RESEARCH ARTICLE

Negative frequency dependent selection contributes to the maintenance of a global polymorphism in mitochondrial DNA

Zorana Kurbalija Novičić[1,2], Ahmed Sayadi[1], Mihailo Jelić[3] and Göran Arnqvist[1]*

Abstract

Background: Understanding the forces that maintain diversity across a range of scales is at the very heart of biology. Frequency-dependent processes are generally recognized as the most central process for the maintenance of ecological diversity. The same is, however, not generally true for genetic diversity. Negative frequency dependent selection, where rare genotypes have an advantage, is often regarded as a relatively weak force in maintaining genetic variation in life history traits because recombination disassociates alleles across many genes. Yet, many regions of the genome show low rates of recombination and genetic variation in such regions (i.e., supergenes) may in theory be upheld by frequency dependent selection.

Results: We studied what is essentially a ubiquitous life history supergene (i.e., mitochondrial DNA) in the fruit fly *Drosophila subobscura*, showing sympatric polymorphism with two main mtDNA genotypes co-occurring in populations world-wide. Using an experimental evolution approach involving manipulations of genotype starting frequencies, we show that negative frequency dependent selection indeed acts to maintain genetic variation in this region. Moreover, the strength of selection was affected by food resource conditions.

Conclusions: Our work provides novel experimental support for the view that balancing selection through negative frequency dependency acts to maintain genetic variation in life history genes. We suggest that the emergence of negative frequency dependent selection on mtDNA is symptomatic of the fundamental link between ecological processes related to resource use and the maintenance of genetic variation.

Keywords: Balancing selection, Mitochondria, mtDNA, Polymorphism, Negative frequency dependent selection

Background

Species diversity within ecological communities, which should erode over time as a result of stochastic events, competitive exclusion and unstable host-enemy dynamics, is to a large extent maintained by various mechanisms that generate negative frequency-dependent feed-back whereby rarer species are favored [1, 2]. Similarly, genetic polymorphism should be eroded by natural selection and genetic drift [3] and it has long been recognized that negative frequency-dependent selection (NFDS), where rare variants have a selective advantage over more common ones, is in theory the most powerful processes capable of maintaining genetic polymorphism within populations and

species [4, 5]. Yet, while frequency-dependency has a central position in ecology, this is not true to the same extent in population genetics [2].

Lewontin [5] suggested that NFDS should quite generally act to maintain polymorphism whenever there is (i) temporal or spatial environmental variation, (ii) density-dependent resource competition and (iii) G × E interactions, conditions which should be the rule rather than the exception in natural populations [6]. In fact, theory suggests that frequency-dependent selection is an emergent and general property of density-dependency whenever competition is asymmetric [7, 8] such that, in essence, the most severe competitor for a given genotype is itself. Broadly defined, NFDS can result from inherent properties of specific genotypes or be an emergent result of density-dependent processes and environmental heterogeneity in space or time [6, 9]. However, because

* Correspondence: goran.arnqvist@ebc.uu.se
[1]Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Center, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden
Full list of author information is available at the end of the article

recombination disassociates combinations between alleles at multiple loci, such that a given individual may carry a mix of common and rare alleles across many genes, NFDS is often regarded as a weak force in maintaining polymorphism in major fitness components such as polygenic life history traits. Empirical efforts to understand NFDS have therefore been largely devoted to specific traits or genes with major effects [10], such as self-incompatibility in plants [11], immunity genes in vertebrates [12], rare male mating advantage [13], sexually dimorphic signaling [14, 15], mimicry [16] and anti-predatory coloration [17]. However, recent studies have shown that recombination rate varies markedly across the genome and linked sites sometimes tend to co-segregate as blocks, i.e. supergenes, for example as a result of chromosomal inversions [18]. Detecting NFDS is very challenging [10, 19] but several studies have implied a more general role for NFDS in cases where recombination is reduced, such as in bacteria [20] and in species where key phenotypes are affected by inversions [21–23]. The fact that supergenes can have major effects on variation in life history traits [21, 24, 25];) supports the possibility NFDS may contribute to the maintenance of polymorphism also in life history genes [4, 26, 27].

Here, we ask whether the classic, well characterized and striking pattern of polymorphism in mtDNA within natural populations of the fruit fly *Drosophila subobscura* (see Fig. 1) is the result of NFDS. Mitochondrial DNA (mtDNA) can be regarded as a life history supergene [24], which is maternally inherited, haploid and carries 13 protein coding genes which co-segregate with several sites that affect mitochondrial transcription and translation [28]. The mtDNA genes encode for parts of the ATP-producing OXPHOS pathway of the inner mitochondrial membrane and are as such key metabolic genes and important candidate genes for rate-dependent life history traits in eukaryotes [24, 29]. Segregating mitochondrial DNA haplotypes were long assumed to be functionally equivalent and thus neutral, but this view has changed markedly during the last decade (e.g., [24, 30, 31]). In *D. subobscura*, for example, a number of previous studies have shown that individuals carrying the two major mtDNA haplotypes HI and HII show differences in a range of key life history traits, such as metabolic rate [32], sex-specific juvenile survival [33], male fertility [34], development time [35], adult longevity [33, 35], desiccation resistance [35] and in estimates of fitness [35–37]. A few non-experimental studies of other insects [38–40], as well as an experimental study of non-sympatric haplotypes [41], have indicated that segregating mtDNA variation may be maintained by NFDS but experimental evidence for this is currently lacking.

We initiated a set of experimental evolution lines ($N = 12$), using a 2×2 crossed design, aimed at experimentally testing whether NFDS acts on sympatric mtDNA haplotypes of *D. subobscura* and, in addition, to assess whether resource conditions affect NFDS.
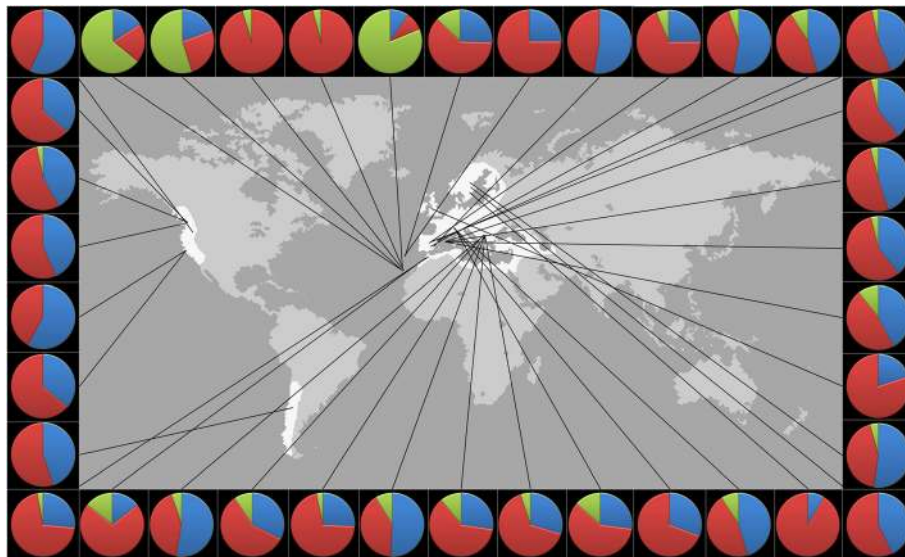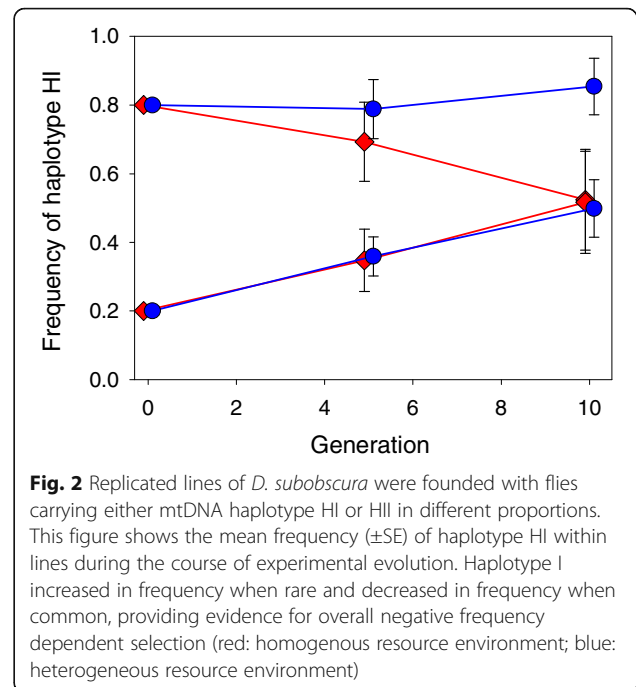

Fig. 1 Mitochondrial DNA variation within natural populations of *D. subobscura* is heavily dominated by the two haplotype groups HI (blue) and HII (red) in all three disjunct parts of its global distribution (white). Apart from populations on the Canary Islands, where other haplotypes (green) are common, the mtDNA haplotype frequencies are similar across populations, with on average 37% of all individuals carrying HI and 58% HII haplotypes. This observation alone suggests that balancing selection is contributing to the maintenance of genetic variation in mtDNA. Several experimental studies have shown that flies carrying HI and HII differ in key life history traits, such as as metabolic rate, development time, adult longevity and desiccation resistance. Distribution map reproduced with permission from Rodríguez-Trelles et al. [42]

Previous non-experimental work in this species has shown that mtDNA haplotype frequencies are polymorphic and stable over time in cage cultures, in a manner consistent with NFDS [38]. Our laboratory populations were (i) founded by contrasting starting frequencies of the two dominating mtDNA haplotype groups HI and HII (20/80% vs. 80/20%) (Fig. 1) and were (ii) allowed to evolve under competitive conditions in a more or less homogenous environment with regards to food resource quality. Flies used were from mitonuclear introgression lines constructed from a single natural population and all flies used share the same outbred nuclear genetic background, and were allowed to evolve in the laboratory for 10 non-overlapping generations. We sequenced and assembled the mitogenomes used (three HI and three HII; Additional file 1) and then employed PoolSeq to estimate genotype frequencies in samples of flies from all laboratory populations at generation 5 and 10. We make two explicit predictions. Most importantly, we predict that haplotypes should tend to increase in frequency when rare and decrease when common if NFDS contributes to the maintenance of polymorphism for the HI and HII haplotype groups. Moreover, if NFDS results from resource competition and $G \times E$ interactions [5], then the strength of NFDS should be affected by our food resource treatment.

Results

The key test for NFDS is one showing that the relative fitness of a genotype is inversely proportional to its frequency. As suggested by the general pattern of genotype frequency shifts during the experiment (Fig. 2), there was indeed an overall significant effect of starting frequency on mtDNA haplotype frequency changes (Table 1). The mean frequency of haplotype HI increased significantly in frequency when rare and decreased when common, and this was consistent over the course of our experiment (Fig. 3a). While the effects of starting frequency on frequency changes seemed weaker in the more heterogeneous resource environment, an overall effect of resource conditions on frequency change was not supported by these analyses (Table 1; non-significant SF × E interaction).

Analyses of selection coefficients derived from our data provided further support for our main prediction of overall NFDS across our 12 laboratory populations, as the estimate of mean S_I was larger than zero (global t-test of H_0: $S_I = 0$, $t_{11} = 2.47$, $P = 0.031$; bootstrap mean $S_I = 0.16$, 95% CI: 0.03–0.27). In contrast to the analyses of haplotype frequency changes, however, the analyses of selection coefficients also provided support for the effect of starting frequency on selection being contingent upon the environmental conditions (Table 1; significant SF × E

Fig. 2 Replicated lines of *D. subobscura* were founded with flies carrying either mtDNA haplotype HI or HII in different proportions. This figure shows the mean frequency (±SE) of haplotype HI within lines during the course of experimental evolution. Haplotype I increased in frequency when rare and decreased in frequency when common, providing evidence for overall negative frequency dependent selection (red: homogenous resource environment; blue: heterogeneous resource environment)

interaction). The latter effect was largely due to an apparent absence of selection in one of our four treatment groups (Fig. 3b), where lines showed little change in mean haplotype frequency over time (Fig. 3a), rendering NFDS to be stronger overall in the homogenous compared to the heterogeneous resource environments. A closer inspection of the data supported this interpretation. Mean S_I was non-zero under homogenous food conditions when HI was started as both common (bootstrap mean $S_I = 0.22$, 95% CI: 0.09–0.34) and rare (bootstrap mean $S_I = 0.24$, 95% CI: 0.09–0.39). In contrast, mean S_I was non-zero under heterogeneous food conditions when HI was started as rare (bootstrap mean $S_I = 0.25$, 95% CI: 0.16–0.33) but not when started as common (bootstrap mean $S_I = -0.07$, 95% CI: $-0.19 - 0.05$).

Average population size during the experimental evolution was $N = 773$ (SE = 56) individuals. Population size increased markedly during the first five generations of our experiment and grew more rapidly to a larger size in populations experiencing the more heterogeneous environment (Additional file 1).

Discussion

Our main results provide direct experimental support for the hypothesis that a classic and enigmatic genetic polymorphism (Fig. 1) in what is essentially a general life history supergene [24] is at least partly attributable to NFDS, as the relative fitness of haplotypes was inversely proportional to their starting frequency in our experimental evolution lines. While most classic studies of NFDS in other regions of the genome involve quite

Table 1 Repeated measures ANOVAs of the effects of starting frequency (rare or common) and environmental conditions (homogenous or heterogeneous) on changes in the frequency of mtDNA haplotype HI and the strength of frequency dependent selection in the experimental evolution lines. Generation represents the within-line effect of generation 0–5 versus 5–10
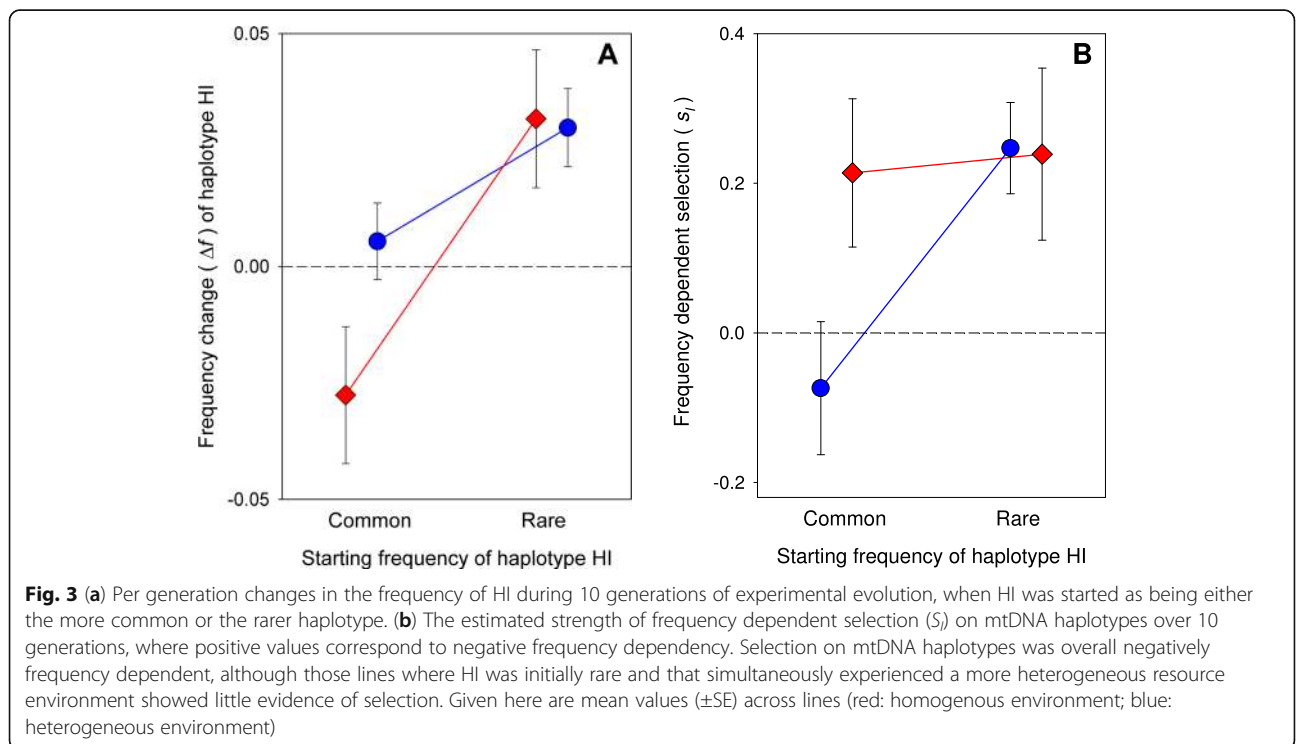
	Frequency change of HI (Δf)				Frequency dependent selection coefficient (S_l)			
Between-line effects:	df	MS	F	P	P_{perm} [a]	df	MS	F	P	P_{perm} [a]
Starting frequency (SF)	**1**	**1.1E-02**	**8.091**	**0.022**	**0.021**	1	2.1E-02	0.048	0.832	0.849
Environment (E)	1	1.5E-03	1.123	0.320	0.336	1	4.5E-02	0.104	0.756	0.786
SF × E	1	1.8E-03	1.402	0.270	0.276	1	**2.3E+00**	**5.394**	**0.049**	**0.034**
Residual	8	1.3E-03				8	4.3E-01			
Within-line effects:										
Generation (G)	1	4.5E-06	0.001	0.972		1	2.1E-01	0.308	0.594	
G × SF	1	2.4E-06	0.001	0.979		1	4.0E-01	0.585	0.466	
G × E	1	1.4E-04	0.043	0.842		1	1.5E-01	0.214	0.656	
G × SF × E	1	4.9E-04	0.146	0.713		1	1.6E+00	2.348	0.164	
Residual	8	3.4E-03				8	6.8E-01			

[a] Permutation tests of between-line effects
Both entries in bold face are significant at alpha = 0.05

specific biological mechanisms [10], our findings are consistent with a more general role for NFDS. Life history strategies represent multidimensional phenotypes, characterized by differences in growth, reproduction and ageing, and it is likely that the relative fitness of a given life history strategy frequently and importantly depends upon the strategies adopted by other individuals in the population [43, 44]. In cases where variation in life history strategies are affected by a relatively restricted set of

supergenes, or major-effect genes, we expect NFDS to contribute to standing genetic variation.

Because mtDNA is haploid and maternally transmitted, its role as a supergene is in some ways distinct. Yet, the fact that it harbours a non-recombining block of key life history genes suggests that it can serve as a model for the fate of life history supergenes [24]. Interestingly, mtDNA often shows within-population polymorphism [30] and recent studies have revealed that much of this

Fig. 3 (**a**) Per generation changes in the frequency of HI during 10 generations of experimental evolution, when HI was started as being either the more common or the rarer haplotype. (**b**) The estimated strength of frequency dependent selection (S_l) on mtDNA haplotypes over 10 generations, where positive values correspond to negative frequency dependency. Selection on mtDNA haplotypes was overall negatively frequency dependent, although those lines where HI was initially rare and that simultaneously experienced a more heterogeneous resource environment showed little evidence of selection. Given here are mean values (±SE) across lines (red: homogenous environment; blue: heterogeneous environment)

Kurbalija Novičić *et al. BMC Evolutionary Biology* (2020) 20:20

Page 5 of 9

variation is in fact functional [31]. This implies that NFDS on mtDNA may be common, a possibility that is supported by both laboratory observations of allopatric haplotypes [41] and instances of geographically widespread sympatric coexistence between distinct mtDNA haplotype groups in other taxa (e.g., [45]). Needless to say, this has important ramifications for studies that implicitly or explicitly assume that mtDNA segregates as a neutral genetic marker [46].

In addition to the main effect of starting frequency, we also found support for an effect of environmental conditions on the strength of NFDS in our experiment. While this provides general support for a role of competition as a generator of NFDS, our expectation was to see stronger NFDS under more heterogeneous resource conditions in line with Lewontin's [5] tenet. However, we found the opposite. In theory, the strength of NFDS should be collectively determined by the intensity of competition, the magnitude of $G \times E$ interactions for competitive ability and the degree of environmental heterogeneity. While our aim was to standardize net quantitative level of resources across food treatments, and to vary only the degree of heterogeneity, the observed population dynamics in our experimental evolution lines (Additional file 1) showed that this was not fully successful. Population size grew more rapidly and to a larger size in populations experiencing the more heterogeneous environment, implying that the absolute level of density dependent regulation was lower under these conditions at least during the first half of the experiement (Additional file 1). Previous experimental studies of other genes under NFDS in *Drosophila* have demonstrated that NFDS is weakened under less competitive conditions [26] and we suggest that the weaker NFDS seen in our heterogeneous environment may represent the result of less competitive conditions overall. We note, however, that alternative explanations include (i) the possibility that the equilibrium frequencies of haplotypes varied across food regimes, such that the equilibrium frequency of HI was considerably higher under the more heterogeneous environment, (ii) an impact of positive selection for HI in the the more heterogeneous environment and (iii) environment-specific complex mitonuclear epistatic effects (i.e., a $G \times G \times E$ interaction [47]). In either case, this particular facet of our results does provide support for a functional link between the resource environment and the selective regime.

The main candidate gene responsible for the many documented functional differences between the sympatric HI and HII haplotype groups in *D. subobscura* (e, g., [32–36]) seems, remarkably, to not be a protein coding gene but rather a gene that affects translational events. The HI and HII haplotype groups show only two consistent differences (Additional file 1). One is a synonymous substitution within Nad5, which has

served as the diagnostic basis for previous genotyping using restriction enzymes. The other is a SNP in 12S rRNA, located in a region homologous to the base of stem 15 in the mammalian 12S rRNA. Interestingly, the region in question is polymorphic in humans and there is evidence of both disease variants [48] and adaptive polymorphisms in 12S rRNA [49]. The specific HI/HII SNP in 12S rRNA of *D. subobscura* is associated with differences in the predicted secondary structure of the rRNA (Additional file 1), which may affect ribosomal protein synthesis and therefore also rate-dependent life history traits that are dependent upon the abundance of metabolic enzymes.

Conclusions

Our work supports the view that we should consider balancing selection through NFDS a prime suspect among those processes that maintain genetic variation in life history genes in nature [3]. Moreover, the fact that ecological resources affected the NFDS regime in *D. subobscura* firmly places ecology at the root of maintenance of genetic variation of life history traits in this particular system. The emergence of NFDS on mtDNA may thus more generally be symptomatic of the fundamental and inevitable link between ecology and evolution [6–8, 43].

Methods

Construction of Mito-nuclear introgression lines

To isolate the genetic effects of mtDNA from the nuclear genome, our experiments were based on mito-nuclear introgression lines (MNILs). The construction of our MNILs is explained in detail in Kurbalija Novicic et al. [32]. Briefly, all MNILs used were created from isofemale lines (IFLs), each of which was founded by a single wild-collected mated female originating from a common single natural population (Sicevo Gorge-Serbia, S 43°19′55.58″ N 22°0837.98″). Lines were first genotyped for their mtDNA haplotype using restriction enzymes [50] and we selected six IFLs, three of which carried HI haplotypes and three HII haplotypes, each of which was then backcrossed with a common seventh IFL ("D"). In each MNIL, backcrossing was performed by pairing 10 virgin females from a given IFL with 20 males from the D line. We used this repeated introgressive backcrossing procedure for 12 subsequent generations to replace the nuclear genome of a given IFL with the common outbred D nuclear genome (> 99.95% replaced). Note that IFLs were not inbred or otherwise made isogenic. To exclude the possibility of contamination during introgression, the mtDNA integrity of all MNILs were validated at generation 5, 8 and 12 by genotyping a sample of flies from each MNIL. We also screened for the presence of *Wolbachia* in all MNILs, by a PCR assay using 16S rDNA *Wolbachia*-specific

primers [51] using methods detailed in García-Martínez et al. [37]. We used two different *Drosophila* strains containing *Wolbachia* as positive controls (*D. melanogaster* stock no. 5, Bloomington Stock Centre, *D. simulans*, Riverside strain). These PCR assays were negative for all of our MNILs as well as for the D line. All lines were maintained and all experiments performed under constant laboratory conditions, at 19 °C, 60% relative humidity, light of 300 lx, and at a photoperiod of 12 h light: 12 h dark.

Founding of the experimental evolution populations

Prior to the experiments, we amalgamated the three MNILs carrying HI into one HI source population and the three HII MNILs into a HII source population to homogenize potential nuclear genetic variation across MNILs. This was done by mixing 100 adult flies from each MNIL, in two population cages (i.e., $N = 3 \times 100$ flies per cage) (Plexiglas cages, L25 cm x W15 cm xH15 cm, with 3 dishes each containing 30 ml of cornmeal) and maintaining these for one subsequent full generation under standard laboratory conditions. The two source populations, thus, carried either the HI or the HII mtDNA haplotype, expressed in an outbred and common nuclear genetic background (i.e., D). Virgin flies from these two source populations were then sexed and used to found the experimental evolution populations.

We initiated $N = 12$ experimental evolution populations. In each population, $N = 100$ virgin flies (sex ratio 1:1) aged 3–5 days were introduced from the source populations into a population cage (L25 cm x W15 cm xH15 cm). We varied the starting frequency of HI and HII haplotypes and food resource conditions across populations, using a 2×2 crossed design ($N = 3$ populations per cell), in the following manner. In half of the cages, 80% of the founding flies were from the HI source population and 20% from the HII source population. In the other half, 20% of the founding flies were from the HI source population and 80% from the HII source population. In terms of food resource conditions, the amount of medium (both by volume and surface area) was identical in the two treatment groups while within population variation in nutrient concentration was manipulated as follows. Half of the cages (homogenous food conditions) contained 3 identical food dishes, each with 30 ml of standard cornmeal medium (YC) containing 1.5% of yeast. The other half of the cages (heterogeneous food conditions) also contained 3 dishes each with 30 ml standard medium, but these differed in yeast concentration (YL-0.375%, YC- 1.5%, YH- 6%).

Maintenance of the experimental evolution populations

Populations were maintained in the laboratory in a manner that ensured discrete generations, 40 days/cycle [40], at 19 °C, 60% relative humidity and a 12 h light: 12 h dark cycle. At day 40, the three old food dishes were replaced by three new ones and flies were allowed in total 9 days for oviposition [40]. The new dishes, containing eggs and larvae, were then cleared from any adults and transferred to a new cage to start the next generation. All adult flies in each old cage were then counted, once they were dead.

Sequencing and mitogenome assemblies

Prior the haplotype frequency estimation (see below), we sequenced and assembled the all mtDNA haplotypes used. DNA was extracted from the six IF lines (HI: 1, 3, 5; HII: 21, 25, 29) used to create the MNILs, using a salt-ethanol precipitation protocol. Flies were first gently macerated and placed in preparation buffer (100 mM NaCl, 10 mM Tris-HCl, pH = 8.0, 0.5% SDS) together with proteinase K, vortexed and incubated at 50 °C overnight. Samples were then frozen overnight. To precipitate DNA, we added saturated NaCl several times before adding 95% ethanol, and we spun the DNA into a pellet. The DNA pellet was suspended in TE 4 buffer (pH = 7.6). DNA quality and quantity was assessed using NanoDrop, Qubit and Bioanalyzer, followed by fragment length assessment on an agarose gel.

Sequencing libraries were prepared from 100 ng DNA using the TruSeq PCRfree DNA library preparation kit. The six samples were then sequenced to 125 bp paired-end reads in two lanes on an Illumina HiSeq2500 system using v4 sequencing chemistry. In total, we sequenced on average 194 million reads for each library. Mitogenomes from the six samples were then assembled using a 5% subset of the total number of reads from each library. Reads were fed to the MITObim V 1.8 algorithm [52] and the MIRA V 4.0.2 [53] assembler, to perform guided assemblies, using the *Drosophila pseudoobscura* mitogenome (GenBank: FJ899745.1) as a reference genome. All obtained assemblies were circular mitogenomes with a size of almost 16kbp. The final assemblies were then aligned using ClustalW [54] and MAFFT [55], and manually curated to obtain a final polished assembly for each haplotype. The assembled mitogenomes were annotated using DOGMA [56] and MITOS [57], using default parameter settings, and finally manually curated.

To assess the validity of our mitogenome assembly, all *Drosophila subobscura* mtDNA sequences available on GenBank (Cox1, Cox2, Cox3, Cob, Nad1, Nad2, Nad3, Nad5, rrnL, rrnS, the A + T rich region, and several trna), in total covering more than 50% of the total assembly, were aligned against our mitogenomes. Without exception, these showed > 99% sequence identity.

Several unique SNPs were found in each of the six mitogenome haplotype (see below), of which two consistently distinguished the HI and HII haplotype groups. Coverage depth for each SNP was verified by mapping

Kurbalija Novičić *et al. BMC Evolutionary Biology*   (2020) 20:20

Page 7 of 9

back the reads used for the mitogenome assembly using Bowtie v 1.2 [58]. These efforts confirmed all SNPs identified in the assembly step. Here, we focus on the two main haplotype groups I and II, which show a striking and consistent pattern of within-population polymorphism across populations (Fig. 1) and functional phenotypic differences (see Introduction). Although SNPs do occur within each of the two haplotype groups, such SNPs are rare (e.g. [59]) and are not consistently polymorphic.

Estimation of the mtDNA haplotype frequencies

We used pool-seq to estimate haplotype frequency evolution, by sequencing samples of flies from the 5th and 10th generation of each experimental evolution population. In each sample, 105 randomly selected flies per cage were pooled and subjected to DNA extraction (in groups of 15) and sequencing library preparations as described above. The $N = 24$ samples were then sequenced to 125 bp paired-end reads in two lanes on an Illumina HiSeq2500 system, using v4 sequencing chemistry. Our pool-seq effort was designed to provide sufficient sequencing depth for accurate estimation of mtDNA haplotype frequencies, but does not allow detailed analyses of the nuclear genome.

We sequenced, on average, 66 million reads for each library. Reads from each library were then mapped back to the six assembled mitogenomes, retaining only unique and zero-mismatch mappings using Bowtie v 1.2 [58]. The number of reads mapping to the two SNPs that distinguished the HI and HII types (in Nad5 and in 12S rRNA) were then counted and used as our estimate of the relative frequency of each main haplotype (HI or HII) in each sample.

Statistical analyses

For each sample, all reads mapping to the two diagnostic SNPs (see below) were counted as being of either type HI or type HII. The proportion of HI reads for the two different SNPs were very closely correlated indeed across the 24 samples ($r = 0.987$), such that both markers provided virtually identical estimates. Here, we used the mean proportion across both SNPs here to estimate the frequency of HI present in each pool-seq sample.

Evolution can be defined as changes in genotype frequencies within a population, and our design allowed us to derive two temporally separated repeated measures of per generation haplotype frequency changes in our each evolving line as either $\Delta f_{0-5} = (f_5 - f_0)/5$ or $\Delta f_{5-10} = (f_{10} - f_5)/5$ where subscripts denote the generation at which the sample was collected. In addition, we estimated $\Delta f_{0-10} = (f_{10} - f_0)/10$ to assess net haplotype frequency change. Because only two genotypes were involved, the frequency of HI: $f_I = 1 - f_{II}$, and we thus restrict our analyses to changes in the frequency of one of the haplotypes. For each estimate of Δf, we also derived a frequency dependent selection coefficient (S_I) corresponding to the strength of selection required to cause the observed change in haplotype frequencies (see below).

Each evolving line represents an observational unit in our design, and we hence analysed our data using repeated measures ANOVAs (i.e., within-subjects ANOVAs). Here, each line represents the subject, and starting frequency of HI (0.2 or 0.8) and environmental condition (homogenous or heterogeneous) were two between-subjects factors, and the two repeated measures (of Δf or S_I) taken at different generational intervals was treated as a within-subjects factor. In these analyses, the focal between-subjects factors test for effects of our experimental treatments and the within-subjects factor tests whether the pattern of evolution changed during the course of our experiment. This analytical strategy draws on the fact that we track frequency dynamics in replicated and independent lines and the non-focal effect of random genetic drift, which is predicted to be substantial for mtDNA dynamics, forms a part of the residual term of our inferential models. Conventional F-tests of the between-subjects factors were validated using permutation tests, based on 9999 random permutations of data. Mean S_I for different groups were assessed using 95% CI's from 9999 bootstrap replicates of data.

Estimates of the strength of frequency dependent selection

We also wished to more explicitly assess whether the strength of frequency dependent selection on HI and HII differed across the two environmental experimental treatments (homogenous / heterogeneous). To allow this, we derived simple measures of frequency dependent selection from our empirical data using the following rationale. Previous explicit modelling of non-experimental data in this system has shown that the best fit with haplotype frequency dynamic data occurs where there is no positive or negative selection on these mtDNA haplotypes but where there is negative frequency dependent selection which is equally strong on both haplotypes [38]. We consider the two mtDNA haplotypes H_I and H_{II}, with the frequencies p_I and p_{II} ($p_I + p_{II} = 1$) and fitnesses W_I and W_{II}. The per-generation change in p_I is then given by

$$\Delta p_I = \frac{p_I p_{II}(W_I - W_{II})}{\overline{W}}$$

Observations from both natural populations (see Fig. 1; Additional file 1) and replicated laboratory cage populations [38, 40] also suggest that selection is symmetrical with an equilibrium for the two haplotypes H_I and H_{II} in

the vicinity of $p_I = p_{II} = 0.5$. Assuming (i) an equilibrium frequency of $p_I = p_{II} = 0.5$, (ii) that haplotype fitness is linearly related to haplotype frequency and (iii) that selection is symmetrical, all of which are supported by previous studies [38, 40], we attain

$$W_I = 1 - p_I s_I$$

$$W_{II} = 1 - p_{II} s_I$$

where s_I is the frequency dependent selection coefficient. Given that $\overline{W} = p_I W_I + p_{II} W_{II}$ we can then estimate s_I from our empirical observations of Δp_I as

$$s_I = \frac{-\Delta p_I p_I - \Delta p_I p_{II}}{-\Delta p_I p_{II}{}^2 - p_I p_{II}{}^2 + p_I{}^2 p_{II} - \Delta p_I p_I{}^2}$$

Defined this way, s_I assumes an arbitrary scale but is positive when of Δp_I changes towards the attractor and negative when it changes away from the attractor. For each cage, we estimated two independent repeated measures of s_I based on the observed haplotype frequency changes between generations 0–5 and 5–10, respectively. We note that our measure will be precise when the true equilibrium is near $p_I = p_{II} = 0.5$ but more approximate if the true equilibrium deviates from this condition.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12862-020-1581-2.

Additional file 1. Supplementary results.

Abbreviations
ATP: Adenosine triphosphate; G × E: Genotype by environment; IFL: Isofemale line; MNIL: Mitonuclear introgression line; mtDNA: Mitochondrial DNA; NFDS: Negative frequency dependent selection; OXPHOS: Oxidative phosphorylation; SNP: Single-nucleotide polymorphism

Availability of data and materials
All mitogenome assemblies have been deposited at Genbank (see Additional file 1). Raw sequencing data has been deposited at the NCBI Sequence Read Archive database (SRA) under the BioProject accession number PRJNA589517. Data from our experiments has been published in Mendeley Data and are available at DOI: https://doi.org/10.17632/59gxcyvzkc.1.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Center, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden. [2]Department of Neuroscience, Psychiatry, Uppsala University Hospital, Entrance 10, 751 85 Uppsala, Sweden. [3]Faculty of Biology, University of Belgrade, Studentski trg 16, Belgrade 11000, Serbia.

References
1. Chesson P. Mechanisms of maintenance of species diversity. Annu Rev Ecol Syst. 2000;31:343–66.
2. Doebeli M, Ispolatov I. Complexity and diversity. Sci. 2010;328:494–7.
3. Charlesworth, B., and K. A. Hughes. 2000. The maintenance of genetic variation in life-history traits. Pp. 369–392 in R. S. Singh and C. B. Krimbas eds. Evolutionary genetics. Cambridge Univ. Press, Cambridge, UK.
4. Ayala FJ, Campbell CA. Frequency-dependent selection. Annu Rev Ecol Syst. 1974;5:115–38.
5. Lewontin RC. The genetic basis of evolutionary change. Press, NY: Columbia Univ; 1974.
6. Heino M, Metz JA, Kaitala V. The enigma of frequency-dependent selection. Trends Ecol Evol. 1998;13:367–70.
7. Kisdi E. Evolutionary branching under asymmetric competition. J Theor Biol. 1999;197:149–62.
8. Svardal H, Rueffler C, Hermisson J. A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. Theor Pop Biol. 2015;99:76–97.
9. Gromko MH. What is frequency-dependent selection? Evol. 1977;31:438–42.
10. Llaurens V, Whibley A, Joron M. Genetic architecture and balancing selection: the life and death of differentiated variants. Mol Ecol. 2017;26:2430–48.
11. Castric V, Vekemans X. Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. Mol Ecol. 2004;13:2873–89.
12. Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proc Roy Soc Lond B. 2010;277:979–88.
13. Ehrman L, Spassky B, Pavlovsky O, Dobzhansky T. Sexual selection, geotaxis, and chromosomal polymorphism in experimental populations of *Drosophila pseudoobscura*. Evol. 1965;19:337–46.
14. Bleay C, Comendant T, Sinervo B. An experimental test of frequency-dependent selection on male mating strategy in the field. Proc. Roy. Soc. Lond. B. 2007;274:2019–25.
15. Olendorf R, Rodd FH, Punzalan D, Houde AE, Hurt C, Reznick DN, Hughes KA. Frequency-dependent survival in natural guppy populations. Nat. 2006; 441:633–6.
16. Thompson MJ, Jiggins CD. Supergenes and their role in evolution. Heredity. 2014;113:1–8.
17. Bond AB, Kamil AC. Spatial heterogeneity, predator cognition, and the evolution of color polymorphism in virtual prey. Proc Natl Acad Sci U S A. 2006;103:3214–9.
18. Ritz KR, Noor MA, Singh ND. Variation in recombination rate: adaptive or not? Trends Genet. 2017;33:364–74.

Kurbalija Novičić *et al. BMC Evolutionary Biology*   (2020) 20:20

Page 9 of 9

19. Fijarczyk A, Babik W. Detecting balancing selection in genomes: limits and prospects. Mol Ecol. 2015;24:3529–45.
20. Rainey PB, Buckling A, Kassen R, Travisano M. The emergence and maintenance of diversity: insights from experimental bacterial populations. Trends Ecol Evol. 2000;15:243–7.
21. Schwander T, Libbrecht R, Keller L. Supergenes and complex phenotypes. Curr. Biol. 24:R288–94.
22. Chouteau M, Llaurens V, Piron-Prunier F, Joron M. Polymorphism at a mimicry supergene maintained by opposing frequency-dependent selection pressures. Proc Natl Acad Sci U S A. 2017;114:8325–9.
23. Lamichhaney S, et al. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). Nat Genet. 2016;48:84–8.
24. Ballard JWO, Melvin RG. Linking the mitochondrial genotype to the organismal phenotype. Mol Ecol. 2010;19:1523–39.
25. Lowry DB, Willis JH. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol. 2010;8:e1000500.
26. Fitzpatrick MJ, Feder E, Rowe L, Sokolowski MB. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. Nat. 2007;447:210–2.
27. Snyder TP, Ayala FJ. Frequency-dependent selection at the Pgm-1 locus of *Drosophila pseudoobscura*. Genetics. 1979;92:995–1003.
28. Wallace DC. Mitochondrial DNA in evolution and disease. Nature. 2016;535: 498–500.
29. Stojković B, Sayadi A, Đorđević M, Jović J, Savković U, Arnqvist G. Divergent evolution of lifespan associated with mitochondrial DNA evolution. Evol. 2016;71:160–6.
30. Bazin E, Glemin S, Galtier N. Population size does not influence mitochondrial genetic diversity in animals. Sci. 2006;312:570–2.
31. Dobler R, Rogell B, Budar F, Dowling DK. A meta-analysis of the strength and nature of cytoplasmic genetic effects. J Evol Biol. 2014;27:2021–34.
32. Kurbalija Novicic Z, Immonen E, Jelic M, Andjelkovic M, Stamenkovic-Radak M, Arnqvist G. Within-population genetic effects of mtDNA on metabolic rate in *Drosophila subobscura*. J Evol Biol. 2015;28:338–46.
33. Jelić M, et al. Sex-specific effects of sympatric mitonuclear variation on fitness in *Drosophila subobscura*. BMC Evol Biol. 2015;15:135.
34. Castro JA, Oliver P, Christie JS, Picornell A, Ramon M, & Moya A. Assortative mating and fertility in two Drosophila subobscura strains with different mitochondrial DNA haplotypes. Genetica, 2003;119(3):295–301.
35. Christie JS, et al. Fitness and life-history traits of the two major mitochondrial DNA haplotypes of *Drosophila subobscura*. Heredity. 2004;93:371–8.
36. Fos M, Dominguez MA, Latorre A, Moya A. Mitochondrial DNA evolution in experimental populations of *Drosophila subobscura*. Proc Natl Acad Sci U S A. 1990;87:4198–201.
37. García-Martínez J, Castro JA, Ramón M, Latorre A, Moya A. Mitochondrial DNA haplotype frequencies in natural and experimental populations of *Drosophila subobscura*. Genetics. 1998;149:1377–82.
38. Arnqvist G, Kurbalija Novičić Z, Castro JA, Sayadi A. Negative frequency dependent selection on sympatric mtDNA haplotypes in *Drosophila subobscura*? Hereditas. 2016;153:15.
39. MacRae AF, Anderson WW. Evidence for non-neutrality of mitochondrial DNA haplotypes in *Drosophila pseudoobscura*. Genetics. 1988;120:485–94.
40. Olivier P, Balanyà J, Ramon MM, Picornell A, Serra L, Moya A, Castro JA. Population dynamics of the 2 major mitochondrial DNA haplotypes in experimental populations of *Drosophila subobscura*. Genome. 2005;48:1010–8.
41. Kazancıoğlu E, Arnqvist G. The maintenance of mitochondrial genetic variation by negative frequency-dependent selection. Ecol Lett. 2014;17:22–7.
42. Rodríguez-Trelles F, Rodríguez M, Scheiner S. Tracking the genetic effects of global warming: *Drosophila* and other model systems. Conserv Ecol. 1998;2:2.
43. Brommer JE. The evolution of fitness in life-history theory. Biol Rev. 2000;75:377–404.
44. Wolf M, Van Doorn GS, Leimar O, Weissing FJ. Life-history trade-offs favour the evolution of animal personalities. Nat. 2007;447:581–4.
45. Kvie KS, Hogner S, Aarvik L, Lifjeld JT, Johnsen A. Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*). Ecol. Evol. 3:126–44.
46. Ballard, J.W.O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Mol. Ecol.13:729–744.
47. Arnqvist G, Dowling DK, Eady P, Gay L, Tregenza T, Tuda M, Hosken DJ. Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. Evol. 2010;64:3354–63.
48. Jacobs HT. Disorders of mitochondrial protein synthesis. Human Mol Gen. 2003;12:R293–301.
49. Ruiz-Pesini E, Wallace DC. Evidence for adaptive selection acting on the tRNA and rRNA genes of human mitochondrial DNA. Hum Mutat. 2006;27:1072–81.
50. Jelic M, et al. Absence of linkage disequilibria between chromosomal arrangements and mtDNA haplotypes in natural populations of *Drosophila subobscura* from the Balkan Peninsula. Genome. 2012;55:214–21.
51. O'Neill SL, Giordano R, Colbert AM, Karr TL, Robertson HM. 16SrRNA phylogenetic analysis of the bacterial endosymbiont associated with cytoplasmatic incompatibility in insects. Proc Natl Acad Sci U S A. 1992;89: 2699–702.
52. Hahn C, Bachmann L, Chevreux B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads: a baiting and iterative mapping approach. Nucleic Acids Res. 2013;41:e129.
53. Chevreux B, Wetter T, Suhai. Genome sequence assembly using trace signals and additional sequence information. German Conf Bioinf. 1999;99:45–56.
54. Larkin MA, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23: 2947–8.
55. Katoh K, Standley D. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
56. Wyman SK, Jansen RK, Boore JL. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 2004;20:3252–5.
57. Bernt M, et al. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9.
58. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10:R25.
59. Erić P, Jelić M, Savić-Veselinović M, Kenig B, Anđelković M, Stamenković-Radak M. Nucleotide diversity of Cyt b gene in *Drosophila subobscura* Collin. Genetika. 2019;51(1):213–26.

Publisher's Note