Rapidly Evolving Mitochondrial Genome and Directional Selection in Mitochondrial Genes in the Parasitic Wasp *Nasonia* (Hymenoptera: Pteromalidae)

Deodoro C. S. G. Oliveira,[1] Rhitoban Raychoudhury,* Dennis V. Lavrov,† and John H. Werren**

*Department of Biology, University of Rochester; and †Department of Ecology, Evolution, and Organismal Biology, Iowa State University

We sequenced the nearly complete mtDNA of 3 species of parasitic wasps, *Nasonia vitripennis* (2 strains), *Nasonia giraulti*, and *Nasonia longicornis*, including all 13 protein-coding genes and the 2 rRNAs, and found unusual patterns of mitochondrial evolution. The *Nasonia* mtDNA has a unique gene order compared with other insect mtDNAs due to multiple rearrangements. The mtDNAs of these wasps also show nucleotide substitution rates over 30 times faster than nuclear protein-coding genes, indicating among the highest substitution rates found in animal mitochondria (normally $<$10 times faster). A McDonald and Kreitman test shows that the between-species frequency of fixed replacement sites relative to silent sites is significantly higher compared with within-species polymorphisms in 2 mitochondrial genes of *Nasonia*, *atp6* and *atp8*, indicating directional selection. Consistent with this interpretation, the Ka/Ks (nonsynonymous/synonymous substitution rates) ratios are higher between species than within species. In contrast, *cox1* shows a signature of purifying selection for amino acid sequence conservation, although rates of amino acid substitutions are still higher than for comparable insects. The mitochondrial-encoded polypeptides *atp6* and *atp8* both occur in F0F1ATP synthase of the electron transport chain. Because malfunction in this fundamental protein severely affects fitness, we suggest that the accelerated accumulation of replacements is due to beneficial mutations necessary to compensate mild-deleterious mutations fixed by random genetic drift or *Wolbachia* sweeps in the fast evolving mitochondria of *Nasonia*. We further propose that relatively high rates of amino acid substitution in some mitochondrial genes can be driven by a "Compensation-Draft Feedback"; increased fixation of mildly deleterious mutations results in selection for compensatory mutations, which lead to fixation of additional deleterious mutations in nonrecombining mitochondrial genomes, thus accelerating the process of amino acid substitutions.

Introduction

The mitochondrial genome of animals is maternally inherited, generally nonrecombining with other mitochondrial lineages, and contains 13 protein-coding genes; all mtDNA-encoded polypeptides work in close association with nuclear-encoded subunits in 4 out of the 5 protein complexes involved in oxidative phosphorylation (Boore 1999). This is the terminal stage of the cellular respiration, where most of the ATP is produced, and it is evident that malfunction of such fundamental complexes would be lethal or at least severely affect fitness. Therefore, the mitochondrial genome is generally thought to evolve primarily under constant purifying selection. The possibility of directional selection, however, is gaining wider acceptance (Gerber et al. 2001; Bazin et al. 2006; Meiklejohn et al. 2007).

Several reports have indicated a preponderance of nonsynonymous/synonymous polymorphism within species compared with nonsynonymous/synonymous differences between species. This phenomenon has been demonstrated in mitochondrial genes of *Drosophila* and apes among others (Ballard and Kreitman 1994; Nachman et al. 1994; Rand et al. 1994; Templeton 1996; Hasegawa et al. 1998). This pattern is not consistent with expectations of the neutral theory, which predicts the within-species polymorphism to be the same as the between-species rate of divergence. The relatively higher within-species amino acid diversity has been interpreted as being slightly deleterious mutations that are eventually removed from the population before fixation (Nachman et al. 1994; Ballard and Kreitman 1995; Hasegawa et al. 1998).

The smaller effective population size of the mtDNA is often invoked to explain a faster accumulation of mutations in the mitochondrial genome relative to the nucleus (Hartl and Clark 1997). However, several other lineage-specific conditions that could affect mutation rate and fixation of mutations due to hitchhiking (genetic draft) may also be important; such as metabolic rate, DNA repair efficiency, generation time, fluctuations in population size, and presence of *Wolbachia* (e.g., Laird et al. 1969; Martin and Palumbi 1993; Shoemaker et al. 2004).

The order Hymenoptera has been reported to have unusually high mitochondrial substitution rates and more frequent gene rearrangements compared with other insects (Jermiin and Crozier 1994; Dowton and Austin 1995), and such elevated mutation rates could have consequences for their evolution. Although the honeybee *Apis mellifera* was the third insect mtDNA genome to be sequenced (Crozier RH and Crozier YC 1993), only recently other complete or nearly complete mtDNA genomes of Hymenoptera have become available (Castro and Dowton 2005; Castro et al. 2006).

We sequenced the nearly complete mtDNA of the 3 closely related species of the parasitic wasp *Nasonia*, *Nasonia vitripennis*, *Nasonia giraulti*, and *Nasonia longicornis*, and investigated the patterns of evolution by sequencing multiple strains per species for a subset of genes. In summary, our data show the following: 1) the AT-rich mtDNA of *Nasonia* has gone though a series of rearrangements in gene order affecting not only the relative position of some tRNAs but also several protein-coding genes; 2) impressive accelerated rates of molecular

[1] Present address: Deodoro C. S. G. Oliveira; Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, Spain

Key words: *Nasonia*, mitochondrial genome, gene organization, rates of evolution.

E-mail: Deodoro.Oliveira@uab.cat.

Mol. Biol. Evol. 25(10):2167–2180. 2008
doi:10.1093/molbev/msn159
Advance Access publication July 24, 2008

© The Author 2008. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

evolution in all mitochondrial genes, over 30 times higher than nuclear genes, giving *Nasonia* one of the highest relative mitochondrial substitution rates for any organism; 3) a significantly higher between-species frequency of fixed replacement sites relative to silent sites in comparison to within-species polymorphism for 2 mitochondrial genes, *atp6* and *atp8*—this is the opposite of what is normally found for mitochondria of animals and indicates directional selection in these genes.

We discuss factors that can lead to higher amino acid substitutions in mitochondrial genomes, including genetic hitchhiking of deleterious mutations during selective sweeps of the maternally inherited endosymbiont *Wolbachia*, and the possibility of a "Compensation-Draft Feedback," where deleterious mutations spawn compensatory mutation sweeps, which lead to fixation of additional deleterious mutations, thus creating a feedback (described in details in Discussion). The process will be greatly accelerated in mitochondrial genomes because of elevated mutation rate and complete linkage of the genome.

Materials and Methods

Samples Examined

The standard laboratory strains of the 3 species of *Nasonia* were chosen for mtDNA sequencing, RV2(u) (*N. giraulti*), IV7(u) (*N. longicornis*), and AsymC (*N. vitripennis*). A second strain of *N. vitripennis*, HiCd12, was also sequenced; this strain is interesting because it produces a high number of gynandromorphy sex-determination defects that show an influence of maternal cytotype (Kamping et al. 2007). Several field strains of *Nasonia* used for molecular analysis are described in table S1 (Supplementary Material online), as much as possible they cover the range of distribution of the 3 *Nasonia* species. Sequence information for 3 species in the closely related genus *Trichomalopsis* were included as outgroups, *Trichomalopsis sarcophagae*, *Trichomalopsis dubia*, and *Trichomalopsis* sp. All species cited are parasitic wasps belonging to the family Pteromalidae.

Primer Design, Polymerase Chain Reaction Amplification, and Sequencing

DNA was extracted from single wasps using the DNAeasy kit (Qiagen, CA) or Puregene DNA Purification Kit (Gentra Systems, MN). We used regular Taq (Invitrogen, CA) for polymerase chain reactions (PCRs) to yield amplicons up to 2 kb, and longer PCRs were performed with AmpliTaq (Invitrogen) or LA-PCR (TaKaRa Bio Inc, Japan). Standard conditions described in these kits were used. Sequencing was done at the Cornell University Biotechnology Resource Center or at Functional Genomics Center, University of Rochester.

Initial PCR and sequencing was done with animal mitochondrial primers described in Burger et al. (2007). Species-specific primers were designed based on these sequences and used in long PCR to amplify the regions downstream of *cob* and *cox1* in each species and upstream of *cox1* and downstream of *cox3* in *N. giraulti* and *N. vitripennis* (AsymC). Consequently, several additional *Nasonia*-specific primers were designed based on results of the expressed sequence tag (EST) projects (see below), covering most of the mtDNA (table S2, Supplementary Material online). They were successfully used in multiple combinations producing smaller or larger amplicons.

We used 2 mtDNA regions for analysis of polymorphism and divergence. The complete *atp6* and *atp8* genes were amplified together in a single PCR—primers NMCox3C and NMCox2b, fragment size ~1,200 bp. The sequencing was done with 4 primers, the 2 external primers cited above plus NMATP6a and NMATP6b. These internal primers were employed to generate high-quality full-length sequences. A similar approach was used for the *cox1* region—a fragment of 1,048 bp was amplified with the external primers NMCox1c and NMCox1f (also used for sequencing); internal primers for sequencing only were NMCox1d and NMCox1e. GenBank accession numbers are presented in table S1 (Supplementary Material online).

Identification and Annotation of Mitochondrial and Nuclear Genes

An unpublished EST data set for 2 *Nasonia* species, *N. vitripennis* and *N. giraulti*, was very useful for this study (Desjardins CA, Oliveira DCSG, Dang PM, Hunter WB, Colbourne JK, Tettelin H, Werren JH unpublished data). Sequences for all mitochondrial protein-coding genes and the 2 rRNAs were found in the ESTs. These sequences were used to design *Nasonia*-specific primers as described above (table S2, Supplementary Material online). Furthermore, 13 coding regions of functionally diverse nuclear genes present in the ESTs allowed us to estimate rates of divergence for nuclear genes in *Nasonia*. Blast searches using either the insect orthologs or available *Nasonia* sequences were employed to identify genes of interest (Altschul et al. 1990). Individual EST reads were pulled from each of the *N. vitripennis* and *N. giraulti* databases, and they were assembled using Sequencher 4.7 software (Gene Codes Corp., MI) and annotated manually. We used the online program tRNAscan-SE v.1.21 (Lowe and Eddy 1997) to annotate the mitochondrial tRNAs.

Genetic Divergence, Molecular Evolution, and Phylogenetics

The software packages DnaSP 4.0 (Rozas et al. 2003) and Mega 4.0 (Tamura et al. 2007) were used to compute DNA polymorphism and divergence, including the number of synonymous substitutions per synonymous site (Ks) and the number of nonsynonymous substitutions per nonsynonymous site (Ka). DnaSP was also used to perform the McDonald and Kreitman (MK) neutrality test (McDonald and Kreitman 1991). Maximum likelihood (ML) estimates of synonymous and nonsynonymous substitution rates were obtained with the package of programs PAML version 4 (Yang 2007); 4 nucleotide frequencies (F1X4) were assumed to calculate codon frequencies in tree-based Ka/Ks estimates.

Phylogenetic trees were constructed using Paup* 4.0 (Swofford 2002). For the maximum parsimony (MP)

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

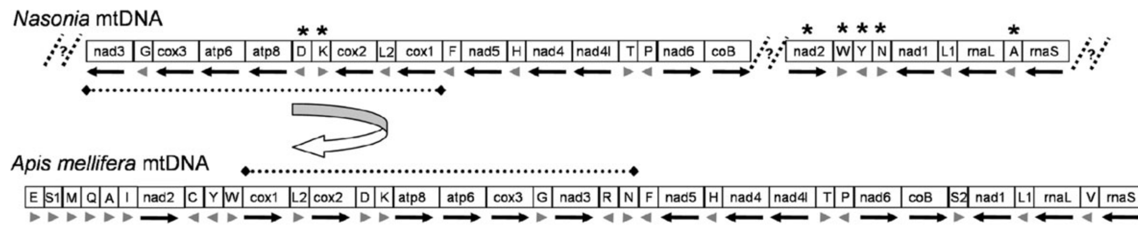

FIG. 1.—The mitochondrial genome of *Nasonia*. Linearized map showing gene order in the 2 fragments of the *Nasonia* mtDNA. The complete mitochondrial genome of the honeybee *Apis mellifera* (GenBank accession number L06178) is shown for comparison. A large inversion encompassing 6 protein-coding genes and 4 tRNAs is highlighted. Asterisks indicate other rearrangements in gene order, including the position of *nad2*. Unknown junctions are indicated with a question mark. See text for details.

method, heuristic searches for combined or individual mitochondrial partitions were done with 1,000 replicates (Tree Bisection-Reconnection algorithm). For the ML method, 100 replicates were conducted using the HKY85 model. Measures of support for each branch were generated for the combined MP analyses; bootstrap proportions (Felsenstein 1985, 1988) and Jackknife (Farris et al. 1996) methods were performed with 1,000 replicates and 100 random sequence additions. MacClade 4.08 (Maddison DR and Maddison WP 2000) was used to visually expect the alignment and to place amino acid replacements in the tree branches.

Results

Nasonia M*itochondrial Genome*

The *Nasonia* mitochondrial genome was sequenced with a combination of long and short PCR. Initially, the *N. giraulti*, *N. vitripennis* (AsymC strain), and *N. longicornis* mtDNA genomes were amplified by long PCR using primers designed for well-conserved regions of several mitochondrial genes; however, the success of this approach was limited mostly to the region downstream of *cob* and *cox3*. The completion of an EST project of *N. giraulti* and *N. vitripennis* facilitated the mtDNA sequencing because mitochondrial sequences for all 13 proteins and the 2 rRNAs were found to be present in the ESTs . This allowed us to use *Nasonia*-specific primers (table S2, Supplementary Material online) to expand the mtDNA. A second strain of *N. vitripennis*, HiCd12, was entirely sequenced with the *Nasonia*-specific primers. As detailed below, with this procedure 2 large mitochondrial regions were obtained for each species. However, several PCR conditions using all possible combinations failed in joining these 2 fragments. Difficulties in finishing mitochondrial genomes of Hymenoptera have been reported before (Castro and Dowton 2005; Castro et al. 2006), with a possible explanation being the presence of long duplicated segments. In this regard, the availability of message from the EST project made us confident that the mitochondrial genes reported here are expressed.

All 13 protein-coding genes and the 2 rRNAs are present in the 2 large fragments obtained for *N. vitripennis* (AsymC) and *N. giraulti*, the 2 genomes with the most complete coverage (fig. 1). The longer fragment, ~9,900 bp, contains 11 protein-coding genes, and the shorter fragment, ~4,500 bp, contains the 2 rRNAs and the 2 remaining protein-coding genes, *nad1* and *nad2*. In total, we were able to generate ~14,400 bp of mitochondrial sequences for *Nasonia* or about 1-kb short of a typical bilaterian mtDNA size (Boore 1999). For *N. vitripennis* (AsymC) and *N. giraulti*, only 2 protein-coding genes at the ends of the sequenced fragments, *nad2* and *nad3*, do not have full coding sequence (table 1). The large ribosomal RNA (*rnl* or 16s rRNA) is also complete, whereas only partial sequence is available for the small rRNA (*snl* or 12s rRNA). Out of the 22 tRNAs normally present in animal mitochondrial genome, only 13 were identified in sequenced regions (table 1 and fig. 1). This is not surprising, given that the missing links contain regions expected to have clusters of tRNAs—mtDNAs from other Hymenoptera species contains 5–6 tRNAs at the 5′ of *nad2* and 2–3 tRNAs at the 3′ of *nad3* (Boore 1999; Chandra et al. 2006). Table 1 also presents details for the other 2 mtDNAs described here, for which slightly smaller portions of mtDNA were obtained. The sequence of *nad2* was the only protein-coding gene not generated for *N. longicornis*. Finally, for the second strain of *N. vitripennis*, HiCd12, a partial sequence of *nad2* was generated; however, the region linking *nad1* and *nad2* is missing. All sequences have been submitted to GenBank (accession numbers EU746609–EU746617).

Table 1 presents a comparative summary of the mtDNA gene features for the 3 species of *Nasonia*. All predicted initiation codons have been observed in mtDNA of other bilaterian animals, most are methionine and a few are isoleucine (Crozier RH and Crozier YC 1993); isoleucine (ATT) initiation codons are found in *cox2* and *atp8*. The latter gene shows interspecies polymorphism, as in *N. vitripennis*, it starts with a methionine (ATA) but starts with isoleucine in *N. longicornis* and *N. giraulti*. Incomplete stop codons, single T, are common in DNA sequence, and completion of the stop codons by posttranscriptional polyadenylation is clear when the EST-originated sequences are examined. Only one predicted tRNA gene, tRNA-Tyr, has the same size in *N. vitripennis* and *N. giraulti* (it is missing for *N. longicornis*), all other 12 tRNAs differ in size between species, but tRNA size is always the same for the 2 *N. vitripennis* strains. The *rnl* also varies in size when the 3 *Nasonia* species are compared. In addition, it accounts for the only dissimilarity in gene size encountered between the 2 *N. vitripennis* strains: HiCd12 has one nucleotide less than AsymC. In protein-coding genes, divergence in size is restricted to 3 genes: the *N. longicornis nad6* is 2 codons shorter than in the other 2 species, the *N. vitripennis cox2* is

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Table 1
Comparative Gene Features in the 4 *Nasonia* Mitochondrial Genomes

		Nasonia vitripennis[a]				*Nasonia giraulti*			*Nasonia longicornis*		
		AsymC	HiCD12			RV2			IV7		
Gene	Strand	Length[b]	Length	Start	Stop	Length	Start	Stop	Length	Start	Stop
rns	−	722[c]	603[c]			737[c]			601[c]		
trnA	−	64	64			64			66		
rnl	−	1,331	1,330			1,311			1,320		
trnL1(UAG)	−	67	67			71			70		
nad1	−	933	667[c]	ATA	TAA	933	ATA	TAG	693[c]	ATA	NA
trnN	+	67	NA			68				NA	
trnY	−	68	NA			68				NA	
trnT	+	67	NA			66				NA	
nad2	+	975[c]	758[c]	NA	TAA	975[c]	NA	TAA		NA	
Missing link											
nad3	−	349[c]	144[c]	ATA	NA	349[c]	ATA	NA	129	ATA	NA
trnG	−	68	68			66			67		
cox3	−	783	783	ATG	TAA	783	ATG	TAA	783	ATG	TAA
atp6	−	672	672	ATG	TAA	672	ATG	T	672	ATG	T
atp8	−	156	156	ATA	TAA	156	ATT	TAA	162	ATT	TAA
trnD	−	65	65			67			66		
trnK	+	74	74			73			73		
cox2	−	666	669	ATT	TAA	666	ATT	TAA	666	ATT	TAA
trnL2(UAA)	−	68	68			66			66		
cox1	−	1,557	1,557	ATA	T	1,557	ATA	T	1,557	ATA	T
trnF	−	65	65			66			64		
nad5	−	1,686	1,686	ATA	T	1,689	ATA	TAG	1,689	ATA	TAA
trnH	−	65	65			63			64		
nad4	−	1,338	1,338	ATG	T	1,338	ATG	T	1,338	ATG	T
nad4l	−	291	291	ATA	TAA	291	ATA	TAA	291	ATA	TAA
trnT	+	65	65			64			65		
trnP	−	65	65			65			66		
nad6	+	546	546	ATG	TAA	546	ATG	TAA	540	ATG	TAA
cob	+	1,137	831[c]	ATG	NA	1,137[c]	ATG	TAA	838	ATG	NA

[a] Start and stop codons are the same in the 2 *N. vitripennis* mitochondrial genomes.
[b] The length of protein-coding genes is given excluding the stop codon.
[c] Genes with partial sequence.

1 codon shorter than in the other 2 species, and *atp8* is longer in *N. longicornis*. The *atp8* gene was found to be polymorphic in size within *N. longicornis* strains. The cause of this variation is an AAT triplet that behaves like a microsatellite without losing the frame, and this region is translated into a variable stretch of 5–8 asparagines in the *N. longicornis* strains examined.

The *N. longicornis* mitochondrial genome displays a slightly higher level of total AT content than the other 2 species of *Nasonia*, 86% compared with 83%. These figures are similar to those normally found in Hymenoptera. The mtDNA of the honeybee *A. mellifera*, for example, is composed of 85% AT (Crozier RH and Crozier YC 1993). Hymenopteran species are, however, more AT-rich in their mtDNAs when compared with the other insect orders. The normal base composition for insects excluding Hymenoptera falls in the range of 70–78% AT content (Chandra et al. 2006). A notable exception is *Drosophila melanogaster*, with an 82% AT content due to its exceptionally long control region (AT-rich region)—96% AT in the 4.6-kb long control region (Lewis et al. 1994). Because the control region is missing for *Nasonia*, the AT bias for the complete mitochondrial genome might be even more pronounced.

The *Nasonia* Mitochondrial Genome Has a New Gene Order—Comparison with Other Insects

All 3 species of *Nasonia* have the same gene order (fig. 1). When compared with other insects, it is clear that a series of gene rearrangements have occurred in the evolutionary history of these wasps. The most striking one is a large inversion that affects the relative position of 6 protein-coding genes. Although tRNAs differing in location are not uncommon in insects, the relative position of the 13 protein-coding genes is strongly conserved (Boore 1999; Chandra et al. 2006). Moreover, a second rearrangement is necessary to account for the position of *nad2*. Among the tRNAs, the most impressive novelty is the presence of the tRNA-Ala instead of the tRNA-Val in the middle of the 2 rRNAs. Because the presence of tRNA-Val between the small and large RNAs was a condition previously encountered in all arthropods (Boore 1999). The organization occurring in *Nasonia* for tRNA-Asp and tRNA-Lys is also interesting. This region, the junction of *cox2* and *atp8*, has been described as a "hot spot" for rearrangements in Hymenoptera, in reference to variation in the relative orientation and number of tRNAs found in this location (Dowton and Austin 1999). The arrangement for tRNA-Asp and tRNA-Lys in *Nasonia* (fig. 1) is shared only

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Table 2
Average (Standard Deviation) Interspecific Pairwise Sequence Divergence for All Mitochondrial Genes

	Pi	Ks (JC)	Ka (JC)	Ka (JC)/Ks (JC)
Nasonia				
mtDNA genes[a]				
Nasonia vitripennis × *Nasonia giraulti*	0.148 (0.038)	0.519 (0.086)	0.100 (0.058)	0.197 (0.112)
N. vitripennis × *Nasonia longicornis*	0.148 (0.047)	0.476 (0.064)	0.108 (0.073)	0.227 (0.132)
N. longicornis × *N. giraulti*	0.099 (0.017)	0.389 (0.094)	0.051 (0.028)	0.152 (0.117)
Nuclear genes[b]				
N. vitripennis × *N. giraulti*	0.009 (0.003)	0.031 (0.012)	0.003 (0.003)	0.146 (0.250)
Drosophila[c]				
mtDNA genes	0.046 (0.006)	0.189 (0.47)	0.014 (0.007)	0.083 (0.055)

NOTE.—DNA polymorphism between pairs of sequences in protein-coding regions; Pi, uncorrected nucleotide diversity, and estimates of Ka (the number of nonsynonymous substitutions per nonsynonymous site) and Ks (the number of synonymous substitutions per synonymous site) using the Jukes and Cantor (1969) correction.

[a] Estimates involving *N. longicornis* were calculated for only 12 mitochondrial genes because data for *nad2* are missing. Data for *N. vitripennis* were obtained using sequences of the AsymC strain, results for HiCD12 are similar (data not shown).

[b] Thirteen functionally diverse nuclear genes.

[c] Average sequence divergences between *Drosophila melanogaster* (U37541) and *Drosophila simulans* (AF200833) for all 13 mitochondrial proteins is shown for comparison. See text for details.

with *Encarsia formosa* (Family Aphelinidae; Dowton and Austin 1999); this might be a synapomorphy for the superfamily Chalcidoidea, to which these parasitic wasps belong.

The *Nasonia* Mitochondrial Genome Evolves Fast—Between-Species Comparisons

The most prominent characteristic of the *Nasonia* mtDNA is that it is evolving very fast. *Nasonia vitripennis* forms the sister group to the other 2 species, and the average uncorrected nucleotide diversity (Pi) for all mitochondrial protein-coding genes is 15% when compared with *N. longicornis* or *N. giraulti*, whereas *N. longicornis* and *N. giraulti* show 10% divergence form each other in their mtDNA (table 2). These numbers are higher than those previously estimated by van Opijnen et al. (2005); this is not surprising because they used only the *cox1* region which, as described below, is the most conserved gene in the *Nasonia* mitochondria (table 2 and fig. 2). It is important to note that these 3 species of *Nasonia* are very close morphologically and can be easily crossed in the laboratory (Darling and Werren 1990; Breeuwer and Werren 1990, 1995). Therefore, it was surprising to find out that their mitochondrial genomes were so divergent.

The fast accumulation of mutations in the mtDNA of *Nasonia* becomes more evident when balanced by the divergence in the coding region of nuclear genes (table 2). When only synonymous positions are considered, mitochondrial protein-coding genes are evolving 17–35.5 times faster than nuclear genes in *Nasonia*. Nuclear divergence estimates were first performed using EST data for 13 genes that are available for *N. vitripennis* and *N. giraulti* —divergence of 3.1% in nuclear genes and 52% in mitochondrial genes (*N. vitripennis–N. giraulti*; with Jukes and Cantor [JC] correction). Although an EST data set has not been generated for *N. longicornis*, we have preliminary data in an ongoing project dealing with nuclear diversity and divergence in the genus *Nasonia* (these results will be presented elsewhere). Based on 7 nuclear gene regions, Ks for *N. longicornis–N. giraulti* is approximately 1.1%, whereas the mtDNA Ks for these 2 species is 39% (table 2). It assigns to *Nasonia* a mutation rate in the mitochondrial DNA that is 35 times more accelerated than the nuclear genome. Due to possible saturation problems in the *N. vitripennis–N. giraulti* mitochondrial divergence, the 35-fold greater mitochondrial/nuclear divergence found in *N. giraulti–N. longicornis* is our current best estimator of the relative mitochondrial substitution rate within the genus *Nasonia*. Assuming proportional rates of synonymous substitution in the nucleus and in the mitochondria throughout the evolution of the genus *Nasonia*, it is possible to estimate that the mitochondrial synonymous divergence should be approximately 110%, or 2.8 times higher (3.1%/1.1%) between *N. giraulti–N. vitripennis* (or *N. vitripennis–N. longicornis*) than that of *N. giraulti–N. longicornis*, whereas it is only 1.3 times higher (52%/39%), suggesting that saturation is an issue between *N. giraulti* and *N. vitripennis*. The amino acid replacement rate is also accelerated for mitochondrial-encoded proteins; it is over 30-fold higher than for nuclear-encoded proteins (table 2). However, the Ka/Ks ratios in the mtDNA and nuclear genes are of similar magnitude (table 2).

A Comparison between *Nasonia* and *Drosophila* Mitochondrial Genomes Evolution

For comparative purposes, we generated similar data for the mitochondrial genome of *D. melanogaster* and *Drosophila simulans*. The nuclear Ks for these flies is about 8%—average of 19 nuclear metabolic genes reported by Flowers et al. (2007), much larger than between the *Nasonia* species (Ks = 1.1% between *N. giraulti–N. longicornis* and 3.1% between *N. giraulti–N. vitripennis*). Yet, these flies have a much less divergent mtDNA (Ks = 19%), even when compared with *N. giraulti–N. longicornis*, the 2 closely related *Nasonia* species (Ks = 39%; table 2 and fig. 2). This yields a relative mitochondrial to nuclear synonymous substitution ratio for *N. giraulti–N. longicornis* of 35.5 compared with only 2.4 for the *D. melanogaster—D. simulans* species pair. The *Nasonia* lineage has a relative mitochondrial substitution rate approximately 15 times

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

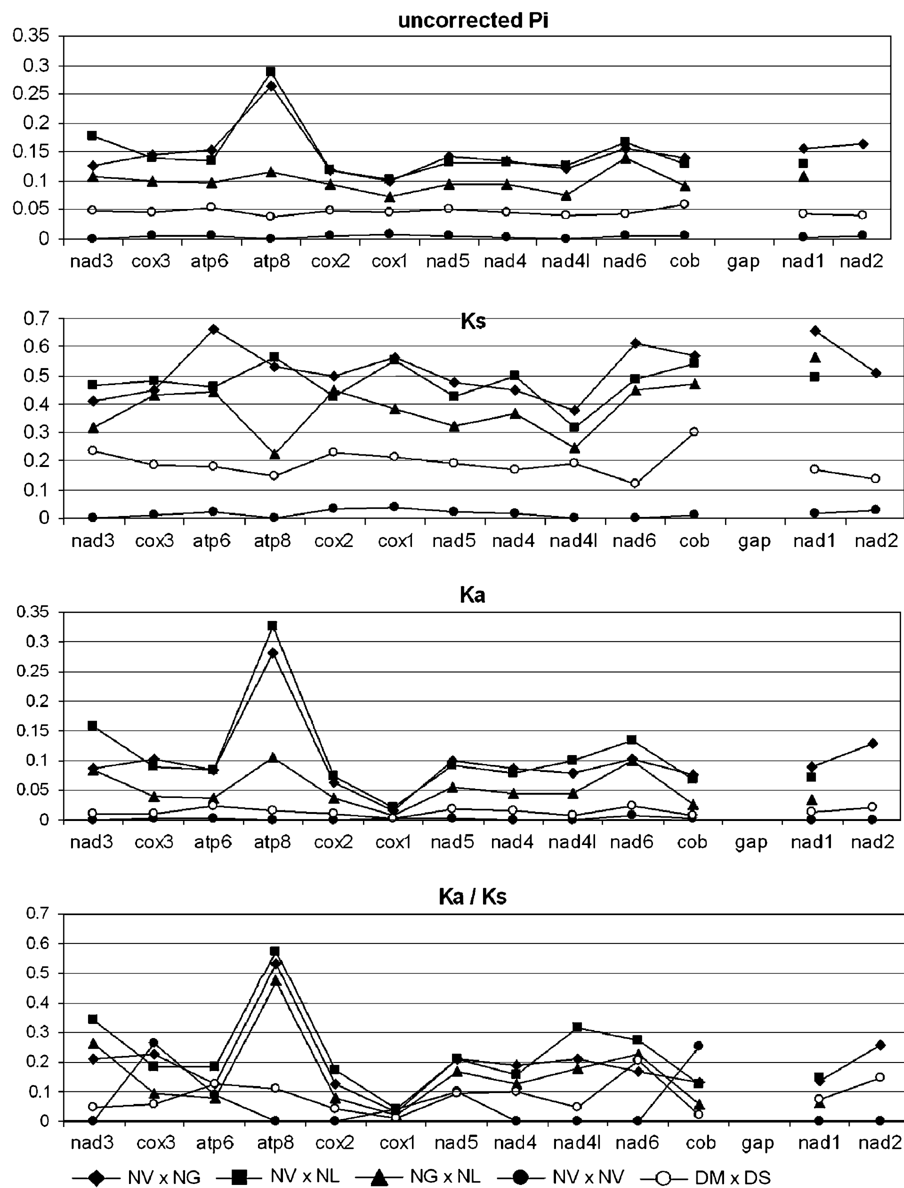

FIG. 2.—Pairwise comparison of sequence divergence for each mitochondrial gene. Genes are ordered according to their position in the mitochondrial genome of *Nasonia*. The mtDNA sequences of *Nasonia vitripennis* (NV) strain AsymC was used for between-species calculation; *Nasonia vitripennis* × *Nasonia vitripennis* compares AsymC and HiCD12 strains of *N. vitripennis*. Values of Ks and Ka are corrected by the JC method. Similar statistics for *Drosophila melanogaster* (DM; U37541) and *Drosophila simulans* (DS; AF200833) is shown for comparison.

greater than *Drosophila*. This contrasts with the estimated times of divergence, which are higher for *Drosophila*. The splitting time between *D. melanogaster–D. simulans* has been estimated to be around 2.5 MYA (Hey and Kliman 1993; Russo et al. 1995). The *Nasonia* species are presumed to be much younger. Two independent calculations estimated that *N. giraulti–N. longicornis* speciated around 0.4 MYA, one based on nuclear ribosomal ITS (internal transcribed spacer; Campbell et al. 1993) and one based on codiverging cytoplasmic bacteria *Wolbachia* (Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH unpublished data). Consistent with this view, hybrids of *Nasonia* species are fertile (Breeuwer and Werren 1995), in contrast to hybrids of *D. melanogaster* and *D. simulans*. A simple calculation, assuming that all mtDNA synonymous mutations behave in a clock-like manner, reveal that the *Nasonia* mtDNA (39%/0.4 MYA) evolves 13 times faster than does *Drosophila* mtDNA (19%/2.5 MYA). The mitochondrial genomes of these 2 *Drosophila* species also show lower Ka and uncorrected Pi values than *Nasonia* (table 2 and fig. 2).

Although, nucleotide diversity between *Nasonia* species is very high for any pair of mitochondrial protein-coding genes (fig. 2), 2 genes deviate markedly from the average values: *atp8* and *cox1*. The highest Ka/Ks ratios were found always for *atp8*. Interestingly, in the case of *N. longicornis atp8* and the *N. giraulti atp8* alleles, both a higher Ka and a lower Ks account for the elevated Ka/Ks. This small mitochondrial gene, 156 bp, accumulates amino acid substitutions 3 times faster than the *Nasonia*

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Table 3
The MK Test of Neutrality for 3 mtDNA Genes

	Fixed		Polymorphic			
	S	R	S	R	*P*-value[a]	NI[b]
atp6						
Nasonia longicornis × *Nasonia giraulti*	29	15	36	10	0.24	0.537
Nasonia vitripennis × *N. longicornis*	37	35	42	10	0.0012**	0.252
N. vitripennis × *N. giraulti*	41	37	47	4	0.000001***	0.094
atp8						
N. longicornis × *N. giraulti*	2	10	13	9	0.029677*	0.138
N. vitripennis × *N. longicornis*	9	30	12	9	0.011619*	0.225
N. vitripennis × *N. giraulti*	10	27	10	7	0.035156*	0.259
cox1						
N. longicornis × *N. giraulti*	49	2	98	2	0.603745	0.500
N. vitripennis × *N. longicornis*	65	10	109	6	0.062252	0.358
N. vitripennis × *N. giraulti*	70	8	90	4	0.143079	0.389
The MK test excluding rare (<10%) polymorphism[c]						
atp6						
N. longicornis × *N. giraulti*	36	16	22	6	0.439242	0.613
N. vitripennis × *N. longicornis*	39	35	31	6	0.001606**	0.215
N. vitripennis × *N. giraulti*	50	38	29	2	0.000117***	0.091
atp8						
N. longicornis × *N. giraulti*	3	10	6	5	0.205983	0.250
N. vitripennis × *N. longicornis*	10	30	5	5	0.245880	0.333
N. vitripennis × *N. giraulti*	13	27	6	2	0.044639*	0.160
cox1						
N. longicornis × *N. giraulti*	58	2	53	0	0.497472	ND
N. vitripennis × *N. longicornis*	73	10	62	1	0.023869*	0.117
N. vitripennis × *N. giraulti*	84	8	45	1	0.171594	0.233

NOTE.—The MK test compares the ratio of replacement (R) to silent (S) fixed substitutions between species and the ratio of replacement (R) to silent (S) polymorphism within species, which should be the same under neutral evolution.

[a] Fisher's exact test (2-tailed): *$0.01<P<0.05$; **$0.001<P<0.01$; ***$P<0.001$.

[b] NI is defined as the ratio of the number of polymorphic R sites to the number of fixed R sites divided by the ratio of the number of polymorphic S sites to the number of fixed S sites (Rand and Kann 1996). NI < 1 indicates an excess of amino acid divergence relative to polymorphism.

[c] Rare segregating within-species polymorphism, frequency smaller than 10%, were removed and only common polymorphism was used for the MK test, the fixed between-species mutations were corrected accordingly. This procedure was adapted from Gojobori et al. (2007) and was done in an attempt to correct for possible saturation in the silent sites (see Results for details).

mtDNA average (fig. 2). This can be considered an indication of positive selection acting at the *atp8* gene or due to relaxed selection. In the mitochondrial genome of *D. melanogaster* and *D. simulans*, the *atp8* gene does not show evidence of accelerated evolution, either by positive or relaxed selection (fig. 2). In fact, only 2 amino acid changes occurred during the history of those 2 species of *Drosophila*, which is reasonable evidence that the *atp8* gene is not always prone to accumulate slight deleterious mutations. On the other side of the spectrum, *cox1* shows the smallest Ka/Ks rates of any mitochondrial gene in *Nasonia* and consequently is under stronger purifying selection. Despite the fact that *cox1* is the *Nasonia* mitochondrial gene least susceptible to accumulate changes in the protein sequence, out of 519 amino acids there are still 11 replacements between *N. longicornis–N. giraulti*, 17 between *N. giraulti–N. vitripennis*, and 23 between *N. longicornis–N. vitripennis*. In *D. melanogaster–D. simulans*, *cox1* is also the slowest evolving mitochondrial gene; there are only 2 amino acid changes and the Ka/Ks is smaller than found for *Nasonia* (fig. 2; *D. melanogaster–D. simulans* = 0.008, *N. giraulti–N. longicornis* = 0.024, *N. giraulti–N. vitripennis* = 0.030, *N. longicornis–N. vitripennis* = 0.040). The comparison brings further confirmation that the *Nasonia* mtDNA is evolving rapidly, and this likely indicates the influence of an elevated mutation rate in these species and (or) hitchhiking during mitochondrial sweeps (discussed below).

Directional Selection in 2 Mitochondrial Genes: *atp6* and *atp8*

To further investigate whether changes in amino acid sequence of these genes are due to relaxed or directional selection, we sequenced mitochondrial genes from several field strains for each of 3 species of *Nasonia*. Three mtDNA genes were chosen for sequencing *atp8*, *atp6*, and *cox1*. When the *atp8* gene, the mtDNA gene with higher Ka/Ks value, is submitted to the MK test of neutrality, the test indicates that *atp8* has a significantly higher number of fixed replacements when compared with within-species polymorphism (table 3); this is true for all 3-way comparisons within *Nasonia*. The higher between-species number of polymorphism is interpreted as evidence of directional selection. The *atp6* gene has an even more significant

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Table 4
Sequence Diversity (Within Species) and Divergence (Between Species) for 3 Mitochondrial Genes

		atp8				*atp6*				*coxI*			
		Pi	Ka	Ks	Ka/Ks	Pi	Ka	Ks	Ka/Ks	Pi	Ka	Ks	Ka/Ks
Within species													
Nasonia giraulti	JC	0.013	0.006	0.047	0.116	0.010	0.001	0.045	0.021	0.011	0	0.053	0
	MNG	—	0.006	0.030	0.215	—	0.001	0.028	0.039	—	0	0.034	0
	ML	—	0.005	0.105	0.047	—	0.001	0.078	0.011	—	0	0.142	0
Nasonia longicornis	JC	0.025	0.014	0.077	0.183	0.012	0.004	0.043	0.101	0.018	0.001	0.083	0.004
	MNG	—	0.017	0.047	0.358	—	0.005	0.027	0.192	—	0.001	0.054	0.007
	ML	—	0.014	0.195	0.069	—	0.004	0.124	0.030	—	0.001	0.366	0.001
Nasonia vitripennis	JC	0.009	0.005	0.026	0.190	0.010	0.001	0.044	0.015	0.010	0.001	0.044	0.024
	MNG	—	0.006	0.017	0.322	—	0.001	0.029	0.028	—	0.001	0.030	0.040
	ML	—	0.005	0.043	0.107	—	0.001	0.106	0.006	—	0.001	0.107	0.008
Between species													
N. giraulti/*N. longicornis*	JC	0.115	0.108	0.207	0.523	0.092	0.037	0.396	0.094	0.078	0.005	0.444	0.011
	MNG	—	0.127	0.120	1.063	—	0.044	0.223	0.198	—	0.006	0.261	0.023
	ML	—	0.100	0.608	0.165	—	0.033	1.631	0.020	—	0.004	3.419	0.001
N. giraulti/*N. vitripennis*	JC	0.264	0.286	0.514	0.556	0.146	0.082	0.580	0.141	0.111	0.016	0.672	0.023
	MNG	—	0.340	0.294	1.157	—	0.098	0.315	0.311	—	0.019	0.373	0.050
	ML	—	0.270	2.071	0.131	—	0.073	2.538	0.029	—	0.014	3.544	0.004
N. longicornis/*N. vitripennis*	JC	0.283	0.326	0.499	0.653	0.129	0.081	0.435	0.186	0.107	0.020	0.593	0.033
	MNG	—	0.390	0.280	1.393	—	0.097	0.244	0.396	—	0.023	0.338	0.068
	ML	—	0.318	1.733	0.184	—	0.071	2.276	0.031	—	0.016	3.514	0.005

NOTE.—DNA polymorphism in protein-coding regions between and within species was calculated as the average of all pairwise comparisons (21 sequences for *N. vitripennis*, 11 for *N. giraulti*, and 16 for *N. longicornis*); nucleotide diversity, Pi (uncorrected), and estimates of Ka, the number of nonsynonymous substitutions per nonsynonymous site, and Ks, the number of synonymous substitutions per synonymous site are shown. The Ka and Ks values were corrected for multiple hits with the Nei–Gojobori method (JC correction) or by the MNG that accounts for the transition/transvertion (R) ratio, as implemented in Mega 2.0 (Tamura et al. 2007). The average within-species R for each gene was used in both within-species and between-species MNG corrections: *atp8* $R = 2.16$, *atp6* $R = 3.8$, *cox1* $R = 5.3$. Pairwise ML estimates were obtained with PAML 4 (Yang 2007) using the program yn00.

excess of replacements when *N. vitripennis* is included in the analysis, but not between *N. longicornis–N. giraulti* (table 3). This result is surprising because *atp6* does not show an elevated Ka/Ks compared with average mitochondrial genes (fig. 2). This finding opens the possibility that similar phenomenon might occur in other mitochondrial genes not investigated in this work. The *cox1* gene does not show a significant increase in fixed nonsynonymous differences compared with within species, a pattern compatible with a gene that is under purifying selection. Furthermore, all genes, even *cox1*, have the neutrality index (NI) lower than 1 (table 3). An NI value lower than 1 is consistent with the idea that the *Nasonia* mtDNA is not evolving under a strictly neutral process and that positive selection can be an important force shaping mtDNA divergence (Bazin et al. 2006; Meiklejohn et al. 2007).

Saturation at silent sites in highly divergent sequences is a concern that has been raised for the contingence-based MK test (Hasegawa et al. 1998). This problem seems to be even more likely in AT-rich genomes like the mtDNA of *Nasonia*. Multiple silent changes at the same position are often counted only as polymorphic and not as fixed substitutions, although the common nucleotide observed in each species are not the same. In an attempt to correct for an underestimation of fixed silent substitutions, we removed rare polymorphism from the MK test (table 3). A similar method to distinguish rare and common polymorphism has been applied by others (Fay et al. 2002; Gojobori et al. 2007; Shapiro et al. 2007), although their purpose was to exclude rare polymorphic deleterious amino acid substitutions. Our data show that besides removing both rare silent and replacement polymorphism for all genes, this method indeed increases the amount of fixed silent changes, affecting little the amount of fixed replacements (table 3). Results for the *atp8* gene are the most affected and only the MK test for *N. vitripennis–N. giraulti* remains statistically significant, whereas *N. longicornis–N. giraulti* and *N. vitripennis–N. longicornis* comparisons became nonsignificant, although NI stays below 1 in all cases. There is no alteration in the MK test of *atp6*, and the highly significant excess of fixed replacements found for *N. vitripennis–N. longicornis* and *N. vitripennis–N. giraulti* still indicates that directional selection has shaped the evolution of this gene. Surprisingly, the *N. vitripennis–N. longicornis* MK test of *cox1* appears now to have a significant excess of replacements. Exclusion of rare polymorphism affects *cox1* by removing most of within-species segregating replacements, which likely are mild-deleterious mutations.

Ka and Ks Ratios among and between *Nasonia* Species

Another approach that is likely to correct in part for the possible saturations at silent positions is to compare the Ka and Ks rates between and within species. Similarly to the MK test, under a neutral hypothesis, the Ka/Ks ratio should be the same within and between species (Hasegawa et al. 1998). Three different methods to estimate the within- and between-species Ka and Ks rates were compared (table 4). They are either correct by the Nei–Gojobori method (JC) that does not account for the transition/transversion ratio (R) or by a modified Nei–Gojobori method (MNG) with

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

takes *R* into account. The estimated within-species *R* for *Nasonia* is used (see below), which better reflects the probability of a mutation. ML pairwise estimates were also obtained (table 4, Yang and Nielsen 2000).

Correcting for the transitions and transversion rates lowers the Ks and increases the Ka, comparing JC and MNG, what approximately doubles all Ka/Ks values (table 4). The *atp8* Ka/Ks corrected by the MNG method reaches values >1. Consistently, the between-species Ka/Ks are 3–12 times higher than the within-species ratios for the 3 genes analyzed. Again, this is opposite to the common finding in other animal mitochondria of a higher within-species Ka/Ks due to segregating mild-deleterious mutations (Hasegawa et al. 1998; Nachman 1998; Rand and Kann 1998). A higher between-species Ka/Ks is most likely due to fixation of beneficial replacements. There are 2 exceptions, *N. longicornis* Ka/Ks for *atp6* and the *N. vitripennis* Ka/Ks for *cox1* are of similar magnitude to the respective between-species values.

The observation that between-species Ks corrected by the MNG method are smaller than corrected by JC suggests that saturation at synonymous sites might be less a concern (table 4). However, the *N. vitripennis*–*N. giraulti* (or *N. vitripennis*–*N. longicornis*) Ks is never as high as 2.8 times the *N. longicornis*–*N. giraulti* Ks, a level that would be predicted if we use the nuclear divergence comparison. This observation may argue that saturation has not been completely compensated for in those comparisons. Alternatively, the difference could be revealing a real decrease in genetic diversity in synonymous positions caused by recurrent sweeps of the mitochondrial genome (Bazin et al. 2006; Meiklejohn et al. 2007).

ML pairwise estimates, on the other hand, greatly increased values of Ks, not only between species (2.9- to 7.7-fold, compared with JC) but also within species (1.6- to 4.4-fold), and therefore, Ka/Ks ratios are much smaller (table 4). Although less pronounced, the general trend of higher between-species Ka/Ks ratios was preserved, particularly for *atp8*. Nonsynonymous substitution rates are very similar to those obtained with other methods. However, there are some inconsistencies with the ML estimates; the between-species Ks values showed a 5.6-fold variation among the 3 mtDNA genes, and for *cox1* the 2 closely related species *N. longicornis*–*N. giraulti* appear to be as divergent from each other, by synonymous substitutions, as they are from *N. vitripennis*. These results are difficult to explain because a homogeneous synonymous divergence is expected for all mitochondrial genes. The findings could reflect differential purifying selection on synonymous sites across the mitochondrial genome.

Codon-based ML was also used to compare a model that fits a single or 2 independent Ka/Ks ratios, a between-species ratio and a within-species ratio, in the phylogenetic tree branches (fig. 3; Hasegawa et al. 1998). Consistent with previous results, the between-species ratio is almost twice as high for *atp6*: Ka/Ks between species = 0.085; Ka/Ks within species = 0.046. The 2 Ka/Ks models also provide a better fit of the data: 2 Ka/Ks log-likelihood = -1757.75; 1 Ka/Ks log-likelihood = -1759.03. However, the likelihood ratio test (2Δlog-likelihood) fails to reject the null hypothesis. Therefore, the 2 ratios model is not significantly better than a model that fits only 1 Ka/Ks ratio throughout the tree. Results for the other 2 genes *atp8* and *cox1* are comparable (data not shown).

The ML results suggest that a high substitution rate and a rather neutral process could explain the pattern of evolution of the *Nasonia* mtDNA. We argue, however, that several mitochondrial sweeps are difficult to model under an ML framework because they violate the stationary assumption in the ML analyses. Therefore, the ML framework is less appropriate for these data.

Transition and Transversion Rates

When among- and between-species *R* are compared for all sites, among-species *R* are much higher than the between-species *R* in *Nasonia*. The respective within-species and between-species *R* (average for the 3 species or average for the 3 possible between-species comparison) are 5.3 and 0.75 for *cox1*, 3.8 and 0.65 for *atp6*, and 2.16 and 0.35 for *atp*8. This suggests that transitions are happening more frequently than expected in the mtDNA, mostly A $>$ G or T $>$ C due to the high AT content. Results obtained using only 4-fold degenerated sites corroborated the finding for all position (data not shown). An implication of this finding for silent positions is that Cs and Gs comprise a large amount of segregating within-species polymorphisms that have a low chance of becoming fixed, which could reflect selection favoring an AT-rich mitochondrial genome. Further, it also implies that some replacement pathways are much more likely to occur and to become fixed, especially if beneficial ones. Transition bias is well known for insect and mammal mtDNAs (e.g., DeSalle et al. 1987 and references therein). As shown here for *Nasonia*, in the faster evolving hymenopteran mtDNA, the transition bias is only clearly observed in within-species sequences. Because transversions erase the transition record, more divergent sequences appear to have an unrealistic transversion bias (Dowton and Austin 1997).

Phylogenetics and Within-Species Diversity

The phylogenetic relationship of the 3 species of *Nasonia* is well established (Campbell et al. 1993; van Opijnen et al. 2005). It is worth to note, however, that the large number of interspecific fixed nucleotide substitutions makes the mitochondrial genes of *Nasonia* powerfully diagnostic. The combined consensus MP tree recovers reciprocal monophyly for the 3 *Nasonia* species with robust support (fig. 3*A*); an ML analyses tree gives identical results (data not shown). This is also true for any gene partition, that is, *cox1*, *atp6*, and even the short *atp8* gene is able to resolve the evolutionary relationship of the 3 species of *Nasonia* with strong support (data not shown). In fig. 3*B*, all amino acid changes for each gene were placed in the MP tree. It shows in a graphic manner that *atp6* and *atp8* accumulated many more replacements than *cox1*. For the 2 genes in the F0F1–ATP synthase complex, the branch leading to *N. vitripennis* seems particularly rich in new amino acids.

The largest amount of intraspecific variation in the mtDNA was found for *N. longicornis* (fig. 3 and table 4),

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

FIG. 3.—(*A*) Consensus tree of the 6 most parsimonious trees for the combined mitochondrial sequences—total of 2,256 aligned nucleotides, 506 parsimony-informative characters, tree length = 988, consistency index = 0.7581, and retention index = 0.9652. Bootstrap support is shown for relevant braches, Jackknife support (with 33% deletion) was also estimated and it was always 100% for the same branches (data not shown). (*B*) Amino acid replacements for each of the 3 studied genes were placed in the trees braches under a most parsimonious framework. Tickers bars are unambiguous changes, and lighter bars are replacements for which the polarity of the change could not be determined. Brach length is proportional to all possible changes.

in agreement with previous report in van Opijnen et al. (2005). In fact, it appears that there are 2 well-supported diverging clades of mtDNA haplotypes in *N. longicornis*. Two clades of mtDNA haplotypes are also present in *N. vitripennis*. *Nasonia vitripennis* has been collected in Eurasia and North America—whereas the other species are exclusively North American in distribution. Most of the polymorphism of *N. vitripennis* is found in European strains, whereas the North American strains have very little polymorphism; a result that is in agreement with a study of *N. vitripennis* diversity performed by Grillenberger et al. (personal communication). One clade of mtDNA haplotypes seems to be encountered only in Europe, whereas the other is shared by European and North American strains. This pattern is consistent with a recent invasion and rapid expansion through out North America by

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

N. vitripennis. An alternative explanation could be a *Wolbachia*-induced mitochondrial sweep that has recently occurred in North America and is still underway in Europe.

Discussion

Here we show that the *Nasonia* mtDNA is evolving in a surprising fast pace; it shows several gene rearrangements compared with other insects, between-species variation in gene size, and a very high interspecific divergence at the DNA and protein level. It has been demonstrated that the dynamic of gene order rearrangements in the mtDNA correlates well with the rate of sequence substitution (Xu et al. 2006). Also, it has been known that genes and proteins in the mitochondria of parasitic Hymenoptera species have evolved at a faster rate than in the other insect orders (Jermiin and Crozier 1994; Dowton and Austin 1995, 1997; Castro et al. 2002). Why do parasitic hymenoptera show an elevated mitochondria substitution rate? Castro et al. (2002) tested the hypothesis that the parasitic lifestyle may enhance mutation rates by comparing parasitic and nonparasitic species in Hymenoptera and in Diptera. They did not find an elevated mitochondrial substitution rate in the Diptera, but they did find such an effect in Hymenoptera. These features, demonstrating a rapidly evolution of mitochondrial genomes in parasitic Hymenoptera, become even more apparent when the closely related species in the genus *Nasonia* are compared.

It is clear that the smaller effective population size of the mitochondria is not enough to explain the high rates of evolution. Two possibilities seem reasonable at this point. First, during its evolutionary history, *Nasonia* might have gone through one (or several) mitochondrial bottleneck(s); this would accelerate the fixation of mildly deleterious mtDNA polymorphisms that would not go to fixation in large populations. Such bottlenecks could result from selection for mitochondrial haplotypes or selective sweeps of the cytoplasmically inherited bacterium *Wolbachia*. *Wolbachia* have been shown to quickly sweep in a population and to carry along the mtDNA (Turelli et al. 1992; Shoemaker et al. 2004). All *Nasonia* field strains are infected with at least 2 types of *Wolbachia* (van Opijnen et al. 2005; Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH unpublished data). However, if such sweeps are common, it is clear that the elevated mitochondrial mutation rate found in *Nasonia* quickly restores intraspecific sequence variation. For example, *N. longicornis* has a high within-species polymorphism level (Ks $>$ 2.7%, table 4) despite having 3 different *Wolbachia* strains that have been acquired since evolution of the species (van Opijnen et al. 2005; Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH unpublished data). Furthermore, evidence presented here indicates that *atp6* and *atp8* have undergone directional selection, and therefore, the amino acid substitutions cannot be explained merely by mildly deleterious mutations.

The between-species excess of replacements/silent sites (compared with within-species) found for *Nasonia* is uncommon for animal mtDNA (Nachman 1998; Rand and Kann 1998; Gerber et al. 2001). Several reports have demonstrated that animal mtDNAs deviate from neutral evolution; however, the normal pattern is a within-species excess of replacements (Ballard and Kreitman 1994; Nachman et al. 1994; Rand et al. 1994; Templeton 1996; Hasegawa et al. 1998). We have found the opposite in *Nasonia*. To account for the excess number of segregating within-species amino acid polymorphism in most species, several authors have argued that the polymorphisms are mild deleterious and ultimately eliminated from the population before fixation (Nachman et al. 1994; Ballard and Kreitman 1995; Hasegawa et al. 1998). The finding that *Nasonia* has an excess of interspecies amino acid changes is therefore quite remarkable and the most likely explanation is that directional selection has played an important role in increasing the amino acid divergence in the genus. Although our results might be partly inflated due to the saturation between species at the synonymous sites, our analyses were performed with identical approaches that gave opposite results for other animals. Furthermore, we employed several adjustments (tables 3 and 4) to compensate for problems of saturation, and the results remained consistent with directional selection in 2 mtDNA genes, *atp8* and *atp6*.

Recently, a case was made for frequent adaptive evolution in invertebrate mitochondrial genomes (Bazin et al. 2006; Meiklejohn et al. 2007). It was based in the fact that invertebrate mtDNAs have a significant lower NI than vertebrates. When NI (polymorphic *R*/fixed *R*)/(polymorphic S/fixed S) is smaller than 1, it is indicative of an excess of amino acid variation between species. The lower NI in invertebrates is believed to be due to their larger population size, resulting in a higher rate of beneficial mutations. Larger populations both increase the probability of fixation of mildly beneficial mutations in mitochondria (due to larger effective population size) and increase the probability that beneficial mutations occur in mitochondrial haplotypes with low levels of mutational load (Peck 1994). According to this argument, recurrent fixation of these favorable alleles causes a decrease in the amount of genetic diversity, that is, it reduces the amount of synonymous polymorphism that is reflected in a smaller NI. However, the high mutation rate and high level of within-species polymorphism found in *Nasonia* argues against this explanation.

The Compensation-Draft Feedback Model

Rand et al. (2004) have pointed out that the accumulation of mildly deleterious mutations in mitochondrial genomes will result in selection for compensatory mutations, which will further increase the rate of nonsynonymous mitochondrial substitutions. Here we extend this model by incorporating the positive feedback that will result from hitchhiking of additional deleterious mutations (genetic draft) due to the spread of compensatory mutations. We refer to this as the Compensation-Draft Feedback model (fig. 4). Genetic draft is the fixation of mutations by hitchhiking during linked selection (Gillespie 2000). Because of the lack of recombination in the mitochondrial genome, any sweep of a mitochondrial haplotype to fixation will likely result in fixation of mildly deleterious alleles. Compensatory mutations, either within the same protein or interacting

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

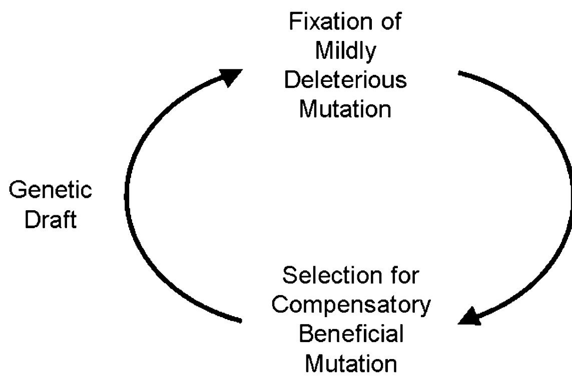

FIG. 4.—The process of Compensation-Draft Feedback is shown. Fixation of a mildly deleterious mutation favors compensatory mutations, which in turn can result in fixation of new mildly deleterious mutations by genetic draft (hitchhiking of harmful mutations in the nonrecombining mitochondrial genome). The process can be started initially by selection for a beneficial mutation, increase by drift of a mildly deleterious mutation, or selective sweeps due to associated cytoplasmically inherited microorganisms such as *Wolbachia*. Higher mitochondrial mutation rates will accelerate the process.

mitochondrial proteins, will result in a second sweep, which again can drag along mildly deleterious alleles, resulting in selection for additional compensatory mutations. Thus, any mitochondrial sweep is likely to result in a cascade of adaptive and nonadaptive amino acid substitutions, the Compensation-Draft Feedback. The role of Compensation-Draft Feedback will be strongly influenced by the mitochondrial mutation rate and, therefore, may be particularly important in species with elevated mitochondrial mutation rates, such as *Nasonia*. Compensation-Draft Feedback can also occur between mitochondrial- and nuclear-encoded genes. For example, fixation of a mildly deleterious allele in the mitochondrial component of the electron transport chain could lead to a compensatory mutation in an interacting nuclear component, which then leads to further compensatory mutations in interacting mitochondrial-encoded genes, thus resulting in further Compensation-Draft Feedback in the mitochondrial genome.

Compensation-Draft Feedback can be initiated by positive selection for a replacement substitution in a mitochondrial protein, by fixation of mildly deleterious mutations by drift, or by selection on nuclear-encoded genes that interact with mitochondrial proteins, thus favoring compensatory mutations. *Wolbachia* can also initiate Compensation-Draft Feedback. Maternally inherited endosymbiont *Wolbachia* are very common in insects (Werren and Windsor 2000; Hilgenboecker et al. 2008) and cause mitochondrial sweeps by hitchhiking when *Wolbachia* spread through host populations (Turelli et al. 1992; Shoemaker et al. 2004). In addition, successive sweeps of *Wolbachia* can occur during adaptation of the bacteria to hosts (Weeks et al. 2007) or multiple invasions of different *Wolbachia* strains (Werren et al. 1995). Therefore, *Wolbachia* may also promote Compensation-Draft Feedback by genetic hitchhiking of mitochondria.

As mentioned above, directional selection of functionally coadapted mitochondrial alleles can help to explain the excess of between-species replacements in *Nasonia*. Besides being physically linked, all mitochondrial genes are biologically associated. The mitochondrial genes encode for polypeptides that are components of protein complexes responsible for the oxidative phosphorylation. These complexes are composed of several polypeptides that are encoded by both mitochondrial and nuclear genes. The accumulation of mild-deleterious mutations in these proteins is likely to negatively affect fitness. Compensatory mutations within the same or interacting polypeptides would therefore be positively selected and rapidly reach fixation. Consistent with the view of coevolution between nuclear and mitochondrial genes, hybrid breakdown in *Nasonia* involves a marked nuclear–cytoplasmic interaction (Breeuwer and Werren 1995). The finding that each species of *Nasonia* has a very divergent mtDNA suggests the possibility of intergenomic coadaptation (discussed by Rand et al. 2004). The *Nasonia* hybrids show cytonuclear incompatibilities that could be due to defective interaction between the nuclear-encoded and the fast evolving mitochondrial-encoded proteins of the oxidative phosphorylation system. Therefore, we suggest that directional selection of functionally adapted mitochondrial alleles also explains the excess of between-species replacements. This hypothesis implies that the mitochondrial genome might have a direct role in the speciation process for these wasps.

In conclusion, an analysis of the mitochondrial genomes of *Nasonia* species reveal many interesting features, including unusual rearrangements, a high mutation rate, elevated nonsynonymous evolution, and evidence of directional selection. We propose that the high mitochondrial mutation rate in *Nasonia* provides substantial variation that may make it a good model for studying mitochondrial evolution and coevolution with nuclear genomes.

Supplementary Material

Tables S1 and S2 are available at *Molecular Biology Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Martha Raj, Brian Palmisano, and Daniel Reed for assistance with the laboratory work. We thank David Rand and anonymous referees for comments. This research was supported by grants NIH 5R01 GM070026 and Indiana's 21st Century Research and Technology Fund UND250086 to J.H.W.

Literature Cited

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Ballard JWO, Kreitman M. 1994. Unraveling selection in the mitochondrial genome of *Drosophila*. Genetics. 138:757–772.

Ballard JWO, Kreitman M. 1995. Is mitochondrial DNA a strictly neutral marker? Trends Ecol Evol. 10:485–488.

Bazin E, Glemin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science. 312:570–572.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27:1767–1780.

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Breeuwer JAJ, Werren JH. 1990. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature. 346:558–560.

Breeuwer JAJ, Werren JH. 1995. Hybrid breakdown between two haplodiploid species: The role of nuclear and cytoplasmic genes. Evolution. 49:705–717.

Burger G, Lavrov DV, Forget L, Lang BF. 2007. Sequencing complete mitochondrial and plastid genomes. Nat. Protoc. 2:603–614.

Campbell BC, Steffen-Campbell JD, Werren JH. 1993. Phylogeny of the *Nasonia* species complex (Hymenoptera: pteromalidae) inferred from an internal transcribed spacer (ITS2) and 28S rDNA sequences. Insect Mol Biol. 2:225–237.

Castro LR, Austin AD, Dowton M. 2002. Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera. Mol Biol Evol. 19:1100–1113.

Castro LR, Dowton M. 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: symphyta: pergidae). Mol Phylogenet Evol. 34:469–479.

Castro LR, Ruberu K, Dowton M. 2006. Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: vanhorniidae) and *Primeuchroeus spp*. (Aculeata: chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: hymenoptera). Genome. 49:752–766.

Chandra SBC, Vlk JL, Kapatral V. 2006. Comparative insect mitochondrial genomes: Differences despite conserved genome synteny. Afr. J Biotechnol. 5:1308–1318.

Crozier RH, Crozier YC. 1993. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. Genetics. 133:97–117.

Darling DC, Werren JH. 1990. Biosystematics of *Nasonia* (Hymenoptera, Pteromalidae)—2 new species reared from birds nests in North America. Ann Entomol Soc Am. 83:352–370.

DeSalle R, Freedman T, Prager EM, Wilson AC. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J Mol Evol. 26:157–164.

Dowton M, Austin AD. 1995. Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. J Mol Evol. 41:958–965.

Dowton M, Austin AD. 1997. Evidence for AT-transversion bias in wasp (Hymenoptera: symphyta) mitochondrial genes and its implications for the origin of parasitism. J Mol Evol. 44:398–405.

Dowton M, Austin AD. 1999. Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. Mol Biol Evol. 16:298–309.

Farris JS, Albert V, Kallersjo M, Lipscomb D, Kluge AG. 1996. Parsimony jackknifing outperforms neighbor-joining. Cladistics. 12:99–124.

Fay JC, Wyckoff GJ, Wu CI. 2002. Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. Nature. 415:1024–1026.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 39:783–791.

Felsenstein J. 1988. Phylogenies from molecular sequences: Inference and reliability. Annu Rev Genet. 22:521–565.

Flowers JM, Sezgin E, Kumagai S, Duvernell DD, Matzkin LM, Schmidt PS, Eanes WF. 2007. Adaptive evolution of metabolic pathways in *Drosophila*. Mol Biol Evol. 24:1347–1354.

Gerber AS, Loggins R, Kumar S, Dowling TE. 2001. Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? Annu Rev Genet. 35:539–566.

Gillespie JH. 2000. Genetic drift in an infinite population. The pseudohitchhiking model. Genetics. 155:909–919.

Gojobori J, Tang H, Akey JM, Wu CI. 2007. Adaptive evolution in humans revealed by the negative correlation between the polymorphism and fixation phases of evolution. Proc Natl Acad Sci. 104:3907–3912.

Hartl DL, Clark AG. 1997. Principles of Population Genetics. 3rd edition. Sunderland (MA): Sinauer Associates.

Hasegawa M, Cao Y, Yang Z. 1998. Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. Mol Biol Evol. 15:1499–1505.

Hey J, Kliman RM. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol Biol Evol. 10:804–822.

Hilgenboecker K, Hammerstain P, Schlattmann P, Teleschow A, Werren JH. 2008. How many species are infected with Wolbachia?—A statistical analysis of current data. FEMS Microbiol Lett. 281:215–220.

Jermiin LS, Crozier RH. 1994. The cytochrome b region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: Sequence divergence in Hymenoptera may be associated with nucleotide content. J Mol Evol. 38:282–294.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. New York: Academic Press. p. 21–132.

Kamping A, Katju V, Beukeboom LW, Werren JH. 2007. Inheritance of gynandromorphism in the parasitic wasp *Nasonia vitripennis*. Genetics. 175:1321–1333.

Laird CD, McConaughy BL, McCarthy BJ. 1969. Rate of fixation of nucleotide substitutions in evolution. Nature. 224:149–154.

Lewis DL, Farr CL, Farquhar AL, Kaguni LS. 1994. Sequence, organization, and evolution of the A+T region of *Drosophila melanogaster* mitochondrial DNA. Mol Biol Evol. 11:523–538.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl. Acids Res. 25:955–964.

Maddison DR, Maddison WP. 2000. MacClade: Analysis of phylogeny and character evolution. Version 4.0. Sunderland (MA): Sinauer Associates.

Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci. 90:4087–4091.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature. 351:652–654.

Meiklejohn CD, Montooth KL, Rand DM. 2007. Positive and negative selection on the mitochondrial genome. Trends Genet. 23:259–263.

Nachman MW. 1998. Deleterious mutations in animal mitochondrial DNA. Genetica. 102/103:61–69.

Nachman MW, Boyer SN, Aquadro CF. 1994. Nonneutral evolution at the mitochondrial NADH dehydrogenase subunit 3 gene in mice. Proc Natl Acad Sci. 91:6364–6368.

Peck JR. 1994. A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. Genetics. 137:597–606.

Rand DM, Dorfsman M, Kann LM. 1994. Neutral and nonneutral evolution of *Drosophila* mitochondrial DNA. Genetics. 138:741–756.

Rand DM, Haney RA, Fry AJ. 2004. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol Evol. 19: 645–653.

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022

Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. Mol Biol Evol. 13:735–748.

Rand DM, Kann LM. 1998. Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. Genetica. 102/103:393–407.

Russo CA, Takezaki N, Nei M. 1995. Molecular phylogeny and divergence times of Drosophilid species. Mol Biol Evol. 12:391–404.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19:2496–2497.

Shapiro JA, Huang W, Zhang C, et al. (12 co-authors). 2007. Adaptive genic evolution in the *Drosophila* genomes. Proc Natl Acad Sci. 104:2271–2276.

Shoemaker DD, Dyer KA, Ahrens M, McAbee K, Jaenike J. 2004. Decreased diversity but increased substitution rate in host mtDNA as a consequence of *Wolbachia* endosymbiont infection. Genetics. 168:2049–2058.

Swofford DL. 2002. PAUP*: Phylogenetic analysis using parsimony (and other methods). Version 4.0 beta. Washington (DC): Smithsonian Institution.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Templeton AR. 1996. Contingency tests of neutrality using intra/interspecific gene trees: the rejection of neutrality for the evolution of the mitochondrial cytochrome oxidase II gene in the hominoid primates. Genetics. 144:1263–1270.

Turelli M, Hoffmann AA, McKechnie SW. 1992. Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics. 132: 713–723.

van Opijnen T, Baudry E, Baldo L, Bartos J, Werren JH. 2005. Genetic variability in the three genomes of *Nasonia*: Nuclear, mitochondrial, and *Wolbachia*. Insect Mol Biol. 14:653–663.

Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA. 2007. From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. PLoS Biol. 5:e114.

Werren JH, Windsor DM. 2000. *Wolbachia* infection frequencies in insects: Evidence of a global equilibrium? Proc Biol Sci. 267:1277–1285.

Werren JH, Zhang W, Guo LR. 1995. Evolution and phylogeny of *Wolbachia*: Reproductive parasites of arthropod. Proc Biol Sci. 261:55–63.

Xu W, Jameson D, Tang B, Higgs PG. 2006. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. J Mol Evol. 63:375–392.

Yang Z. 2007. PAML: A program package for phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.

Yang Z, Nielsen R. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol. 17:32–43.

Naruya Saitou, Associate Editor

Accepted July 19, 2008

Downloaded from https://academic.oup.com/mbe/article/25/10/2167/1028716 by guest on 20 August 2022