial type, named K, in two populations near Lyon. We investigated its phylogenetic relationship with existing types, particularly its proximity to D and the non-CMS mitotype N. We then tested whether K mitotypes were also associated with a male-sterile phenotype and whether this phenotype was affected by the nuclear background, as expected in the presence of nuclear restorers of fertility.

Material and methods

Discovery of the K mitotype

We sampled *P. acuta* at four stations around the initial discovery site of the male-sterile D type in Lyon, France (October 2019, Table S1). Approximately one hundred G0 individuals from each site were mitotyped after some were allowed to lay separately to preserve progeny. DNA was extracted using a Chelex protocol (Gautam 2022), and we used the David et al. (2022) PCR test designed to identify D and N mitotypes. This involved two PCRs, each with specific primers of a 604-bp fragment (N-specific in the first PCR, D-specific in the second) as well as external universal primers of a 705-bp fragment in the mitochondrial COI gene (Folmer et al. 1994). The 705-bp fragment is amplified only when the specific primers fail to anneal. The resulting PCR products, run side by side on agarose gels, allow mitotype identification based on fragment size patterns (long-short for D, short-long for N, David et al. 2022).

While investigating, some individuals exhibited a long-long COI pattern (Figure S2), suggesting a mismatch with both N- and D-specific primers (Figure S3). Seven such individuals (two from Irigny and five from Erevan) were Sanger sequenced for the COI fragment amplified with "Folmer" universal primers, revealing a distinct sequence, that we termed "mitotype K". Subsequently, a new PCR test was developed for one-step mitotype assessment (N, K, or D) based on COI sequences. The test used a mixture of five primers, including two universal external COI primers and three internal primers specific for N, D, and K. On an agarose gel, the resulting fragment size directly indicated the corresponding mitotype, with a long fragment indicating no match to N, D, or K. The primers and protocol for this "triplex" PCR test are detailed in Figure S4.

Introgression experiments.

Nuclear restoration is present when the propensity of a mitotype to produce male-fertile or male-sterile individuals, depends on the nuclear background to which it is associated. We tested this idea through an introgression experiment over 17 generations, inserting the mitochondrial lineages derived from wild-caught ancestors from Lyon, into the "naive" nuclear background

of a laboratory population (see introgression protocol detailed Supplementary Text 1 and Figure S1). Introgression occurred within “matrilines”, defined as sets of focal individuals sharing the same ancestor by maternal descent, and hence the same mitochondrial DNA; there were 25 matrilines (12 K and 13 N). At each generation, the sperm donors were taken from a laboratory population of *P. acuta*, derived from a mixture of ten natural populations from Montpellier (300km south of Lyon, Noël et al. 2016), with an albino phenotype. The albino population is “naive” because it consists of 100% N-mitotype individuals and has been maintained in captivity for more than 80 generations: selection should not have maintained high frequencies of restorer genes, if they exist and have a cost. Within the matrilines, the nuclear genes (initially from Lyon) were generation after generation replaced by those of the albino population. We however selected only pigmented offspring to propagate each matriline, allowing us to distinguish them from sperm donors. We performed PCR tests at different generations to check that the mitotype remained the same within a matriline over generations (Supplementary Text 1).

At each of the 17 generations, we assessed the male-fertility status of the focal individual on the basis of the offspring laid by albino mates (Figure S1). The presence of pigmented juveniles indicates that the (pigmented) partner from the matriline is male-fertile. Albinism is driven by a single recessive allele: if an albino mother produces a pigmented offspring, the father must be a pigmented individual.

Behavioral and reproductive traits of K and N individuals.

In the introgression protocol, we investigated the sexual phenotype of the newly discovered mitotype K. Both behavioral and fitness traits were assessed for the matrilines K and N at G1 and G5, with two to three individuals from each matriline isolated until sexual maturity. We conducted observations of mating pairs during 45mn, each focal individual from a matriline (pigmented) paired with a N-mitotype albino mate, both previously virgin, with behaviors recorded and timed. After observations, snails were left together for three days, re-isolated for egg-laying in separate boxes for three more days, and then we collected their clutches. Female fitness was assessed by counting the number of eggs laid, male fitness by the number of juveniles sired by the focal in the progeny of its partner. Male fertility status was defined by the ability to stimulate oviposition when inseminating their virgin partner and to sire a majority of offspring (Supplementary Text 1).

Statistical analyses.

We used linear mixed models (LMM) and generalised linear mixed models (GLMM) with the lme4 R package (Bates et al. 2015) to analyse juvenile counts (fitness) and time spent in male or female behaviours. Poisson distributions were used for count variables and a Gaussian model for quantitative variables. To account for relatedness between individuals, a random factor 'matriline' was included, along with an individual random factor to deal with potential overdispersion in GLMMs.

Based on the 17-generation time course of inheritance of the male sterility phenotype within matrilines, we used Maximum likelihood estimation to determine the frequencies of polymorphic restorer alleles in the Lyon population (p_0) and the albino population (p), under the assumption of mono-locus dominant restoration (i.e. recessive variant is sterile in the context of K), the most common restoration determinism observed in plant populations in which restoration is highly prevalent (Delph 2007). The likelihood was computed for each (p_0 , p)

value, representing the probability of the observed male fertility status of founder K individuals and all their descendants within each matriline, conditional on their pedigree - in principle, the frequency of fertile individuals should start from p_0 and progressively evolve towards p with successive generations of introgression (details of likelihood computation in Table S5). Multiple (p , p_0) values (0 to 1 in steps of 0.01) were explored to determine the maximum likelihood, with 95% confidence intervals obtained using the profile likelihood method. All statistical analyses were performed in RStudio v. 4.3.1 (R Core Team 2023).

Mitochondrial genome sequence. To reconstruct the complete mtDNA of type K, short-read whole genome sequencing was performed on one individual (Supplementary Text 1, ENA project PRJEB50799). The genome of this individual was assembled using MEGAHIT (Li et al. 2015, 2016). We then extracted the contig corresponding to the mtDNA by blasting a COI sequence from a K mitotype against the assembly. The contig was then annotated using the MITOS webserver (Bernt et al., 2013). The absence of contamination and chimeras in the assembly was verified by looking at the coverage along the genome (Supplementary Text 1, Figure S4). The assembled and annotated mitochondrial genome were deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB50799. Once the mtDNA was assembled, we sequenced ten individuals carrying the K mitotype. Ten K-mitotype individuals were bred (three from Irigny and six from Erevan; with the long-long band pattern), and sent for Illumina transcriptome sequencing (Supplementary Text 1).

Phylogenetic analyses.

The sequence alignments made in this section are available in Zenodo (10.5281/zenodo.10521893).

Physidae phylogeny. To investigate the phylogenetic origin of *P. acuta* mitotypes, we used mitochondrial protein-coding genes from K, D and N mitotypes, including newly acquired K sequences and D and N sequences from David et al. (2022). In addition, seven gastropod species from GenBank, consisting of three *Physidae* species and five other outgroup species (consistent with David et al. 2022, plus *Physa fontinalis*, Table S2 section 1), were included. The 11 protein-coding genes (excluding ATP8 and NAD4L) were first aligned without *P. acuta* K and D using MAFFT (Kato et al. 2002). K and D mitotypes were then mapped onto this alignment using MACSE (Ranwez et al. 2011). G-blocks (50% gap allowed, Castresana 2000) were used to remove ambiguously aligned sites and coding sequences were concatenated. A phylogenetic tree was constructed using PhyloBayes (Lartillot et al. 2009) under a CAT-GTR+G+I model, focusing on the first and second codon positions to avoid saturation. Two independent chains were run and stopped after 455,785 steps (40,000 generations excluded as burn-in), after checking for convergence with the tracecomp program (effective sample size > 50 and discrepancy between chains < 0.3 for all statistics). The majority consensus tree, obtained from 41,578 trees sampled every 10 steps, was derived using a 10% threshold to resolve polytomies. This tree was used to estimate non-synonymous (dN) and synonymous (dS) substitution rates under the YN98 codon model. YN98 model parameters were estimated using bppml (Guéguen et al. 2013), and dN/dS, together with Kr/Kc, were calculated using substitution mapping with mapNH (Romiguier et al. 2012). Tree processing was performed using the R package ggtree (Yu et al. 2020).

Physa phylogeny. The COI sequence of the K mitotype was blasted against the NT database of GenBank to identify similar sequences. A COI dataset was then assembled using COI sequences retrieved from GenBank with more than 93% of identity to our K sequence (Table S4), seven K sequences from the Irigny and Erevan populations obtained by Sanger sequencing using the same primers as in the PCR tests, the D sequence and the five N sequences used in David et al. (2022), together with the *P. fontinalis* sequence used as an outgroup (Table S2 section 2). A phylogenetic tree was built using the same protocol described above.

Comparison of divergence among gastropods. To the COI dataset, we added COI sequences of other snails retrieved from GenBank and used in David et al. 2022 (Table S2 section 3). The sequences were aligned using MAFFT. Mean divergences between sequences of the K mitotype and between sequences from the three mitotypes (D, N and K) were then obtained for the COI using the APE package in R.

Results

K mitochondrial genome sequence

We developed a fast PCR protocol (see “triplex” PCR protocol, Figure S3) to distinguish a newly described male-sterile mitotype (K) from the previously found N and D mitotypes (respectively normal and male-sterile, David et al. 2022). The K mitotype was found in 12% of the individuals from the Irigny population and 32% from the Erevan population (Table S1). No K individuals were found in the Crêt-2 and Saint-Priest populations.

The seven K mitotype snails sequenced (two from Irigny and five from Erevan) shared an identical COI sequence (0% divergence). By blasting the K COI sequence in GenBank, six sequences between 100% and 93% of identity were found. These were from North America (Idaho), Singapore and South Africa and referenced as *Physa sp.* (Table S4). The average divergence between all 11 K sequences (seven from our populations and the six from Genbank) was 1.58% ($\pm 2.48\%$ SD, uncorrected DNA distance). The K COI sequences from Singapore and South Africa were almost identical to our K sequences from Lyon (all had 100% identity, except one of the two South African samples at 97.4%). The COI sequences from North America were more divergent (nearly 7% divergence from the Lyon K sequences; and 5.9% ($\pm 1.3\%$ SD) from all other K COI sequences combined - the seven from our populations plus the four from both Singapore and South Africa). A COI phylogenetic reconstruction grouped the K sequences into a single clade ($pp = 1$, Figure 1B) composed of two subclades: North American sequences, and all other localities including the French, South African and Singapore populations.

Comparison of divergence among gastropods. Based on COI gene divergences, the newly discovered K mitotype was 22.8 % ($\pm 0.79\%$ SD, 11 individuals) divergent from the normal mitochondrial genome of *P. acuta* (5 individuals of N mitotype, among which the average divergence was $2.2\% \pm 1.6\%$, Table S6). Interestingly, this divergence was even stronger when comparing K with the male-sterile D mitotype (28.1%, $\pm 0.43\%$ SD, average of five D individuals). The same observation, though at a higher divergence scale, was made for 11 additional mitochondrial genes: the K-N DNA divergence ranged from 35% to 57%, while the K-D divergence ranged from 41% to 61% (Table S6). The divergence between K and N (or D) was equivalent to that expected between snail clades that diverged hundreds of millions of years ago (Figure 1C).

Physidae phylogeny. In the phylogeny based on all mitochondrial genes, the K mitotype was not clearly grouped with either N or D (Figure 1A). Although K was clearly within the *Physidae* and more precisely the *Physella* clade (*Physa acuta* + *Physa gyrina*), deep nodes within this group were not well supported (posterior probabilities below 0.9) because the divergences between *P. gyrina* and the three mitotypes of *P. acuta* are probably too high, resulting in long-branch attraction artifacts (Figure 1A). For the terminal branches leading to each of the three mitotypes (K, D and N), the dN/dS ratios were low (0.035, 0.036 and 0.032), as were the Kr/Kc ratios (0.784, 0.798 and 0.725). The three mitogenomes N, K, D are of similar length (14.4 Mb) and they contain the same set of genes; however the K mitogenome differs from the N and D mitogenomes by multiple rearrangements (Figure S6).

Introgression experiments

The control PCRs (N = 175, Supplementary Text 1) carried out during the introgression experiment showed that the mitotype was conserved within a matriline, i.e., the individuals within a matriline identified as K mitotype at G0 maintained the K mitotype in subsequent generations (G1, G2, G7 and up to G25).

Phenotyping

Behavioral and reproductive traits of K and N individuals

In G1, no significant difference in behavior (male or female) was found between K and N individuals (Figure 2A). Similar results were obtained for female and male fitness (Figure 2B and Table S3). However, at G5, K individuals exhibited a significantly reduced male activity compared to N mitotype individuals (Figure 2A). Out of 26 albinos paired with K snails, 18 did not lay any eggs or produced only albino hatchlings (i.e. self-fertilized brood). However, the female behavior and fertility of K snails did not differ from that of N snails, and they produced many offspring sired by albino mates (Figure 2B and Table S3).

Proportion of male sterility in K progressively increased with generation of introgression

The nuclear background of the mitochondrial lineages originating from wild-caught ancestors from Lyon was diluted into the “naive” nuclear background of the Montpellier albino population over 17 generations. This introgression protocol led to an increase in the proportion of male-sterile K individuals; after G11, male sterility stabilized at around 60% (Figure 3A). The proportion of male-sterile individuals in the N matriline was consistently low and unaffected by introgression. From G1 to G17 male sterility was heritable, as male-sterile K mothers produced a majority of male-sterile offspring, significantly more than male-fertile mothers (Chi-squared test, $X = 216.62$, $P < 0.001$). However, male-sterile K mothers still had 27% of male-fertile offspring (Figure 3B).

The MLE of the restorer allele frequency (assuming a single-dominant gene) was 0.26 ± 0.08 in the albino population and 0.56 ± 0.08 in the K individuals from Lyon (Figure S5).

Discussion

Highly divergent K mitotype is associated with CMS and restoration

David et al (2022) reported the first example of cytoplasmic male sterility in animals, linked to a highly divergent mitochondrial genome (the D mitotype). Here, we discovered another such mitotype, K. The mitochondrial nature of K sequences is confirmed by their high coverage, the absence of stop codons or frameshifts in coding sequences, and the maternal transmission of PCR patterns that persisted over more than 20 generations while the nuclear genome was progressively replaced by introgression. We did not obtain evidence for mitochondrial

heteroplasmy either (Supplementary Text 1). The K mitotype shares its essential characteristics with D: a male sterility phenotype and high divergence from the normal mitochondrial type (N) without evidence of accelerated non-synonymous substitutions proportions (dN/dS is stable). However, the expression of male sterility is different from that of D. Indeed, first-generation K individuals, whose genome comes entirely from their natural populations of origin, mostly looked normal: when pairing N snails from an albino laboratory strain with K snails, both individuals mated in both sex roles and could sire offspring. It was only by introgressing the K mitotype into the genomic background of the “naive” albino population from Montpellier (which has been kept for more than 80 generations in the laboratory, without any contact with the K mitotype) that K snails started to show a reduction or suppression of male behavior and fertility. During introgression, K male-sterile mothers tended to produce more male-sterile offspring than male-fertile ones, showing that the male-sterile phenotype associated with the K mitotype is heritable. The results of the introgression experiment indicate the existence of nuclear polymorphisms that modulate the expression of the male sterility phenotype in K individuals, i.e., restorer genes. This stands in contrast to the D mitotype, in which almost all individuals from the natural population in Lyon were male-sterile (a small fraction was fertile), a situation that remained unchanged by introgression into the genomic background of the Montpellier albino stock (David et al. 2022). The K mitotype in *P. acuta* is therefore the first animal example of CMS with clear evidence of nuclear restoration.

High frequency of restoration in P. acuta natural populations

In plants, previous studies suggest that restorer alleles are more often dominant than recessive (reviewed in Delph et al. 2007). We were able to estimate the frequency of a putative restorer allele conferring male fertility to the K mitotype under a mono-dominant scenario. Interestingly the restorer frequency was estimated to be 26% in the Montpellier albino population, reflecting the fact that even after 17 generations of introgression (i.e., when the quasi-totality of the nuclear genome originates from the Montpellier albino stock), a significant fraction of male fertility still exists in K snails and male-sterile mothers continue to produce a fraction of male-fertile offspring. Under the same scenario, the frequency of the restorer allele in the original K population was then estimated to be around 56%. This latter result is consistent with the high frequencies of restorer alleles observed in some gynodioecious plant populations. For example, in natural populations of *Raphanus sativus* *Rf* frequencies range from 41% to 100% (Murayama et al. 2004); in wild rice species it is either 60% or 41.9%, depending on the CMS type (Li et al. 2005).

The presence of restorers in our albino population might appear paradoxical; after more than 80 generations maintained as a pure N mitotype population, one could expect restorer alleles to have been eliminated from this population. Two possibilities could explain the maintenance of the restorer gene: (i) the K mitotype was historically present in the original Montpellier source populations (which served as founders of our laboratory stock 80 generations ago) and the cost of restoration in the absence of CMS is low, so the restorer frequency is decreasing only slowly over generations, (ii) our scenario on the genetic architecture of restoration is too simple; it is possible, for example, that K male-sterility would not be total, even in a naive population without specific restorers (David et al 2022). In this case, there would be no nuclear background producing 100% of male sterile individuals, whatever the mitotype. The phenotype of K would be a partial reduction in male fertility (e.g., a decrease in sperm production, not a complete elimination) leading to a non-zero proportion of male fertile individuals under the conditions of our tests (pairing with virgin individuals). This situation would not be exceptional as threshold models have been used to understand

complicated restoration patterns of inheritance of CMS in plants (Ehlers et al. 2005, Bailey & Delph 2007) – in such models, the genetic variance for a multigenic quantitative trait (liability) determines the fraction of male-sterile individuals in a given cytoplasm/mitotype, through a threshold function : individuals with a trait value above a given threshold will have a hermaphrodite phenotype, while the remaining individuals will have a female phenotype. As in plants, a combination of single locus and complex genetic basis of restoration may apply to *P. acuta*, with different relative prevalence in populations depending on their costs and on the abundance of CMS. Further studies are needed to clarify the genetic determinism of the restorer and its frequency, particularly by selecting nuclear backgrounds with contrasting frequencies of the restorer in the Montpellier populations.

The origin of highly divergent mitotypes and intragenomic conflicts

Our results showed that (i) The K mitotype is not an offshoot of the D mitotype, (ii) K and D are both very distant from the N mitotype and from each other, (iii) the order of genes differs among the three mitogenomes without any clear resemblance shared by the two CMS-associated ones. Unfortunately, the K and D mitotypes have diverged so much that it has become impossible to reconstruct their detailed origin. The K and D mitotypes branch at the root of the *Physa* group (including *acuta* and *gyrina* species), which is likely due to long-branch attraction artifacts and saturation that prevent to fully reconstruct their phylogenetic history. As suggested by David et al. (2022) elevated mutation rates may be responsible for the high divergence of the CMS mitotypes. As similar results are found in the K mitotype, the patterns observed may be driven by the same mechanism. However, an increase in the mutation rate alone would increase the substitution rate, but it should also increase the genetic diversity.

As for the D mitotype, the K mitotype has a high substitution rate but low diversity. The seven mitochondrial sequences from both the Erevan and Irigny populations were almost identical; the COI sequences originating from Singapore and South Africa were also sub-identical to the COI sequences from the Erevan and Irigny populations (0.34% of divergence). The combined observation of a high substitution rate, very low diversity and worldwide distribution suggests that this particular K haplotype has been recently selected and has invaded the *P. acuta* populations. The slightly more divergent (7%) variants of the K mitotype found in North America (Idaho) suggest that more diversity may be found in the native area of *P. acuta*. Indeed, *P. acuta*, although initially described in the Garonne River (France) (Draparnaud 1805), is thought to have originated in North America (where its diversity is maximal, and where sister species are found), and spread to the old world (Europe, Asia, Africa) only recently with human exchanges (it is now probably the most widespread freshwater mollusc, Ebbs et al. 2018). While this invasion has not been accompanied by spectacular losses in molecular diversity (Bousset et al. 2014), the native area of *P. acuta* probably harbours a diversity of CMS mitotypes. The high divergence of these mitotypes has probably hindered their inclusion in phylogeographic studies: people may discard them from analyses thinking their sample is contaminated by another species (Ng et al. 2015), or consider them as a cryptic taxon, basal in the Physid phylogeny (Young et al. 2021).

The logical link between genomic evolution of the mitochondrial DNA and male-sterility is still elusive. First, genomic instability of the mitochondrial DNA (whatever its cellular cause) may fuel the production of male-sterile variants (as in plants). Second, certain mutations could have an impact on the energy metabolism of sperm or sperm-producing cells without affecting females (see Havird et al. 2019). In addition, the fast molecular evolution of the mitochondrial DNA may favour the persistence of such variants over time, in the context of an arms race with nuclear genes. This hypothesis is in agreement with the apparent lack of

accumulated polymorphism within male-sterile mitotypes (K and D) despite a mitogenome-wide increase in mutation rates, which suggest recurrent selection drafts. Finally, one cannot exclude that the instability of the mitogenome and the male-sterility phenotype are two consequences of a common cause (such as a maternally-transmitted cytoplasmic symbiont) – but we have no evidence for it so far.

Accelerated mitogenome evolution has also been observed in some plant clades where CMS is frequent (e.g., *Plantaginaceae* and *Caryophyllaceae*, Cho et al. 2004, Mower et al. 2007). It has been proposed that the driver of this accelerated evolution is antagonistic coevolution, known to drive rapid evolution in other contexts (Brockhurst et al. 2014, Ebert & Fields 2020). Indices of arms race evolution have indeed been observed in the main family of plant restorer genes (Fujii et al. 2011). The finding of male-fertility restorers in *P. acuta* opens up the possibility of a similar interpretation in our system. The arms race could be linked to the penetrance of the male sterile phenotype: a mutation conferring higher penetrance to a CMS mitotype would increase in frequency, while restorers would be selected to reduce the penetrance. In this context, the D mitotype (for which no restorers are yet known) should invade the populations, but -with the little available data- its frequencies seem to be relatively low at both local and global scale compared to K. We still have a very incomplete picture of a potentially highly dynamic system in which new divergent male-sterilizing mitotypes may emerge and spread quickly before being counteracted by restorers. Tracking the dynamics of divergent mitotypes in natural populations over time would be an essential source of information on this issue.

Theoretical studies of gynodioecy (reviewed in Delph et al. 2007) show that CMS polymorphism can be maintained under a form of balanced selection (Charlesworth 2002) when CMS mitotypes have a female advantage and restorers have sufficient costs. In our study, female advantage would result in higher female fitness of K when not restored ($K > N$ in G5); while we would expect $K < N$ when K is restored (in G1), resulting from costs of restoration. We did not find such differences. However this does not mean that differences would not appear in other conditions. Female advantage is frequent, but very variable and not systematically significant in plants, and probably depends on environmental conditions (Dufaÿ & Billard 2012). Restoration costs have been detected directly through their effects on seed (de Haan et al. 1997, Del Castillo et al. 2009) or pollen production (Bailey 2002, Caruso & Case 2013), or inferred through equilibrium models from the variation in sex ratio among natural populations (Case & Caruso, 2010). In *P. acuta*, restoration costs bearing on sperm production or the number of eggs laid should similarly be investigated, contributing to accurately model the evolutionary dynamics of the first animal example male sterility of CMS.

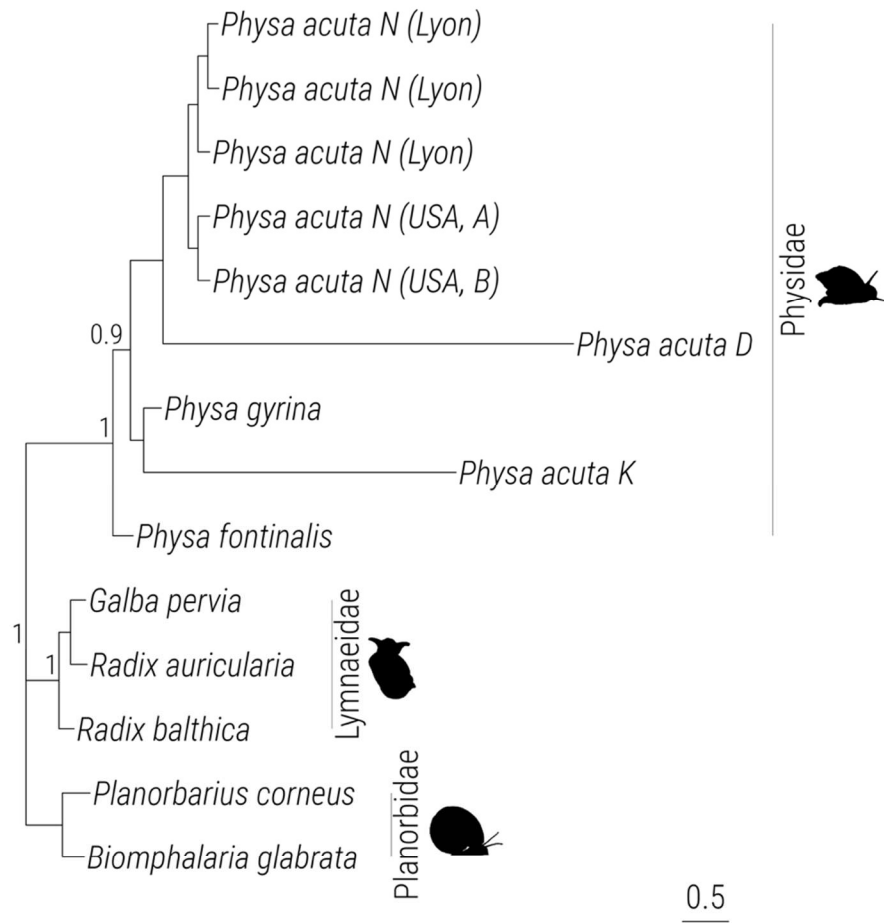
Conclusion

While *Physa acuta* is the first example of complete cytoplasmic male sterility, it is by no means the only example of cyto-nuclear conflicts mediated by mitochondria on animal reproduction. Indeed, the same reasoning, based on the different modes of inheritance of mitochondrial and nuclear genes, predicts that in gonochoric animals, mitochondrial mutations harmful to males can persist or spread if they are not harmful to females, a phenomenon known as the mother's curse (Camus et al. 2022). The presence of nuclear restorers of CMS in *Physa* is paralleled by the evolution of compensatory gene or gene expression in the nucleus of gonochoristic species, mitigating the effects of male-harming mitochondrial mutations; this process manifests itself in the form of reduced male fitness when the mitochondrial DNA meets an incompatible nuclear background (Innocenti et al. 2011, Patel et al. 2016, Camus & Dowling 2018, Perlman et al. 2015). Maternally-transmitted bacterial symbionts sometimes have extreme adverse effects on

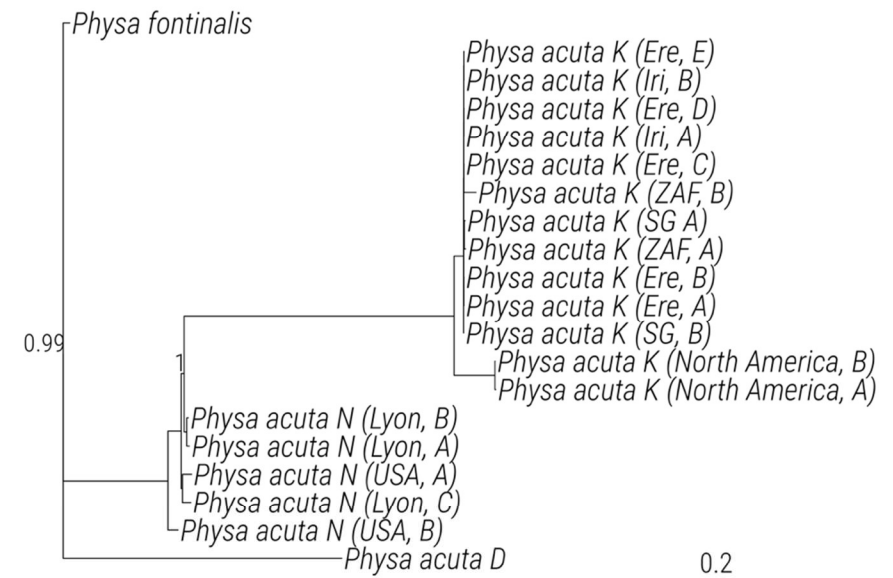
male fitness (e.g. male killing) or feminize males, but to our knowledge, this was not known for mitochondria (Camus et al. 2022). Our findings reaffirm the pivotal role of mitochondria in generating genomic conflicts that impact reproductive processes in both gonochoric and hermaphroditic animals. CMS has long been considered impossible in animals because of the highly compacted and structurally stable mitochondrial genome, in contrast to plants where mitochondrial genomes are larger and prone to structural rearrangements that occasionally give rise to chimeric genes responsible for CMS. Here we show that not only does CMS exist in animals, but its dynamics are not so different from those observed in plants, suggesting that mechanisms of genomic conflicts between the sexes, if they are arguably under different degrees of constraint, are fundamentally common to most sexual organisms.

Figures and captions

A



B



C

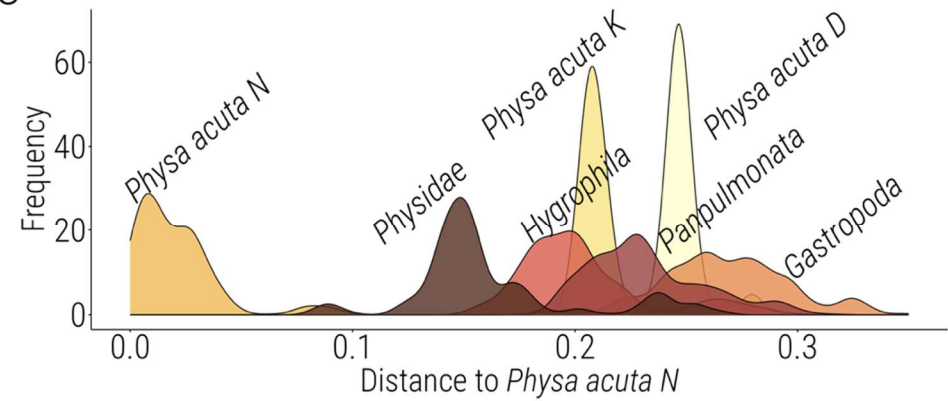


Figure 1: K mitotype is divergent from conspecific individuals and different from the D mitotype. Data used are given Table S2. **A:** Physidae phylogeny based on 11 mitochondrial protein coding genes and built with PhyloBayes (Lartillot et al. 2009) under a CAT-GTR+G+I model selecting only first and second codon positions because saturation. Data include 7 sequences from Genbank as well as 7 sequences of *P. acuta* (5N + 1D + 1K) from Lyon. Species from Lymnaeidae and Planorbidae families were used as outgroups. **B:** COI phylogeny of *Physa acuta*. *Physa fontinalis* was used as an outgroup. Ere: Erevan, Iri: Irigny, ZAF: South Africa, SG: Singapore. Numbers at the nodes represent posterior probabilities; probabilities below 0.9 are not shown. **C:** Distributions of pairwise DNA distances at the COI mitochondrial gene in snails. All pairs include one individual of the N mitotype of *P. acuta*. Data include 304 sequences from Genbank as well as 41 new sequences of *P. acuta* (5D, 29N and 7K).

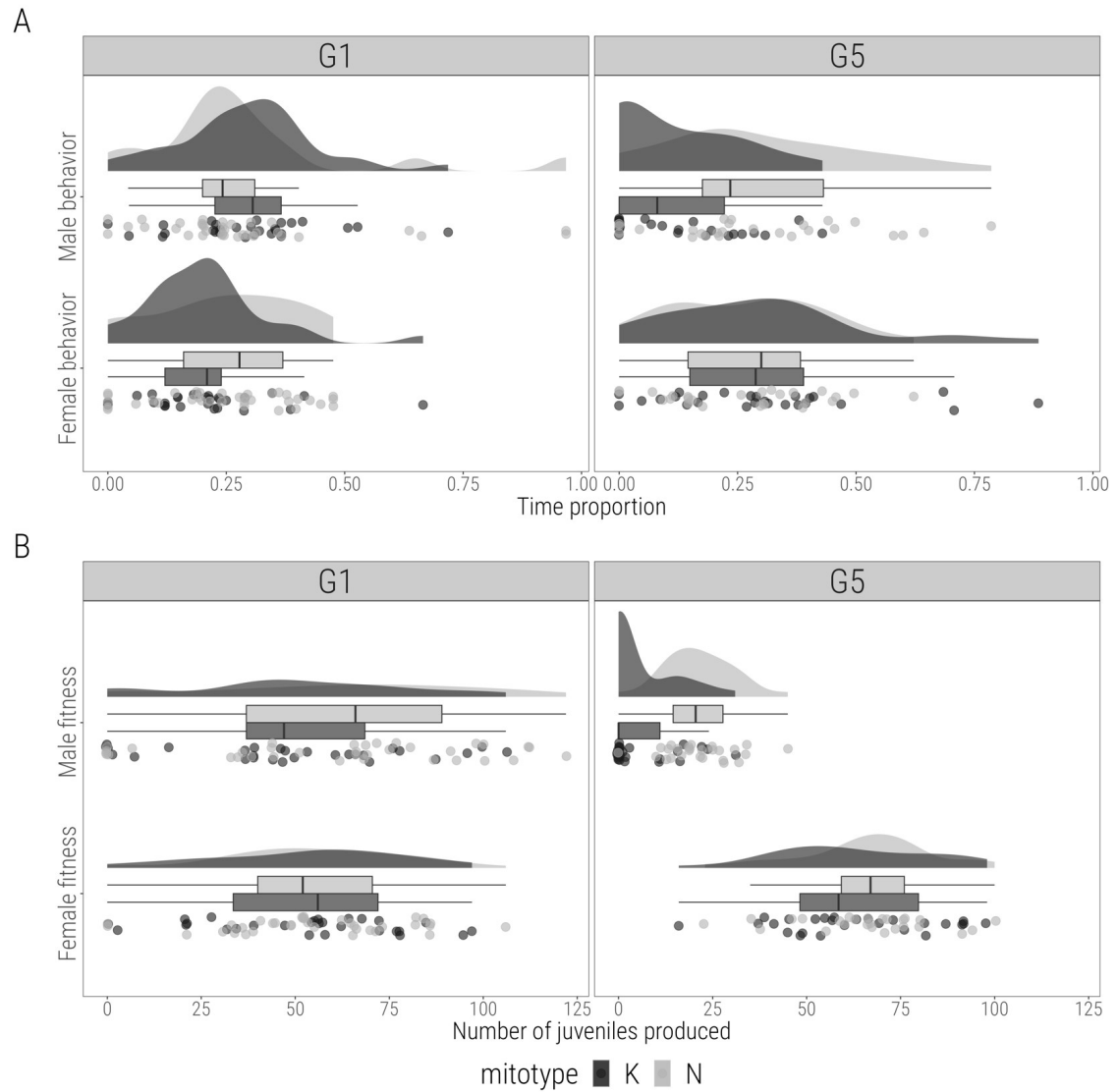


Figure 2: The K mitotype has a male-fertile phenotype at G1 then a male-sterile phenotype at G5. A: Distributions of time spent in male and female behavior by K and N individuals paired with virgin partners at both G1 and G5. **B:** Male fitness, i.e., number of juveniles produced by partner and sired by focal snails. Female fitness, i.e., number of eggs laid by focal snails.

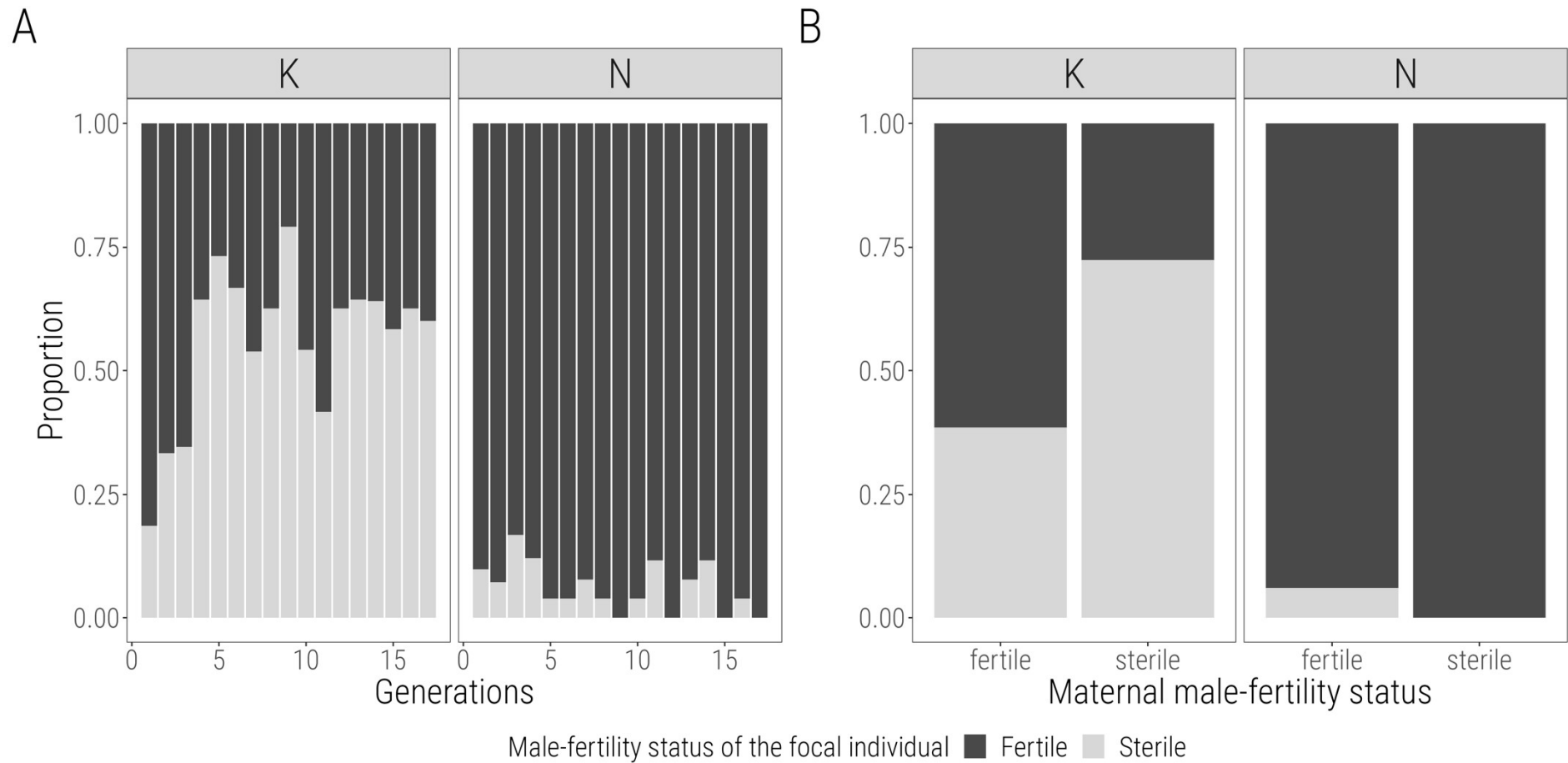


Figure 3: Male-sterility increases in K matriline when introgressed into a naive genomic background, and is heritable. A: Proportion of male-sterile and male-fertile individuals in both K and N mitotype, from first-generation individuals taken from Lyon (nuclear background entirely from Lyon) to 17-th generation individuals whose genome has been near-completely replaced by that of the albino Montpellier laboratory stock. **B:** Male-fertility of daughters of male-sterile and male-fertile mothers (over all 17 generations). In both A and B, semi-sterile and fully male-sterile individuals were pooled into a single “sterile” category.

References

- Adhikari, B., Caruso, C. M., & Case, A. L. (2019). Beyond balancing selection: frequent mitochondrial recombination contributes to high-female frequencies in gynodioecious *Lobelia siphilitica* (Campanulaceae). *New Phytologist*, 224(3), 1381–1393.
- Bailey, M. F. (2002). A Cost of Restoration of Male Fertility in a Gynodioecious Species, *Lobelia Siphilitica*. *Evolution*, 56(11), 2178–2186. <https://doi.org/10.1111/j.0014-3820.2002.tb00142.x>
- Bailey, M. F., & Delph, L. F. (2007). Sex-Ratio Evolution in Nuclear-Cytoplasmic Gynodioecy When Restoration Is a Threshold Trait. *Genetics*, 176(4), 2465–2476. <https://doi.org/10.1534/genetics.107.076554>
- Barrett, S. C. H. (2010). Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537), 99–109. <https://doi.org/10.1098/rstb.2009.0199>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Belhassen, E., Dommée, B., Atlan, A., Gouyon, P. H., Pomente, D., Assouad, M. W., & Couvet, D. (1991). Complex determination of male sterility in *Thymus vulgaris* L.: Genetic and molecular analysis. *Theoretical and Applied Genetics*, 82(2), 137–143. <https://doi.org/10.1007/BF00226204>
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., Pütz, J., Middendorf, M., & Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69(2), 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic acids research*, 27(8), 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>
- Bousset, L., Pointier, J. P., David, P., & Jarne, P. (2014). Neither variation loss, nor change in selfing rate is associated with the worldwide invasion of *Physa acuta* from its native North America. *Biological Invasions*, 16, 1769–1783.
- Brockhurst, M. A., Chapman, T., King, K. C., Mank, J. E., Paterson, S., & Hurst, G. D. D. (2014). Running with the Red Queen: The role of biotic conflicts in evolution. *Proceedings of the Royal Society B: Biological Sciences*, 281(1797), 20141382. <https://doi.org/10.1098/rspb.2014.1382>
- Camus, M. F., & Dowling, D. K. (2018). Mitochondrial genetic effects on reproductive success: signatures of positive intrasexual, but negative intersexual pleiotropy. *Proceedings of the Royal Society B: Biological Sciences*, 285(1879), 20180187.
- Camus, M. F., Alexander-Lawrie, B., Sharbrough, J., & Hurst, G. D. (2022). Inheritance through the cytoplasm. *Heredity*, 129(1), 31–43.
- Case, A. L., & Caruso, C. M. (2010). A novel approach to estimating the cost of male fertility restoration in gynodioecious plants. *New Phytologist*, 186(2), 549–557.
- Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Caruso, C. M., & Case, A. L. (2013). Testing models of sex ratio evolution in a gynodioecious plant: female frequency covaries with the cost of male fertility restoration. *Evolution*, 67(2), 561–566.

- Caruso, C. M., Eisen, K., & Case, A. L. (2016). An angiosperm-wide analysis of the correlates of gynodioecy. *International Journal of Plant Sciences*, 177(2), 115-121.
- Charlesworth, D. (2002). Plant sex determination and sex chromosomes. *Heredity*, 88(2), Article 2. <https://doi.org/10.1038/sj.hdy.6800016>
- Charlesworth, D., & Laporte, V. (1998). The Male-Sterility Polymorphism of *Silene vulgaris*: Analysis of Genetic Data From Two Populations and Comparison With *Thymus vulgaris*. *Genetics*, 150(3), 1267–1282. <https://doi.org/10.1093/genetics/150.3.1267>
- Chase, C. D. (2007). Cytoplasmic male sterility: A window to the world of plant mitochondrial–nuclear interactions. *Trends in Genetics*, 23(2), 81–90. <https://doi.org/10.1016/j.tig.2006.12.004>
- Cho, Y., Mower, J. P., Qiu, Y. L., & Palmer, J. D. (2004). Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. *Proceedings of the National Academy of Sciences*, 101(51), 17741–17746. <https://doi.org/10.1073/pnas.0408302101>
- Cuguen, J., Wattier, R., Saumitou-Laprade, P., Forcioli, D., Mörchen, M., Van Dijk, H., & Vernet, P. (1994). Gynodioecy and mitochondrial DNA polymorphism in natural populations of *Beta vulgaris* ssp *maritima*. *Genetics Selection Evolution*, 26. <https://doi.org/10.1016/j.mito.2014.04.009>
- Dahan, J., & Mireau, H. (2013). The Rf and Rf-like PPR in higher plants, a fast-evolving subclass of PPR genes. *RNA Biology*, 10(9), 1469–1476. <https://doi.org/10.4161/rna.25568>
- David, P., Degletagne, C., Saclier, N., Jennan, A., Jarne, P., Plénet, S., Konecny, L., François, C., Guéguen, L., Garcia, N., Lefébure, T., & Luquet, E. (2022). Extreme mitochondrial DNA divergence underlies genetic conflict over sex determination. *Current Biology*, 32(10), 2325-2333.e6. <https://doi.org/10.1016/j.cub.2022.04.014>
- de Haan, A. A. D., Luyten, Bakx-Schotman, & van Damme,. (1997). The dynamics of gynodioecy in *Plantago lanceolata* L. I. Frequencies of male-steriles and their cytoplasmic male sterility types. *Heredity*, 79(5), 453–462.
- Del Castillo, R. F., & Trujillo, S. (2009). Evidence of restoration cost in the annual gynodioecious *Phacelia dubia*. *Journal of Evolutionary Biology*, 22(2), 306–313. <https://doi.org/10.1111/j.1420-9101.2008.01644.x>
- Delph, L. F., Touzet, P., & Bailey, M. F. (2007). Merging theory and mechanism in studies of gynodioecy. *Trends in Ecology & Evolution*, 22(1), 17–24. <https://doi.org/10.1016/j.tree.2006.09.013>
- Dillon Jr., R. T., Wethington, A. R., Rhett, J. M., & Smith, T. P. (2002). Populations of the European freshwater pulmonate *Physa acuta* are not reproductively isolated from American *Physa heterostropha* or *Physa integra*. *Invertebrate Biology*, 121(3), 226–234. <https://doi.org/10.1111/j.1744-7410.2002.tb00062.x>
- Draparnaud JPR 1805. *Histoire Naturelle des Mollusques Terrestres et Fluviatiles de la France*, D Colas, Paris.
- Dufay, M., & Billard, E. (2012). How much better are females? The occurrence of female advantage, its proximal causes and its variation within and among gynodioecious species. *Annals of Botany*, 109(3), 505-519.
- Ebbs, E. T., E. S. Loker, and S. V Brant. 2018. Phylogeography and genetics of the globally invasive snail *Physa acuta* Draparnaud 1805 , and its potential to serve as an intermediate host to larval digenetic trematodes. *BMC Evol. Biol.* 18:103. *BMC Evolutionary Biology*.

- Ebert, D., & Fields, P. D. (2020). Host–parasite co-evolution and its genomic signature. *Nature Reviews Genetics*, 21(12), Article 12. <https://doi.org/10.1038/s41576-020-0269-1>
- Ehlers, B. K., Maurice, S., & Bataillon, T. (2005). Sex inheritance in gynodioecious species: A polygenic view. *Proceedings of the Royal Society B: Biological Sciences*, 272(1574), 1795–1802. <https://doi.org/10.1098/rspb.2005.3168>
- Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). *DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates*.
- Fujii, S., & Small, I. (2011). The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytologist*, 191(1), 37–47. <https://doi.org/10.1111/j.1469-8137.2011.03746.x>
- Galtier, N. The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC Biol* 9, 61 (2011). <https://doi.org/10.1186/1741-7007-9-61>
- Gautam, A. (2022). DNA Isolation by Chelex Method. In A. Gautam (Ed.), *DNA and RNA Isolation Techniques for Non-Experts* (pp. 79–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-94230-4_10
- Gouyon, P. H., Vichot, F., & Van Damme, J. M. M. (1991). Nuclear-Cytoplasmic Male Sterility: Single-Point Equilibria Versus Limit Cycles. *The American Naturalist*, 137(4), 498–514. <https://doi.org/10.1086/285179>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644–652. <https://doi.org/10.1038/nbt.1883>
- Guéguen, L., Gaillard, S., Boussau, B., Gouy, M., Groussin, M., Rochette, N. C., Bigot, T., Fournier, D., Pouyet, F., Cahais, V., Bernard, A., Scornavacca, C., Nabholz, B., Haudry, A., Dachary, L., Galtier, N., Belkhir, K., & Dutheil, J. Y. (2013). Bio++: Efficient Extensible Libraries and Tools for Computational Molecular Evolution. *Molecular Biology and Evolution*, 30(8), 1745–1750. <https://doi.org/10.1093/molbev/mst097>
- Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—A baiting and iterative mapping approach. *Nucleic Acids Research*, 41(13), e129. <https://doi.org/10.1093/nar/gkt371>
- Innocenti, P., Morrow, E. H., & Dowling, D. K. (2011). Experimental evidence supports a sex-specific selective sieve in mitochondrial genome evolution. *Science*, 332(6031), 845–848.
- Jarne, P., & Auld, J. R. (2006). Animals mix it up too: The distribution of self-fertilization among hermaphroditic animals. *Evolution*, 60(9), 1816–1824. <https://doi.org/10.1111/j.0014-3820.2006.tb00525.x>
- Kamps, T. L., & Chase, C. D. (1997). RFLP mapping of the maize gametophytic restorer-of-fertility locus (rf3) and aberrant pollen transmission of the nonrestoring rf3 allele: *Theoretical and Applied Genetics*, 95(4), 525–531. <https://doi.org/10.1007/s001220050593>
- Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066. <https://doi.org/10.1093/nar/gkf436>

- Kim, Y.-J., & Zhang, D. (2018). Molecular Control of Male Fertility for Crop Hybrid Breeding. *Trends in Plant Science*, 23(1), 53–65.
<https://doi.org/10.1016/j.tplants.2017.10.001>
- Lartillot, N., Lepage, T., & Blanquart, S. (2009). PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, 25(17), 2286–2288. <https://doi.org/10.1093/bioinformatics/btp368>
- Lewis, D. (1941). Male Sterility in Natural Populations of Hermaphrodite Plants. The Equilibrium Between Females and Hermaphrodites to be Expected with Different Types of Inheritance. *The New Phytologist*, 40(1), 56–63.
- Li H.*, Handsaker B.*, Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment map (SAM) format and SAMtools. *Bioinformatics*, 25, 2078-9.
- Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25:1754-60
- Li, D., Liu, C-M., Luo, R., Sadakane, K., and Lam, T-W., (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, doi: 10.1093/bioinformatics/btv033.
- Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H. and Lam, T.W., 2016. MEGAHIT v1.0: A Fast and Scalable Metagenome Assembler driven by Advanced Methodologies and Community Practices. *Methods*.
- Li, S., YANG, G., LI, S., LI, Y., CHEN, Z., & ZHU, Y. (2005). Distribution of Fertility-restorer Genes for Wild-abortive and Honglian CMS Lines of Rice in the AA Genome Species of Genus *Oryza*. *Annals of Botany*, 96(3), 461–466.
<https://doi.org/10.1093/aob/mci197>
- Lloyd, D. G. (1975). The maintenance of gynodioecy and androdioecy in angiosperms. *Genetica*, 45(3), 325–339. <https://doi.org/10.1007/BF01508307>
- McCauley, D. E., & Bailey, M. F. (2009). Recent advances in the study of gynodioecy: The interface of theory and empiricism. *Annals of Botany*, 104(4), 611–620.
<https://doi.org/10.1093/aob/mcp141>
- Mower, J. P., Touzet, P., Gummow, J. S., Delph, L. F., & Palmer, J. D. (2007). Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. *BMC Evolutionary Biology*, 7(1), 135. <https://doi.org/10.1186/1471-2148-7-135>
- Murayama, K., Yahara, T., & Terachi, T. (2004). Variation of female frequency and cytoplasmic male-sterility gene frequency among natural gynodioecious populations of wild radish (*Raphanus sativus* L.). *Molecular Ecology*, 13(8), 2459–2464.
<https://doi.org/10.1111/j.1365-294X.2004.02231.x>
- Murlas Cosmides, L., & Tooby, J. (1981). Cytoplasmic inheritance and intragenomic conflict. *Journal of Theoretical Biology*, 89(1), 83–129. [https://doi.org/10.1016/0022-5193\(81\)90181-8](https://doi.org/10.1016/0022-5193(81)90181-8)
- Ng, T. H., S. K. Tan, and D. C. J. Yeo. 2015. Clarifying the identity of the long-established, globally-invasive *Physa acuta* Draparnaud, 1805 (Gastropoda: Physidae) in Singapore. *BioInvasions Rec.* 4:189–194.
- Noël, E., Chemtob, Y., Janicke, T., Sarda, V., Pélissié, B., Jarne, P., & David, P. (2016). Reduced mate availability leads to evolution of self-fertilization and purging of inbreeding depression in a hermaphrodite. *Evolution*, 70(3), 625-640.
- Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. *PLOS ONE*, 6(9), e22594.
- Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.

- Patel, M. R., Miriyala, G. K., Littleton, A. J., Yang, H., Trinh, K., Young, J. M., ... & Malik, H. S. (2016). A mitochondrial DNA hypomorph of cytochrome oxidase specifically impairs male fertility in *Drosophila melanogaster*. *elife*, 5, e16923.
- Perlman, S. J., Hodson, C. N., Hamilton, P. T., Opit, G. P., & Gowen, B. E. (2015). Maternal transmission, sex ratio distortion, and mitochondria. *Proceedings of the National Academy of Sciences*, 112(33), 10162-10168.
- Romiguier, J., Figuet, E., Galtier, N., Douzery, E. J. P., Boussau, B., Dutheil, J. Y., & Ranwez, V. (2012). Fast and Robust Characterization of Time-Heterogeneous Sequence Evolutionary Processes Using Substitution Mapping. *PLOS ONE*, 7(3), e33852. <https://doi.org/10.1371/journal.pone.0033852>
- Saumitou-Laprade, P., Cuguen, J., & Vernet, P. (1994). Cytoplasmic male sterility in plants: Molecular evidence and the nucleocytoplasmic conflict. *Trends in Ecology & Evolution*, 9(11), 431–435. [https://doi.org/10.1016/0169-5347\(94\)90126-0](https://doi.org/10.1016/0169-5347(94)90126-0)
- Shykoff, J. A., Kolokotronis, S.-O., Collin, C. L., & López-Villavicencio, M. (2003). Effects of male sterility on reproductive traits in gynodioecious plants: A meta-analysis. *Oecologia*, 135(1), 1–9. <https://doi.org/10.1007/s00442-002-1133-z>
- Sisco, P. H. (1991). Duplications Complicate Genetic Mapping of Rf4, a Restorer Gene for cms-C Cytoplasmic Male Sterility in Corn. *Crop Science*, 31(5), crops1991.0011183X003100050036x. <https://doi.org/10.2135/crops1991.0011183X003100050036x>
- Snyder, R. J., & Duvick, D. N. (1969). Chromosomal Location of Rf2, a Restorer Gene for Cytoplasmic Pollen Sterile Maize1. *Crop Science*, 9(2), crops1969.0011183X000900020012x. <https://doi.org/10.2135/crops1969.0011183X000900020012x>
- Touzet, P. (2012). Chapter Four—Mitochondrial Genome Evolution and Gynodioecy. In L. Maréchal-Drouard (Ed.), *Advances in Botanical Research* (Vol. 63, pp. 71–98). Academic Press. <https://doi.org/10.1016/B978-0-12-394279-1.00004-1>
- Tsitrone, A., Jarne, P., & David, P. (2003). Delayed Selfing and Resource Reallocations in Relation to Mate Availability in the Freshwater Snail *Physa acuta*. *The American Naturalist*, 162(4), 474–488. <https://doi.org/10.1086/376889>
- Weeks, S. C. (2012). The role of androdioecy and gynodioecy in mediating evolutionary transitions between dioecy and hermaphroditism in the Animalia. *Evolution*, 66(12), 3670–3686. <https://doi.org/10.1111/j.1558-5646.2012.01714.x>
- Young, M. K., Smith, R., Pilgrim, K. L., & Schwartz, M. K. (2021). Molecular species delimitation refines the taxonomy of native and nonnative physine snails in North America. *Scientific reports*, 11(1), 21739.
- Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. *Current Protocols in Bioinformatics*, 69(1). <https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cpbi.96>

Supplementary Text 1 – Material and Methods

The species

P. acuta (Physidae, Hygrophila, Gastropoda) is a freshwater snail native to North America with a cosmopolitan distribution that lives in various freshwater habitats (Dillon et al. 2002). Individuals lay egg capsules that hatch in 7-10 days and reach sexual maturity in six to eight weeks at 25°C. *P. acuta* is a simultaneous hermaphrodite and preferential outcrosser; sperm and oocytes are produced within the same gland (ovotestis) and, unlike plants, external morphology is identical between male-sterile and male-fertile individuals – the difference can only be made on the basis of the ability to sire offspring in controlled pairings, or on the basis of dissection and sperm counts (David et al. 2022). This species is a preferential outcrosser and two male-fertile individuals, especially if both are virgin, usually mate rapidly when paired (for example, 70% of them had copulated as males (and 88.6% as females) in the first 45mn after encountering a virgin mate in David et al. 2022). Matings are one-sided: at a given time, one individual is acting as a male and the other as a female; but they often switch roles for the next mating, that can take place immediately after the first one. Self-fertilization can occur as a reproductive insurance strategy when sexually mature individuals are deprived of mates; in that case, the onset of reproduction is delayed by about two weeks (“waiting time”) compared to individuals with available mates (Tsitrone et al 2003).

Samples studied

The *P. acuta* snails studied here were sampled in October 2019 at 4 sites near Lyon (Saint-Priest, Erevan, Crêt-2, and Irigny, Table S1), around the population where the male-sterile D mitotype was initially found (Lyon, France). Some of the snails were frozen to extract DNA while others were kept alive. At the same time, we carried out a whole-genome sequencing analysis on several French populations, and found individuals with mitotype K in one of them: “Canet des Maures” (CDM, Table S1). One individual from the CDM population was used to reconstruct the complete mitochondrial genome of type K (see below). For the introgression experiment (see below), we also used a laboratory albino population of *P. acuta*, maintained as a large outbred population for >80 generations, and initially derived from a mix of ten natural populations from Montpellier (300km south of Lyon, Noël et al. 2016). This population is composed only of male-fertile, N-mitotype individuals with an albino phenotype. Albinism is due to a single recessive allele, body colour can thus be used to ascertain paternity: when an albino mother produces a pigmented juvenile, the father must be a pigmented individual. During all experiments, the snails were maintained in 75-ml plastic boxes at 25°C in groundwater under a 12:12 photoperiod and fed ad libitum with boiled organic lettuce. Water was changed once or twice a week.

Introgression experiments

The aim of the introgression protocol was to progressively replace the nuclear DNA originating from wild-caught ancestors from Lyon (focal), by the “naïve” nuclear background of the albino N population from Montpellier (partner), while keeping the ancestral (N or K) mitochondrial lineage from Lyon. The introgressed lines (“matrilines”) were propagated by maternal descent, keeping only pigmented individuals, while the sperm donors came at each generation from the albino population. From G1 to G17 the same protocol was used; focal (individual from the matriline) were pigmented and their partners albino, the male-fertility status of each focal was assessed at each generation on the basis of the offspring of albino partners.

Generating the focal mothers. To generate the first generation of focal individuals, we placed all adults from Erevan and Irigny (i.e., G0 individuals, Table S1) in one large aquarium per population for one week, so that each individual had many opportunities to be inseminated.

We then let individuals lay separately in individual boxes and raised the offspring used as focal in the first generation of introgression (G1).

Introgression. Introgression occurred within “matrilines”, defined as sets of focal individuals sharing the same ancestor by maternal descent, and hence the same mitochondrial DNA; there were 25 matrilines (12 K and 13 N). At each generation, individuals from each matriline (focals) were kept virgin and, once mature, paired with a virgin N-mitotype albino partner. Eggs laid by the focal were then used to generate the next generation. After the first generation focal individuals were heterozygous at the albinism locus and produced a mix of albino and pigmented juveniles in their progeny. We propagated the matrilines by keeping only pigmented individuals, to distinguish them from their mates. We usually kept three pigmented individuals in each matriline at each generation to avoid accidental loss of matrilines due to mortality or failure to reproduce. This process was continued for 17 successive generations.

Assessing male-fertility status. The male fertility status of focal individuals was assessed at each generation on the basis of the offspring laid by albino mates: the presence of pigmented juveniles indicated male-fertility. Starting from the G2, the focal was heterozygous at the albinism locus, thus from the third generation on, a 100% siring success resulted in a mix of 50% albino and 50% pigmented juveniles.

Monitoring mitotypes. All the G0 individuals were sacrificed and their mitotype was determined using PCR tests (Figure S2). Then, throughout the experiment we tested again individuals at G1 (N = 75, three offspring per G0 mother), G2 (N = 50, all the focals and their albino partner) and G7 (N = 50, two offspring per focal mother). In parallel we also constituted, using excess individuals in G1, two large populations by pooling K offspring in one aquarium, and N offspring in another. The aquaria were conserved to this day as separate populations and 24 individuals sampled in each aquarium were mitotyped in G25.

Behavioral and reproductive traits of K and N individuals

In the G1 and G5 generations of the introgression protocol, which respectively represent matrilines before and after introgression, we assessed the sexual phenotype of matriline individuals. Following a methodology akin to that employed for mitotype D, we evaluated both behavioral and fitness traits. Two to three individuals from each N or K matriline were isolated until reaching sexual maturity at 50 days. The assessment of behavioral traits took place during 45-minute observations of mating pairs, starting immediately after pairing. Each pair comprised one pigmented focal individual from the matriline and one N-mitotype albino mate, both partners being virgin prior to mating. The behaviors of both male and female individuals were recorded and timed, focusing on the climbing and sitting actions of the male-acting individual. Subsequent to the observations, the snails were left together for three days, re-isolated for egg-laying in separate boxes for a further three days, and then removed.

The male-fertility status was established through the snails ability to (i) induce oviposition in their virgin partners (the normal consequence of insemination of a virgin partner) and (ii) sire a majority of the offspring laid by the partner. Male-sterile individuals either failed to inseminate, so that the partner produced no or purely self-fertilized (100% albino) offspring, or inseminated their partner but with very little sperm, so that very few outcrossed offspring were produced. This second case, which we defined as less than 10 outcrossed offspring, was termed 'semi-sterility' and could occur either because the outcrossed offspring represented a very low proportion of all offspring or because the number of eggs laid was abnormally low. Complete male sterility was declared when absolutely no pigmented offspring were produced. In the remaining individuals, more than half of the partner's offspring were outcrossed, with a total exceeding 10; these individuals were called male-fertile. In G1, because focal individuals were homozygous pigmented, this number of outcrossed offspring was obtained directly by counting pigmented juveniles. In G5, because focal individuals were heterozygous, the criteria

were adjusted to more than 25% and 5 pigmented juveniles (representing an expected 50% or 10 outcrossed juveniles respectively). In the main text, we report results obtained by contrasting male-fertile with pooled male-sterile and male-semisterile phenotypes. However the semisterile individuals are not numerous and the results remain qualitatively the same when they are excluded or pooled with the fertile category.

Mitochondrial genome sequence

To reconstruct the complete mitochondrial genome of type K, a whole genome sequencing was performed on one individual (CDM029) sample in a population sampled at Canet des Maures (Var, France). The DNA libraries were prepared at the GenSeq technical platform (University of Montpellier, France) using an Illumina kit (*Nextera DNA Flex*). The libraries were sequenced at Montpellier GenomiX (Montpellier, France) with an NovaSeq6000 lane and validated using DNA quantification (concentration and fragment size) on Fragment Analyzer (Standard Sensitivity NGS kit) and qPCR (ROCHE Light Cycler 480). The genome was assembled using MEGAHIT (Li et al. 2015, 2016).

To ensure the absence of contamination and chimera in the assembly of the K mitochondrial genome, coverage was checked along the assembled contig. To do this, short pair-end reads were mapped onto the assembled genome using BWA (Li and Durbin, 2009). The resulting SAM file was converted to BAM using Samtools (Li et al. 2009), then coverage was measured on the mitochondrial contig using the function depth of Samtools. While the average genome coverage is 25X (sd=57), the average coverage of the mitochondrial genome is 555X (sd=63) and is constant throughout the mitochondrial genome (Figure S3).

In order to assess the variability of the K, transcriptomes were sequenced for ten cultivated individuals carrying the K mitotype. We bred ten K-mitotype individuals (three from Irigny and six from Erevan; from different G0 mothers with the long-long band pattern), and sent them for Illumina transcriptome sequencing (extraction and sequencing protocol as in David et al. (2022)). Using Trimmomatic (v 0.39, Bolger et al. 2014), adapters were clipped from the sequences, reads were scanned with a 4-base sliding window and trimmed when the average quality per base was below 15. Resulting reads longer than 36 bases were kept. Paired-end transcriptomes were assembled de novo using Trinity v2.3.2 (Grabherr et al., 2011). Mitochondrial genes of *P. acuta* N were blasted against the assembled transcriptomes in order to find mitochondrial contigs. The longest and most expressed contigs were kept and annotated using the MITOS web server (Bernt et al., 2013) to obtain sequences of protein-coding genes. The genes were aligned using the MAFFT software (Katoh et al. 2002) and their divergence calculated in R using the APE package (Parais *et al.* 2004). As the sequences of the ten K individuals were nearly identical (max 0.2% divergence), we assumed that they belonged to the same mitotype and used only one of them to determine their position in the phylogeny. For the phylogenetic analysis (*Physidae* phylogeny), we used the sequences from the individual E3A1 because the sequences were of better quality (few indeterminate bases). Transcriptome reads and assembly have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB50799. We carried out blasts on the 10 transcriptomes to recover the sequences of mitochondrial protein-coding genes. These blasts did not reveal any cases of heteroplasmy. In fact, in each case where two different contigs were blasted, the second contig had extremely low coverage compared with the first, compatible with cross-contamination. No cases with two normally covered contigs were found.

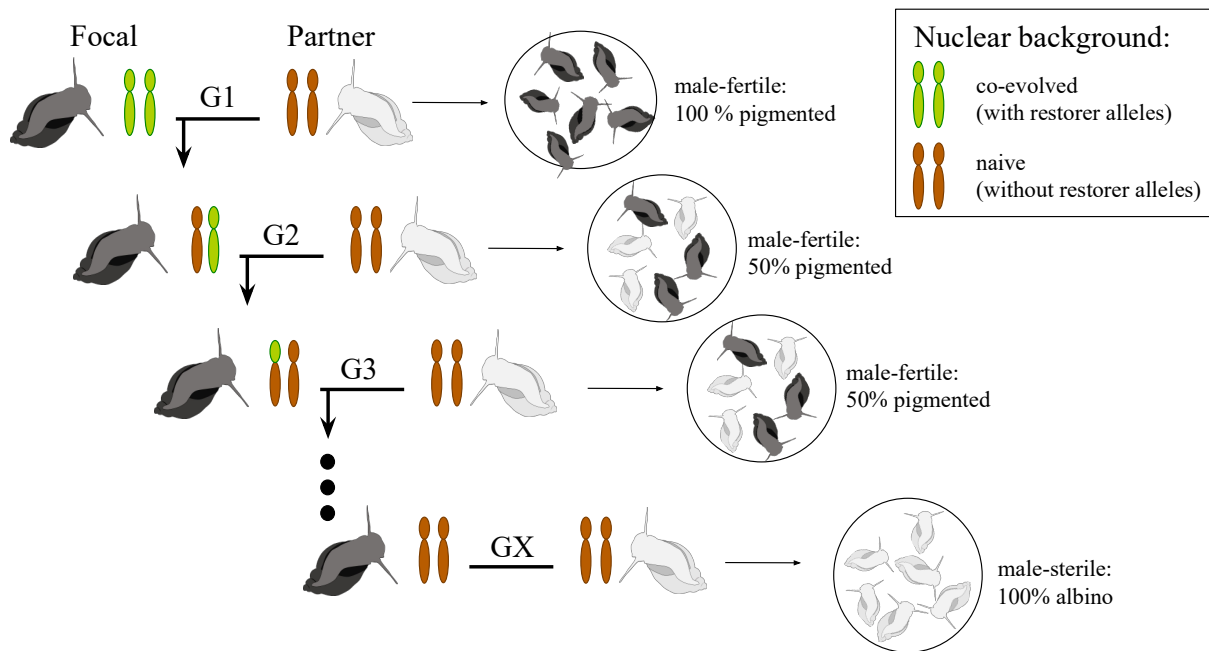


Figure S1: Introgression protocol. Nuclear DNA of the mitochondrial K originating from wild-caught ancestors from Lyon (called “co-evolved”), is progressively diluted into the “naïve” nuclear background of the N albino population kept in Montpellier, that provides sperm donors at each generation. Fathers are of the albino phenotype, mothers are pigmented. The G1 is 100% of Lyon origin (first-generation offspring of wild-caught mothers and fathers), introgressed generations are from G2 to G17. From the G2 on, mothers carry the albino allele in heterozygous form, inherited from the father. From the G3 on, offspring are a mix of albino and pigmented (heterozygous) juveniles at each generation, among which one pigmented individual is selected to become the mother of the next generation. The protocol was continued to G17.

Table S1: *Physa acuta* populations sampled in October 2019 near Lyon (France). Proportions of individuals carrying each of three mitotypes (K, N and D) are assessed from PCR tests made on wild-caught individuals (see Figure S2).

Site	Latitude	Longitude	K proportion	N proportion	D proportion	number of individuals tested
Irigny	N 45°68'20.1"	E4°83'38.7"	12 %	85 %	3%	100
Erevan	N 45°74'42.1"	E 4°81'58.4"	32 %	68 %	0%	124
Crêt-2	N45°80'47.6"	E4°99'69.8"	0 %	100 %	0%	125
Saint-Priest	N45°71'72.3"	E4°93'01.8"	0 %	100 %	0%	114

Table S2: Accessions of mitochondrial sequences used in this study. Section 1: Accessions of complete mtDNA sequences that were used in Fig. 1A, in combination with our own sequences of *P. acuta* from Lyon. For each sample, elements of taxonomy are provided (from lower to higher: family, clade 2, clade 1). For *Physa gyrina*, we reassembled the mtDNA from the available transcriptome reads; the accession provided is that of the reads in the SRA database. **Section 2:** Accessions of COI sequences used in Fig. 1B. **Section 3:** Accession of COI sequences used in Fig. 1C. Shaded accessions correspond to sequences from this study.

Section 1: complete mt DNA sequences used in Figure 1A

Accession	Genus	species	Family	Clade 2	Clade 1
KP098541	Radix	balthica	Lymnaeidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
KP098540	Radix	auricularia	Lymnaeidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
JN564796	Galba	pervia	Lymnaeidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
MT862415	Planorbarius	corneus	Planorbidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
AY380567	Biomphalaria	glabrata	Planorbidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
SRX565294	Physa	gyrina	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
JQ390525	Physa	acuta (isolate A)	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
JQ390526	Physa	acuta (isolate	Physidae	Hygrophila	Heterobranchia,

		B)			Euthyneura, Panpulmonata
MT483691	Physa	fontinalis	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
PRJEB50799	Physa	acuta (Lyon)	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
PRJEB50799	Physa	acuta (Lyon)	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
PRJEB50799	Physa	acuta (Lyon)	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
PRJEB50799	Physa	acuta D	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata
PRJEB50799	Physa	acuta K	Physidae	Hygrophila	Heterobranchia, Euthyneura, Panpulmonata

Section 2: COI sequences used in Figure 1B

Accession	Genus	species	type	Family
PRJEB50799	Physa	acuta (Lyon)	N	Physidae
PRJEB50799	Physa	acuta (Lyon)	N	Physidae
PRJEB50799	Physa	acuta (Lyon)	N	Physidae
JQ390525	Physa	acuta (USA, A)	N	Physidae
JQ390526	Physa	acuta (USA, B)	N	Physidae
PRJEB50799	Physa	acuta	D	Physidae
MT483691	Physa	fontinalis		Physidae
KP182981.1	Physa	acuta (Singapore)	K	Physidae
KP182982.1	Physa	acuta (Singapore)	K	Physidae
ON953193.1	Physa	acuta (South Africa)	K	Physidae
ON953200.1	Physa	acuta (South Africa)	K	Physidae
OK510774.1	Physa	acuta (North America)	K	Physidae
OK510777.1	Physa	acuta (North America)	K	Physidae
PRJEB50799	Physa	acuta (Erevan)	K	Physidae
PRJEB50799	Physa	acuta (Erevan)	K	Physidae
PRJEB50799	Physa	acuta (Erevan)	K	Physidae
PRJEB50799	Physa	acuta (Erevan)	K	Physidae
PRJEB50799	Physa	acuta (Erevan)	K	Physidae
PRJEB50799	Physa	acuta (Irigny)	K	Physidae
PRJEB50799	Physa	acuta (Irigny)	K	Physidae

Section 3: COI sequences of gastropods used in Figure 1C

num	accession	species	species short	genus	family	cat	dist	Pacuta	clade2	clade1	subclass	type
NC_023253	Physa_acuta		acuta	Physa	Physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
JQ390526	Physa_acuta		acuta	Physa	Physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
GU247996	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
GU247995	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
GU247993	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651203	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651193	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651192	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651188	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651186	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651185	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651181	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651174	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY651203	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
AY282589	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038389	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038372	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038371	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038368	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038367	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038366	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038365	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038361	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038360	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038357	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038356	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
EU038355	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769128	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769127	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769126	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769125	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769124	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KJ769123	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
JQ390526	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
JQ390525	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N
KT280445	Physa_acuta		acuta	Physa	physidae			0	Hygrophila	Panpulmonata	Heterobranchia	N

[illegible]

EU038377	Physa(Aplexa)_elongata	elongata	Physa(Aplexa)	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038375	Physa(Aplexa)_elongata	elongata	Physa(Aplexa)	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038363	Physa_pomilia	pomilia	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038354	Physa_pomilia	pomilia	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038353	Physa_pomilia	pomilia	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651196	Physa_hendersoni	hendersoni	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651195	Physa_hendersoni	hendersoni	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651194	Physa_hendersoni	hendersoni	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038398	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038374	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038373	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651200	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651199	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651197	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651191	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651187	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651182	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651178	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651173	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651172	Physa_gyrina	gyrina	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038369	Physa(Stenophysa)_marmorata	marmorata	Physa(Stenophysa)	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
EU038370	Physa(Stenophysa)_marmorata	marmorata	Physa(Stenophysa)	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651190	Physa_fontinalis	fontinalis	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651189	Physa_fontinalis	fontinalis	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651205	Physa_spelunca	spelunca	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
AY651204	Physa_spelunca	spelunca	Physa	physidae	1	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_01	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_02	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_03	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_04	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_05	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_06	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_07	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_08	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	D
P_ac_09	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_10	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_11	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_12	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_13	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	D

P_ac_14	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_15	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_16	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_17	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_18	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_19	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_21	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_22	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_23	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_24	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	D
P_ac_25	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_26	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_27	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_28	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_29	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	D
P_ac_30	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_31	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_32	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_33	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	D
P_ac_34	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
P_ac_35	Physa_acuta	acuta	Physa	Physidae	0	Hygrophila	Panpulmonata	Heterobranchia	N
MH087513	Pseudosuccinea_columella	columella	Pseudosuccinea	Lymnaeidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
KM612193	Lymnaea_stagnalis	stagnalis	Lymnaea	Lymnaeidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
MN601426	Galba_mweruensis	mweruensis	Galba	Lymnaeidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
AY651207	Biomphalaria_obstructa	obstructa	Biomphalaria	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
MF458795	Planorbarius_corneus	corneus	Planorbarius	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
KM612167	Planorbella(Helisoma)_trivolis	trivolis	Planorbella(Helisoma)	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
KF958031	Planorbella(Helisoma)_campanulata	campanulata	Planorbella(Helisoma)	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
EU038397	Gyraulus_parvus	parvus	Gyraulus	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N
EU038390	Glyptophysa_sp	sp	Glyptophysa	Planorbidae	2	Hygrophila	Panpulmonata	Heterobranchia	N

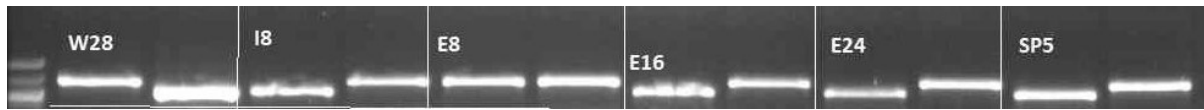


Figure S2: One example of pattern found using the “N versus D” PCR test from David et al. 2022. This test uses two separate amplifications to produce either a long band- short band pattern typical of the D mitotype (example: individual W28, an individual from the Lyon Autoroute population), or a short-long pattern typical of the N mitotype (I8, an individual from Irigny; E16 and E24, two individuals from Erevan; SP5, an individual from Saint-Priest). The E8 individual (from the Erevan population) shows a long-long pattern indicating that none of the internal primers specific to D or N has been able to block the amplification of the long fragment by the external, universal primers. Individuals such as E8 turned out, after sequencing, to belong to a third very divergent mitotype named K.

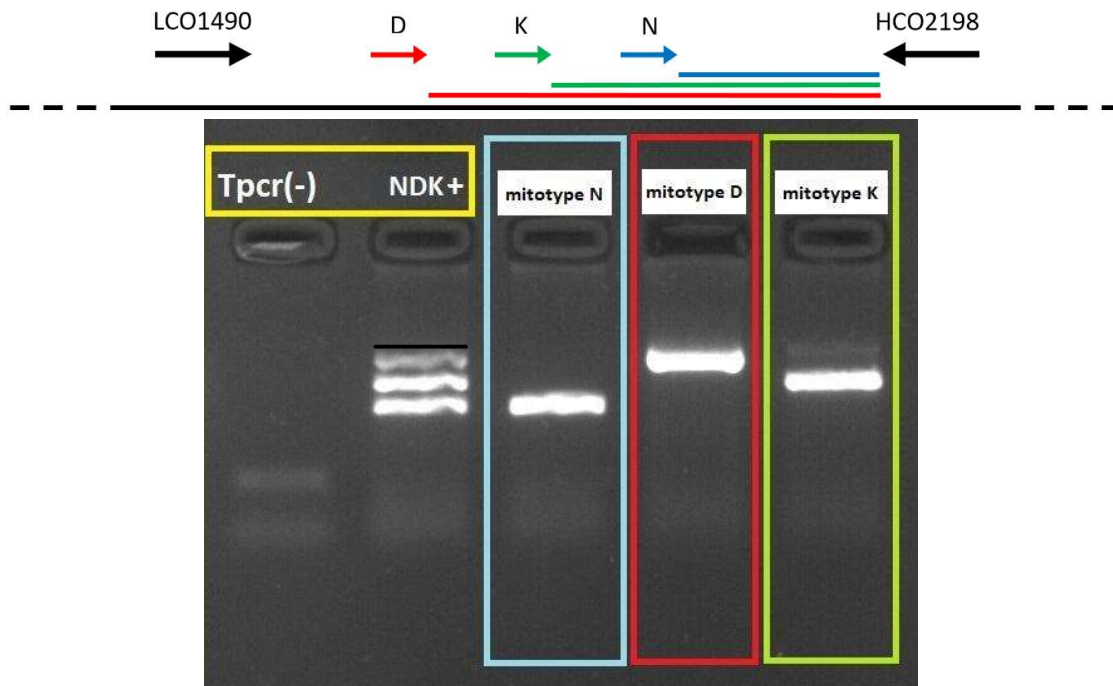


Figure S3: Mitotype identification by PCR. One PCR was performed per individual using five primers per PCR, the two external, universal COI “Folmer” primers and three internal specific primers (ATTACACTTAGCTGGGTTATCA for N, GGGACAACCTAGGGTCTTAACCTAC for D and TTGCTCCCTCCTTCTTTCACA for K mitotypes, blue, red and green respectively). PCR products were run on an agarose gel (2%) giving a typical pattern depending on the mitotype (an example of gel with individuals from three maternal lines of each mitotype is given). Short fragments span 288 bp for N, 607 bp for D and 426 bp for K respectively; if none of the N, D, or K primers matches a 705 bp fragment is amplified (represented by the black line). Amplifications were carried out in 10- μ L final reaction volume comprising 5- μ L of 2X Sigma-Aldrich REDExtract-N-Amp PCR ReadyMix, 0.25- μ L of each primer (10 μ M), 1.75- μ L of PCR-grade Water. The PCR conditions were as follows: 5 min at 95°C, followed by 35 cycles of denaturation for 30 sec at 94°C, annealing for 30 sec at 50°C and extension for 40 sec at 72°C, completed with a final extension step for 10 min at 72°C.

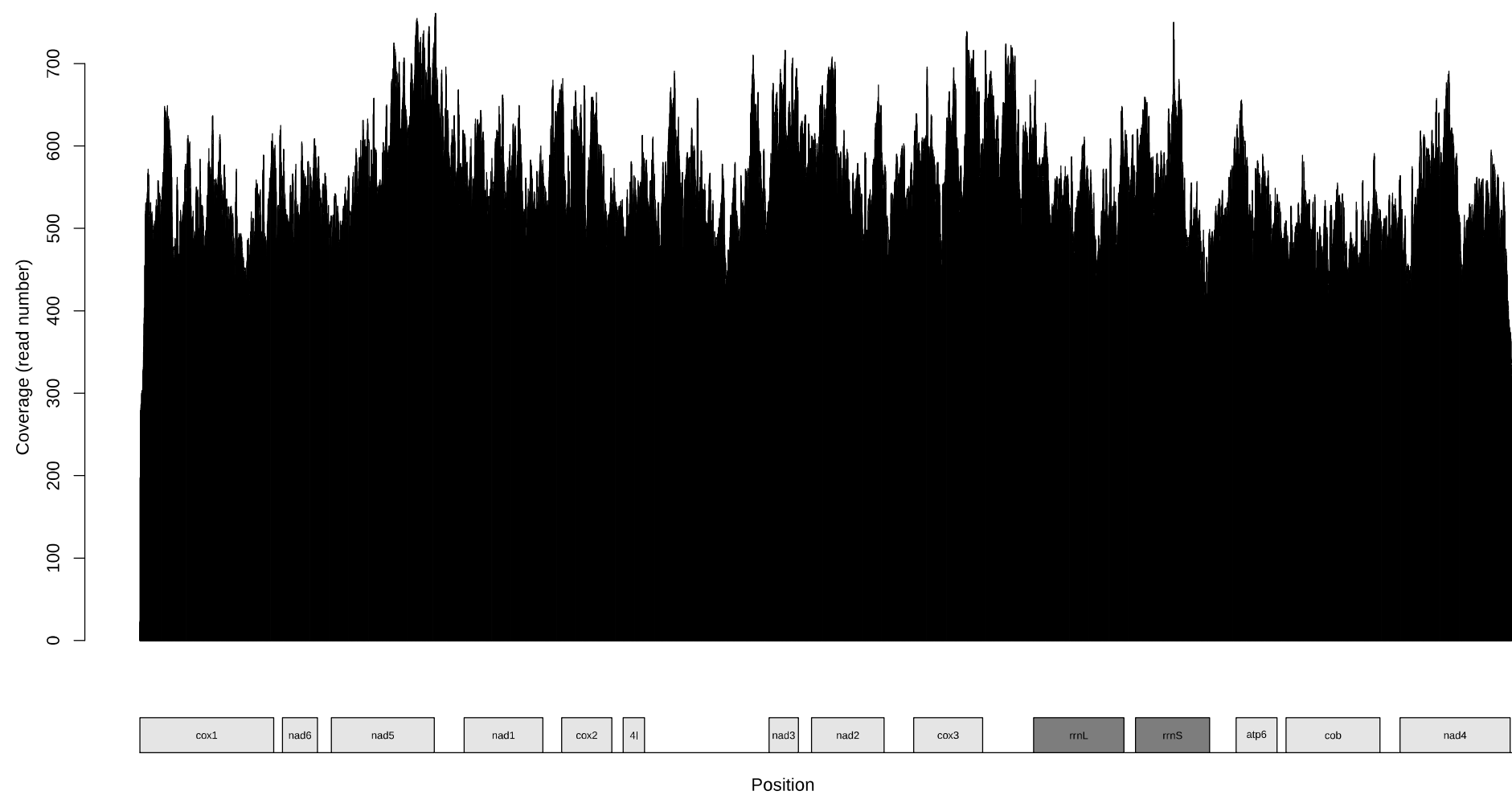


Figure S4: Coverage along the K mitochondrial genome.

Behavioral and copulation data

Traits	Generations	Number of pairs	Mean \pm se success / failures		Test of type effect df = 1	Random effects included	Model type
			Type N	Type K			
Male time	G1	63	0.29 \pm 0.04	0.30 \pm 0.03	$\chi^2 = 0.001$, P = 0.98	Matriline	Gaussian
	G5	52	0.30\pm0.03	0.12\pm0.03	$\chi^2 = 13.85$, P = 1.9 10⁻⁵		
Male copulation (yes/no)	G1	63	28/7	24/4	$\chi^2 = 0.10$, P = 0.75	Matriline, ind	Binomial
	G5	52	22/4	10/16	$\chi^2 = 19.56$, P = 9.8 10⁻⁶		
Female time	G1	63	0.25 \pm 0.03	0.21 \pm 0.03	$\chi^2 = 0.78$, P = 0.38	Matriline	Gaussian
	G5	52	0.26 \pm 0.03	0.30 \pm 0.04	$\chi^2 = 0.88$, P = 0.35		
Female copulation (yes/no)	G1	63	25/10	19/9	$\chi^2 = 0.13$, P = 0.71	Matriline, ind	Binomial
	G5	52	21/5	22/4	$\chi^2 = 0.44$, P = 0.51		
Fitness traits							
Number of pigmented offspring laid by the albino partner	G1	63	60.74 \pm 6.11 (37/4)	49.07 \pm 5.81 (22/5)	$\chi^2 = 1.85$, P = 0.17	Matriline	Gaussian
	G5	52	21.50\pm1.86 (25/1)	6.00\pm1.82 (8/18)	$\chi^2 = 35.58$, P = 2.44 10⁻⁹		
Female fitness	G1	63	55.00 \pm 4.02	53.00 \pm 4.85	$\chi^2 = 0.06$, P = 0.81	Matriline	Gaussian
	G5	52	65.92 \pm 3.49	62.50 \pm 4.04	$\chi^2 = 0.16$, P = 0.69		

Table S3: Results of linear models on behavioral, copulation and reproductive traits of *P. acuta* K and N at G1 and G5. Performances of N and K mitotypes are reported either as trait means and SE (for quantitative traits) or as successes / failures (success = at least one copulation observed, or at least one egg or juvenile produced, failure = no copulation, egg or juvenile). After the first generation focal individuals were heterozygous at the albinism locus and produced a mix of albino and pigmented juveniles in their progeny. Thus the number of pigmented offspring was divided by two from G1 to G5 for the N individuals. The last columns indicate the random effects included in the model and the distribution of the variable. An “individual” (ind) random effect was added when appropriate in order to account for overdispersion (GLMMs). The “matriline” random effect accounts for correlations between related individuals that have inherited copies of the same mitogenome by maternal descent from a common ancestor collected in natura. ^(a)The difference in female fertility between K and N individuals in G5 remained non-significant (means \pm SE : N : 65.88 \pm 3.63 ; K : 60.22 \pm 4.89 ; LRT test $\chi^2 = 0.69$, P = 0.41) when only male-sterile K individuals and male-fertile N individuals were included in the comparison.

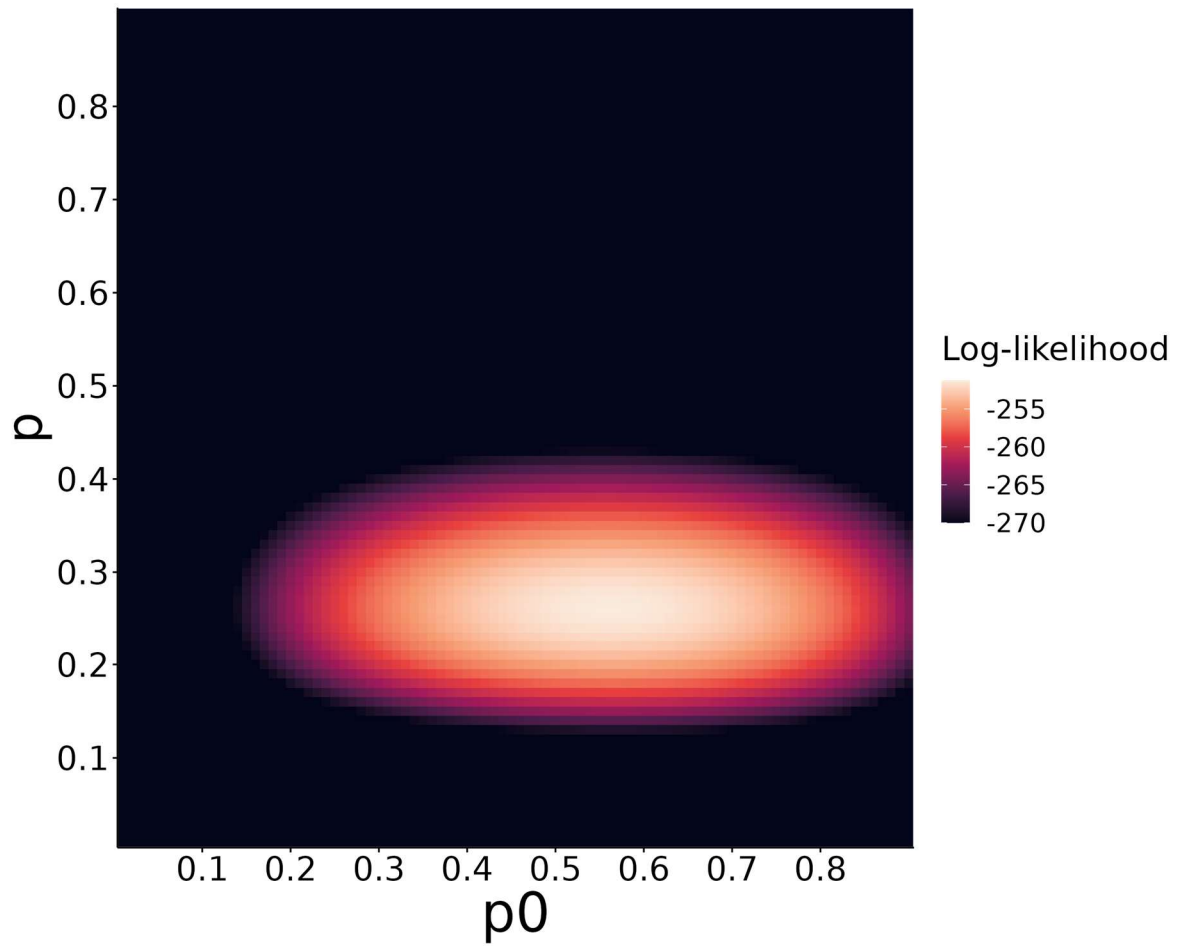


Figure S5: Maximum likelihood estimation for the restorer frequency in the focal population K (p_0) and in the albino population (p). Log-likelihood surface in the (p_0, p) plane.

Table S4: Similar sequences found by blasting the COI sequence of the K mitotype in Genbank.

Species	Accession	% identity	Publication
Physa sp.	KP182981.1	100	Ng, T. H., Tan, S. K., & Yeo, D. C. (2015). Clarifying the identity of the long-established, globally-invasive <i>Physa acuta</i> Draparnaud, 1805 (Gastropoda: Physidae) in Singapore. <i>BioInvasions Record</i> , 4(3).
	KP182982.1	100	
Physa sp.	ON953193.1	100	Molaba, G. G., Molefe-Nyembe, N. I., Taioe, O. M., Mofokeng, L. S., Thekiso, O. M. M., & Mtshali, K (2022). Molecular Characterization of Larval Trematodes from Naturally Infected Freshwater Snails in South Africa. <i>Available at SSRN 4254241</i> .
	ON953200.1	97.37	
Physinae sp.	OK510774.1	93.25	Young, M. K., Smith, R., Pilgrim, K. L., & Schwartz, M. K. (2021). Molecular species delimitation refines the taxonomy of native and nonnative physinine snails in North America. <i>Scientific reports</i> , 11(1), 1-13.
	OK510777.1	93.09	

Table S5: Conditional probabilities used to assess the maximum likelihood estimation.

Maximum likelihood estimation was used to determine the frequency of restorer alleles, if any, in the Lyon population (p_0) and in the albino population (p). The data used are the male-fertility phenotypes of 69 individuals over 17 generations and 12 K matriline – at each generation several daughters of the same mother are raised and mated in parallel (among which only one propagates the line); this increases the sample size, and provides backup individuals in case one of them fails to produce live offspring. We hypothesized that the restored phenotype is controlled by a single dominant allele (f). Male-fertile individuals can be heterozygous (fs) or homozygous (ff) while sterile ones are all ss . The conditional likelihood of each individual (L) is computed iteratively based on the male-fertility status of (i) the mother, (ii) its own phenotype, (iii) all her daughters. From G1 to G17, the probabilities were calculated iteratively as indicated in the table. Let us take the example of a male-fertile offspring of a fertile mother, that only has fertile daughters (top line of the table). The mother is either ff or fs with probabilities $\frac{p(ff)_{mom}}{p(F)_{mom}}$ and $\frac{p(fs)_{mom}}{p(F)_{mom}}$ given by the previous iteration of the computation (where $p(F)_{mom} = p(ff)_{mom} + p(fs)_{mom}$). The focal individual can be ff or fs given that it has produced no sterile daughter; its probability to be ff , conditional on maternal and daughter phenotypes, and knowing that the paternal allele is f with probability p , is $p(ff) = \frac{p(ff)_{mom}}{p(F)_{mom}} p + \frac{p(fs)_{mom}}{p(F)_{mom}} \frac{1}{2} p$. The probabilities of the fs genotype can be found the same way, and the likelihood of the observed fertile daughter (conditional on the phenotype of her mother and daughters) is $p(ff) + p(fs)$. The case of a fertile individual with fertile mother and at least one sterile daughter is treated separately because in that case the individual must be fs so $p(ff)$ is 0. For a sterile individual with a fertile mother (irrespective of daughters' phenotypes) the mother must be heterozygous and the likelihood of the individual is $\frac{1}{2} (1 - p)$.

At G1, because the male-fertility status of the mother was unknown, probabilities were defined as follows. If the focal individual was fertile :

$$\begin{aligned} p(fs) &= 2p_0(1 - p_0) \\ L &= p(ff) + p(fs) \end{aligned}$$

with $p(ff) = 0$ if the individual had at least one sterile daughter, or $p(ff) = p_0^2$ if all the daughters were fertile.

If the focal individual was sterile:

$$\begin{aligned} p(fs) &= 0 \\ p(ff) &= 0 \\ L &= (1 - p_0)^2 \end{aligned}$$

The (conditional) likelihood of each individual was then $\log()$ transformed. All the conditional log-likelihoods were summed up over individuals and matriline to find the sample likelihood, and different sets of values for p and p_0 were explored.

Mother	Daughters	Focal individual	Probabilities and likelihood values
Fertile	all fertile	Fertile	$p(ff) = \frac{p(ff)_{mom}}{p(F)_{mom}} p + \frac{p(fs)_{mom}}{p(F)_{mom}} \frac{1}{2} p$ $p(fs) = \frac{p(ff)_{mom}}{p(F)_{mom}} (1 - p) + \frac{p(fs)_{mom}}{p(F)_{mom}} \frac{1}{2}$ $L = p(ff) + p(fs)$
	≥ 1 sterile	Fertile	$p(ff) = 0$ $p(fs) = \frac{p(ff)_{mom}}{p(F)_{mom}} (1 - p) + \frac{p(fs)_{mom}}{p(F)_{mom}} \frac{1}{2}$ $L = p(ff) + p(fs)$
		Sterile	$p(ff) = 0$ $p(fs) = 0$ $L = \frac{1}{2} (1 - p)$
Sterile	all fertile	Fertile	$p(ff) = 0$ $p(fs) = p$ $L = p(ff) + p(fs)$
	≥ 1 sterile	Fertile	$p(ff) = 0$ $p(fs) = p$ $L = p(ff) + p(fs)$
		Sterile	$p(ff) = 0$ $p(fs) = 0$ $L = (1 - p)$

Table S6: Divergence observed between K mitotype and N and D mitotype at 11 mitochondrial genes. The percentage of divergence on the COX1 gene was estimated on 13 K mitotype (our sequences plus available sequences from Genbank, Table S2), 1 D mitotype and 5 N mitotype sequences (there are many published N sequences but we kept only those for which other mitochondrial genes are available, see M&M). For the other ten genes, we took only one K mitotype as all the available sequences were only from this study and were all identical.

Gene	Divergence K/N	Divergence K/D	Divergence N/N
ATP6	41.2% ($\pm 0.5\%$ SD)	44.1%	7.3% ($\pm 4.1\%$ SD)
COX1	22.8% ($\pm 0.8\%$ SD)	28.1% ($\pm 0.4\%$ SD)	2.2% ($\pm 1.6\%$ SD)
COX2	40.3% ($\pm 1\%$ SD)	42.9%	13.8% ($\pm 7.2\%$ SD)
COX3	43.3% ($\pm 0.3\%$ SD)	42.4%	4.16% ($\pm 2.4\%$ SD)
CYTB	35.8% ($\pm 0.4\%$ SD)	40.9%	4.27% ($\pm 2.3\%$ SD)
NAD1	45.2% ($\pm 0.8\%$ SD)	50.4%	12.50 ($\pm 6.8\%$ SD)
NAD2	57.4% ($\pm 0.6\%$ SD)	57.4%	8.12% ($\pm 4.9\%$ SD)
NAD3	47.8% ($\pm 0.9\%$ SD)	52.3%	10.42% ($\pm 6.2\%$ SD)
NAD4	48.8% ($\pm 0.7\%$ SD)	52.1%	3.99% ($\pm 1.6\%$ SD)
NAD5	42.9% ($\pm 0.2\%$ SD)	47.5%	5.82% ($\pm 3.5\%$ SD)
NAD6	52.9% ($\pm 1\%$ SD)	61.0%	8.00% ($\pm 4.8\%$ SD)

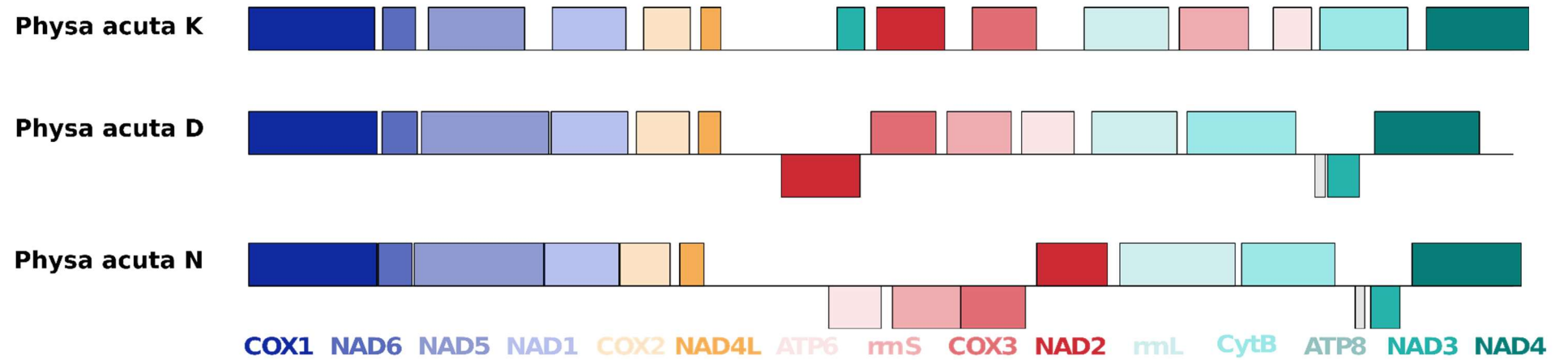


Figure S6: Structural variants in the assembled mitogenomes of D and N individuals of *P. acuta*.