Heredity (2010) 104, 318–326

www.nature.com/hdy

ORIGINAL ARTICLE

Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial–*Wolbachia* sweep in North America

R Raychoudhury[1,4], BK Grillenberger[2,4], J Gadau[3], R Bijlsma[2], L van de Zande[2], JH Werren[1,5] and LW Beukeboom[2,5]

[1]*Department of Biology, University of Rochester; Rochester, NY, USA;* [2]*Evolutionary Genetics, Centre for Ecological and Evolutionary Studies, University of Groningen, Haren, The Netherlands and* [3]*School of Life Sciences, Arizona State University, Tempe, AZ, USA*

Here we report evidence of a mitochondrial–*Wolbachia* sweep in North American populations of the parasitoid wasp *Nasonia vitripennis*, a cosmopolitan species and emerging model organism for evolutionary and genetic studies. Analysis of the genetic variation of 89 *N. vitripennis* specimens from Europe and North America was performed using four types of genetic markers: a portion of the *mitochondrial cytochrome oxidase I* gene, nine polymorphic nuclear microsatellites, sequences from 11 single-copy nuclear markers and six *Wolbachia* genes. The results show that the European populations have a sevenfold higher mitochondrial sequence variation than North American populations, but similar levels of microsatellite and nuclear gene sequence variation. Variation in the North American mitochondria is extremely low ($\pi = 0.31\%$), despite a highly elevated mutation rate ($\sim$35–40 times higher than the nuclear genes) in the mitochondria of *Nasonia*. The data are indicative of a mitochondrial sweep in the North American population, possibly due to *Wolbachia* infections that are maternally co-inherited with the mitochondria. Owing to similar levels of nuclear variation, the data could not resolve whether *N. vitripennis* originated in the New or the Old World.
Heredity (2010) **104,** 318–326; doi:10.1038/hdy.2009.160; published online 20 January 2010

Keywords: biogeography; bottleneck; microsatellites; mitochondrial DNA; selective sweep; *Wolbachia*

Introduction

The parasitic wasp *Nasonia vitripennis* is a model organism for studying developmental biology, sex-ratio studies, evolutionary genetics and host–microbe interaction (Werren, 1980; Breeuwer and Werren, 1990, 1995; Bordenstein *et al*., 2001; Beukeboom and Desplan, 2003; Pultz and Leaf, 2003; Shuker *et al*., 2003; Werren and Loehlin, 2009). Its genome has recently been sequenced, providing a wealth of information for population genetic studies (Werren *et al*., 2009). It has a cosmopolitan distribution and has been found almost everywhere it has been looked for (Whiting, 1967; see also http://www.nhm.ac.uk/research-curation/projects/chalcidoids/). Three additional species have been described in the genus: *N. giraulti* found in eastern North America (NA); *N. longicornis* found in western North America (Darling and Werren, 1990); and the recently identified *N. oneida* from New York state (Raychoudhury *et al*., 2010). Thus, there are four species found in North America, *N. vitripennis*, *N. giraulti*, *N. longicornis* and *N. oneida*, and *N. vitripennis* occurs microsympatrically with the other three. *N. giraulti* and *N. longicornis* are largely allopatric, whereas *N. giraulti* and *N. oneida* have overlapping distributions. In Europe (EU), however, only one species is currently found, *N. vitripennis*. An unresolved question is whether *N. vitripennis* originated in North America along with other known representatives of the genus, or from Eurasia. The divergence of the *Nasonia* genus has been placed in the Pleistocene era around 10^6 years ago (Campbell *et al*., 1993), with *N. vitripennis* diverging first and the other three around 0.4×10^6 years ago (Raychoudhury *et al*., 2010). All the species harbour the maternally inherited, reproductive endosymbiont *Wolbachia,* which causes sperm–egg incompatibilities and has a large role in reproductive incompatibility between several of the species, at least under laboratory conditions (Breeuwer and Werren, 1990; Bordenstein *et al*., 2001). However, elimination of *Wolbachia* by antibiotic treatment leads to the production of viable hybrid offspring (Breeuwer and Werren, 1990). Breeuwer and Werren (1995), Bordenstein *et al*. (2001) and Raychoudhury *et al*. (2010) have shown, using antibiotically cured strains from each species, that *N. vitripennis* has diverged from the other three species genetically to the point where some F2-hybrid breakdown occurs, whereas the three North American species show little or no hybrid breakdown (Bordenstein *et al*., 2001; Raychoudhury *et al*., 2010).

A comparison of the genetic diversity in the three *Nasonia* species in nuclear, mitochondrial and *Wolbachia*

Correspondence: R Raychoudhury, Department of Biology, University of Rochester, 60 Crittenden Boulevard, Apartment 602, Rochester, NY, 14267, USA.
E-mail: rraychou@mail.rochester.edu
[4]*These authors contributed equally to this work.*
[5]*These authors contributed equally to this work.*
Received 3 August 2009; revised 19 October 2009; accepted 22 October 2009; published online 20 January 2010

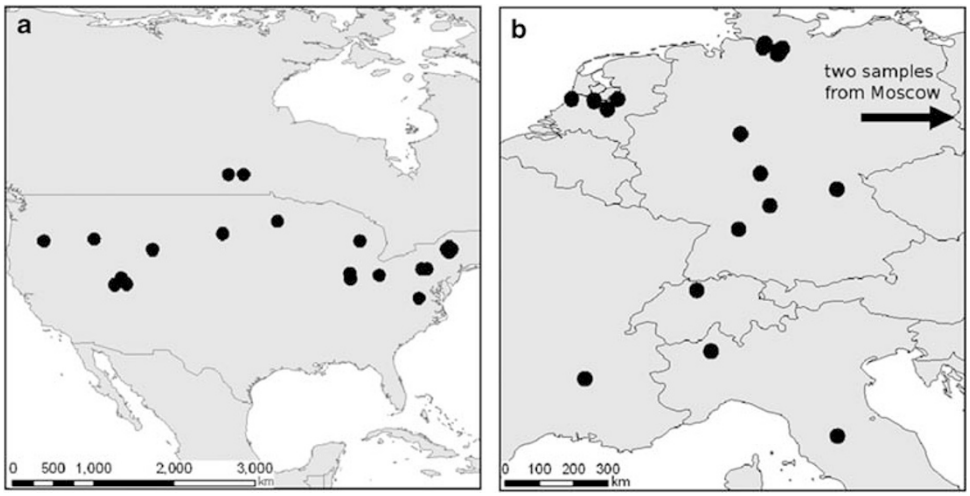

Figure 1 Location of the sampling sites in (**a**) North America and (**b**) Europe.

sequences by van Opijnen *et al*. (2005), revealed a remarkably reduced mitochondrial variation in *N. vitripennis* compared with the other species, whereas the nuclear and *Wolbachia* genomes did not show such a pattern. Oliveira *et al*. (2008) confirmed this pattern in the mitochondria with more strains from both Europe and North America. Moreover, the *N. vitripennis* mitochondria were found to be divided into two distinct clades, one on each side of the Atlantic, with the EU strains showing distinctly more variation than their NA counterparts. In addition, some of the NA haplotypes were nested within the EU mitochondrial haplotype variation. To explain this phenomenon of reduced mitochondrial variation in North America, van Opijnen *et al*. (2005) hypothesized a relatively recent selective sweep of a well-adapted mitochondrial haplotype. The low sequence variation in NA populations is especially noteworthy because Oliveira *et al*. (2008) have shown the *Nasonia* species have an exceptionally high mitochondrial mutation rate relative to the nuclear genes. As mitochondria and *Wolbachia* are co-inherited, a *Wolbachia*-induced sweep can produce very high frequency of the associated mitochondrial haplotypes (Turelli and Hoffmann, 1991; Werren, 1997; Ballard and Whitlock, 2004; Hurst and Jiggins, 2005) and such a sweep could also have produced a reduction in mitochondrial diversity in NA populations. However, an alternative explanation is that NA *N. vitripennis* have undergone a genetic bottleneck, possibly during a founding event from European populations.

In this study, we have collected and analysed additional data to resolve among these possibilities. By comparing genetic diversity indices among a larger sample of NA ($N=52$) and EU ($N=27$) *Nasonia vitripennis* in mitochondria, microsatellites, single-copy nuclear gene sequences and *Wolbachia*, we tested the two alternative hypotheses of mitochondrial–*Wolbachia* sweep versus a bottleneck event in North America. One of the key predictions of a mitochondrial sweep in NA populations would be the presence of significantly less mitochondrial variation compared with EU populations. Another prediction would be relatively similar diversities in microsatellites and nuclear sequences in the two populations. A mitochondrial sweep would asymmetrically affect variation in the cytoplasmic genome compared with the nuclear genome. However, in the event of a bottleneck in NA populations both nuclear and mitochondrial genomes would be expected to show a reduction in variation compared with the EU population. We provide analytical evidence for and against these two competing hypotheses and find that the data clearly support a cytoplasmic sweep in North America.

Materials and methods

Nasonia sampling and DNA extraction

We acquired 52 specimens from North America and 27 from Europe (Figure 1) from bird nests and artificial baits (see supplementary online Table 1 for a complete sample list). The NA specimens were obtained from previous samplings by members of the Werren laboratory (University of Rochester) and from field collections by the senior author in 2005. The EU specimens were obtained from wild derived stocks in the laboratory in Groningen, and from field collections in summer of 2006. All field materials were stored at -80°C or in 95% ethanol. DNA isolation followed a high-salt isolation protocol (Maniatis *et al*., 1982). As some of the laboratory lines had been in culture for a long time and may have lost some genetic variation, we used only one individual per line. This precluded any population genetic estimates that are based on the measured heterozygosity in the sampled population (for example, F_{IS}). For the field samples, we again used only one individual per sampling point (mostly a bird nest box) to avoid sampling of highly related individuals. For the sequencing of nuclear markers, 18 different strains of *N. vitripennis* were chosen. These comprised of 10 strains from North America and 8 from Europe (summarized in Supplementary Table 2).

Mitochondrial DNA analysis

We designed primers for the *N. vitripennis cytochrome oxidase I* (*COI*) gene sequence obtained from GENBANK (LOC100113910). The combination of the primers: NL COI Fwd (5′-GTTATACCTKTWATAATWGGAGGA TTTGG-3′) and NV COI Rev (5′-CTTTGAAAACCACG TTACCC-3′) amplified a ~400-bp fragment in *Nasonia*

using a standard PCR protocol at 52 °C annealing temperature.

The amplification of the fragment was checked on agarose gels (1%) and the PCR product was purified with the Nucleospin II kit (Machery Nagel, Düren, Germany). The sequencing reaction was performed on both strands using the same primers and the Big Dye Termination sequencing kit (Applied Biosystems, Foster City, CA, USA). Purified sequencing products were analysed using an ABI 377 automatic sequencer (Applied Biosystems). For the phylogenetic analysis, we used the ALIGNX program included in the VECTOR NTI software package (Invitrogen, Carlsbad, CA, USA) to align a 399-bp fragment without gaps. The software DNASP4 (Rozas *et al*., 2003) was used to calculate haplotype diversity and nucleotide diversity indices. To find the ancestral haplotype, we used a median joining haplotype network (Bandelt *et al*., 1999) constructed with NETWORK 4.5 (available at: http://www.fluxus-technology.com/sharenet.htm). The PHYLIP software package (available at: http://evolution.gs.washington.edu/phylip.html) was used to construct a neighbour-joining tree rooted with an out group formed by two *Trichogramma* species obtained from GENBANK (*T. ostriniae*: DQ177914, and *T. achaeae*: DQ177918). Two *N. longicornis* (GENBANK accession # EU935415 and EU935416) and two *N. giraulti* (GENBANK accession # EU935417 and EU935418) sequences obtained from this study were also added for comparison. All *N. vitripennis COI* sequences are deposited in GENBANK (accession # EU935326–EU935414; see Supplementary Table 1 for sample names and corresponding accession numbers).

Simulations were performed to estimate the time since the last mitochondrial sweep in NA populations using a Bayesian summary likelihood method (Garrigan *et al*., in press). The method is based on forward-time simulations of a Wright–Fisher model, starting from monomorphism. It estimates the time since the last complete mitochondrial sweep (T_s) for a population that has not yet achieved an equilibrium distribution of mutations. The approach uses aspects of the site frequency spectrum, particularly the proportion of single polymorphisms, to estimate the relaxation time (in generations) from monomorphism to drift–mutation equilibrium. On average, this relaxation time was found to be $4N$ generations, and time to equilibrium is dependent upon the population mutation rate, $\theta = N\mu$, where N is the population effective number of chromosomes and μ is the per locus mutation rate per generation. The method accounts for uncertainty in the population size and mutation rate by randomly sampling values from a gamma prior distribution, and it also assumes a constant effective population size. The Wright–Fisher model simulation stops when both the number of segregating sites and number of singleton polymorphisms in the simulation match those observed from the data, and the generation is recorded and incorporated into the posterior probability distribution of T_s. The posterior distribution for the time since monomorphism can then be estimated from 1000 iterations of the model.

Two different approaches were used to estimate μ, the *CO1* gene mutation rate. First, we took the *Drosphila melanogaster* empirically derived per base mutation rate for mitochondria of 6.2×10^{-8} (Haag-Liautard *et al*., 2008) and extrapolated it to *Nasonia*. However, *Nasonia* is known to have a much higher mitochondrial mutation rate. Oliveira *et al*. (2008) used synonymous divergence rates between *Nasonia* species in mitochondria and nuclear genes to estimate the mitochondrial mutation rate and observed it to be 35-fold higher than the nucleus. Extrapolating to predictions of the time of *D. simulans–D. melanogaster* split and *N. giraulti–N. longicornis* split, indicated that the mitochondrial mutation rate is 13 × greater in *Nasonia* than in *Drosophila*. As an alternative approach, the ratio of mitochondrial/nuclear mutation rates in *D. melanogaster* is 10 × (Haag-Liautard *et al*., 2008), compared with 35 × for *Nasonia*. Assuming that nuclear mutation rates are similar in the *Drosophila* and *Nasonia* lineages, this gives a 3.5 × higher mutation rate for *Nasonia* mitochondria. Adjusting the number of synonymous (neutral) sites in *CO1* (92) gives an estimated μ of 7.4×10^{-5} and 2.0×10^{-5}, respectively.

The EU populations have an empirically estimated $\theta = 8$ and much higher mitochondrial variation than the NA populations. Under the assumption that EU populations are at equilibrium (and have similar population sizes), we estimated the effective number of chromosomes (N) corresponding to the mutation rates to be 107 886 and 400 721, respectively. These parameters were then used in the simulation to estimate the time since monomorphism in NA populations, along with the empirically derived number of segregating sites (8), singleton mutations (4) and sample size (57).

Microsatellite analysis

A total of nine polymorphic microsatellites were examined (di- and trinucleotide repeats with variable repeat length; summarized in Supplementary online Table 3), spread over all five chromosomes to estimate the genetic diversity of the nuclear genome of the EU and the NA populations. The microsatellites Nv 22 and Nv 23 have originally been developed by Pietsch *et al*. (2004), but the primers were redesigned in our laboratory. All other loci were developed in the Groningen laboratory (Beukeboom and Desplan, 2003). As the samples were not distributed evenly over the areas, we grouped samples of neighbouring origin into arbitrary subpopulations to be able to calculate reliable genetic distance estimates. We grouped the samples from the following locations into corresponding subpopulations: Germany = GER; the Netherlands = NL; Switzerland, Italy and France = CH-IT-FR; Canada = CN; Idaho = ID; New York = NY; Utah = UT; Montana, Oregon, Wyoming, Nevada and South Dakota = MT-OR-WY-NV-SD; Pennsylvania and Virginia = PE-VI; Indiana, Michigan, Ohio = IN-MI-OH. The genetic diversity indices were calculated using the F-STAT software package (Goudet, 2001) or manually. Nei's standard genetic distance D_S (Nei, 1987) and Goldstein's $\delta\mu^2$ value (Goldstein and Pollock, 1994) were used to construct neighbour-joining trees between subpopulations within Europe and North America, using the POPULATIONS 1.2.30 software (available at: http://bioinformatics.org/~tryphon/populations/) and 1000 bootstraps over loci.

Nuclear sequence analysis

A total of 11 different nuclear markers were sequenced across 17 different *N. vitripennis* strains. One EU isolate, AsymCX, is the source strain for the genome sequence (Werren *et al*., 2009) and relevant sequences were

downloaded from NCBI. Primers for five of these markers were designed previously by Raychoudhury *et al*. (2009) and the rest were designed for this study. The primers were chosen specifically to provide coverage across the genome and markers across all five chromosomes have been incorporated in this study (summarized in Supplementary Table 4). The condition for all sets of primers were 94 °C for 2 min; 38 cycles at 94 °C for 45 s, 55 °C for 60 s, 72 °C for 1 min 30 s; 72 °C for 10 min; then held at 4 °C. Regular Taq (Invitrogen) polymerase was used for PCR reactions and the amplified products were treated with shrimp alkaline phosphatase and exonuclease I (Amersham, CA, USA) to clean the reactions before sequencing. Sequencing was performed at the Functional Genomics Centre, University of Rochester, using BigDye v2.0 or v3.0 terminator sequencing kit (Applied Biosystem) and an ABI 3700 or 3730xl automated sequencer (Applied Biosystem). The chromatograms were manually inspected and cleaned with Sequencher 4.7 software (Gene Codes, MI, USA) and the sequences were aligned using Bioedit vs 7.0.1 (Hall, 1999). The sequences have been deposited in gene bank with accession numbers GQ896384–GQ896485.

Wolbachia analysis

To identify the *Wolbachia* infection, we used the multi locus sequence typing (MLST) system developed specifically for this bacteria (Baldo *et al*., 2006b). Essentially, information from five genes (*gatB*, *coxA*, *hcpA*, *ftsZ* and *fbpA*) is used to characterize the *Wolbachia* infections by comparison with other sequences on the PubMLST website (http://pubmlst.org/wolbachia/), which archives the sequences. The combination of the five haplotypes forms a unique allelic profile for each *Wolbachia* infection. In addition, we also amplified a part of the *wsp* gene. Primers have been developed for both super group A (*w*NvitA)- and B (*w*NvitB)-specific *Wolbachia* for the five MLST genes and the sequences and conditions can be found on the website. As mitochondria and *Wolbachia* are maternally co-inherited, we screened 21 different field isolates, selected to cover the range of variation in mitochondrial DNA haplotypes, for the presence and characterization of *Wolbachia* with primers from both super groups for the MLST genes as well as *wsp*. The primers and their conditions for the A and B super group *wsp* gene were taken from Zhou *et al*. (1998).

Results

Mitochondrial DNA sequence data

We found a sevenfold higher mitochondrial diversity in EU compared with NA samples (π for EU: 2.36% and for NA: 0.31%; Supplementary Table 5). The level of variation in the NA *N. vitripennis* samples is in the same range as found by van Opijnen *et al*. (2005). It is striking that the diversity in the EU *N. vitripennis* samples found in this study is in the same range as that of *N. longicornis* and *N. giraulti* found by van Opijnen *et al*. (2005) in NA. Tajima's *D* (Tajima, 1989) is −0.88 but not significantly different from 0 for the NA samples and 0.53 (not significant) for the EU samples (Supplementary Table 5). The NA samples differ in their nucleotide sequence (using Nei's D_{XY}) from the EU samples by, on average, 2.3% (compared with 0.3 and 2.4% within NA and EU, respectively), whereas *N. vitripennis* samples differ from the two sister species by, on average, 14.2% (*N. vitripennis*−*N. longicornis* = 13.6%, *N. vitripennis*−*N. giraulti* = 14.8%; Supplementary Table 6). However, there was considerably less *COI* variation in NA populations relative to EU populations, consistent with either the NA founder hypothesis or the NA mitochondrial–*Wolbachia* sweep hypothesis. The low variation in NA population is particularly noteworthy given the 35–40-fold higher synonymous substitution rate in *Nasonia* mitochondrial genes relative to nuclear genes (Oliveira *et al*., 2008; Raychoudhury *et al*., 2009).

Table 1 Mitochondrial haplotype frequencies in NA, EU and pooled (total) samples

Haplotype	*# NA*	*# EU*	*# total*	*freq NA*	*freq total*	*freq EU*
1	0	3	3	0	0.0337	0.0938
2	43	4	47	0.7544	0.5281	0.125
3	0	1	1	0	0.0112	0.0313
4	0	1	1	0	0.0112	0.0313
5	6	0	6	0.1053	0.0674	0
6	0	1	1	0	0.0112	0.0313
7	0	4	4	0	0.0449	0.125
8	0	6	6	0	0.0674	0.1875
9	0	4	4	0	0.0449	0.125
10	0	1	1	0	0.0112	0.0312
11	2	0	2	0.0351	0.0225	0
12	1	0	1	0.0175	0.0112	0
13	1	0	1	0.0175	0.0112	0
14	1	0	1	0.0175	0.0112	0
15	0	1	1	0	0.0112	0.0313
16	0	1	1	0	0.0112	0.0313
17	0	1	1	0	0.0112	0.0313
18	0	1	1	0	0.0112	0.0313
19	1	0	1	0.0175	0.0112	0
20*	1	0	1	0.0175	0.0112	0
21*	1	0	1	0.0175	0.0112	0
22	0	1	1	0	0.0112	0.0313
23*	0	1	1	0	0.0112	0.0313
24	0	1	1	0	0.0112	0.0313
Sample size	57	32	89			

Abbreviations: EU, European; NA, North American.
*Haplotypes containing one non-synonymous mutation.

In total, there were 24 unique mitochondrial haplotypes containing 34 synonymous and three non-synonymous mutations in the sequenced isolates (Table 1). The non-synonymous mutations lead to three haplotypes that were represented by one individual each. The only haplotype that was shared between EU and NA populations was the most common haplotype in North America, but it was not the most frequent in Europe (haplotype # 2; Figure 2 and Table 1). A nested clade analysis (Templeton *et al*., 1992) did not yield any significant geographic clustering of haplotypes within continents. The phylogenetic analysis (Supplementary online Figure 1) showed that all NA samples cluster together with a few interspersed EU samples. The whole NA group is nested within the EU group, and structure among the EU samples is much more pronounced compared with the NA samples.

The estimated divergence time of *N. longicornis*–*N. giraulti* split was used to calibrate divergence times of the two populations of *N. vitripennis*. Two estimates of the *N. longicornis*–*N. giraulti* split are available and are 0.4×10^6 years ago as estimated by Campbell *et al*. (1993) and 0.51×10^6 years ago as estimated by Raychoudhury *et al*. (2009). The total synonymous divergence (k_s) between

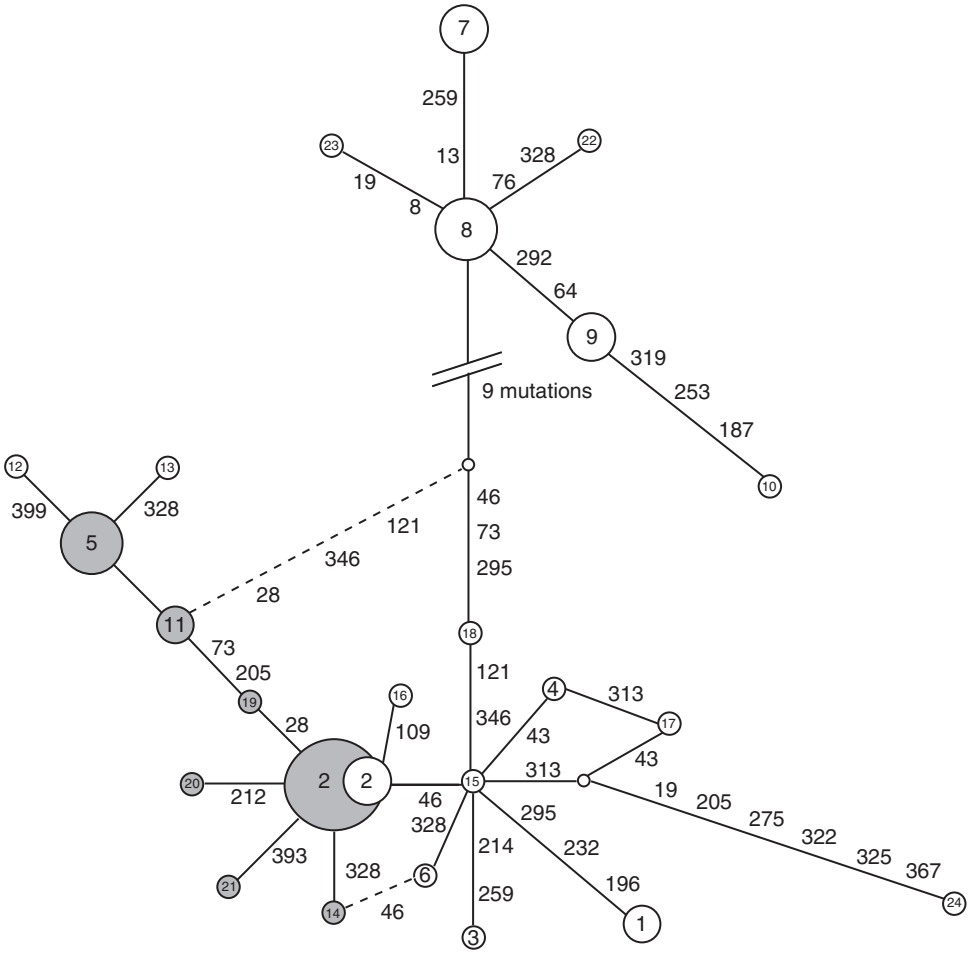

Figure 2 Haplotype network based on the 399-bp mitochondrial *COI* sequence. Numbers in circles represent particular haplotypes, numbers along lines are the nucleotide positions in the sequence that changed. The sizes of the circles reflect the haplotype frequency. Nodes without a label represent intermediate haplotypes that were not sampled. Dashed lines indicate alternative mutation routes. The grey haplotypes have been found in North America, the white ones in Europe. Haplotye # 2 has been found on both continents (see Table 1 and text).

these two species has been estimated to be 38.9% (Oliveira *et al*., 2008), whereas the k_s for the NA and EU populations are 1.32 and 10.55%, respectively. Thus, extrapolating this rate to *N. vitripennis* gives an age of the NA mitochondria of around $0.027–0.035 \times 10^6$ years and the EU mitochondria to be $0.217–0.277 \times 10^6$ years.

A Bayesian analysis method based on mutational spectral patterns (Garrigan *et al*., in press) was used to estimate the time since the mitochondrial sweep in NA populations (see Materials and Methods: Mitochondrial DNA analysis). Using EU populations as a reference and assuming them to be in mutational equilibrium, NA populations were estimated to have undergone a mitochondrial sweep 1945 generations ago (95% confidence limits of 1219–2671) and 3193 (3127–3259) generations ago, based on two estimates of the mitochondrial mutation rate in *Nasonia*. Given the potential error in estimating the parameters in these simulations, these times should be taken with caution. Nevertheless, the estimates support the contention that NA populations have undergone a selective sweep relatively recently and after the NA–EU split, thus accounting for their reduced variation relative to both the EU population and nuclear variation (see below).

Microsatellite data

Nei's gene diversity $\hat{H}$ (Nei, 1987) and allelic richness R are equal in the NA and EU samples, indicating similar levels of nucleotide variation in the two areas. This pattern is consistent over loci with various repeat lengths and motifs (Table 2). This is indicative of an evolutionary force affecting mitochondrial and nuclear genes differentially, as there is a sevenfold difference in the mitochondrial variation between NA and EU populations but not in the nuclear markers. There is low, but highly significant ($P < 0.001$) differentiation between the two continents ($F_{ST} = 0.037 \pm 0.013$ s.e.). As this measure is dependent on the variability of the marker used, we also calculated Hedrick's $G'_{ST} = 0.23$ (Hedrick, 2005), to compare the differentiation to other studies. This corrected differentiation index shows the same level of differentiation between the two continents as did a previous study between two populations within Europe (0.23, Grillenberger *et al*., 2008). Both neighbour-joining trees, based on the genetic distances (D_S and $\delta\mu^2$) between sub-populations within Europe and North America, indicate no clear split between the two continents. We only present the tree based on Nei's D_S, as the qualitative results for the tree based on $\delta\mu^2$ are

Table 2 Summary statistics of microsatellite data

Locus	*# repeats*	$\hat{H}$ *NA*	$\hat{H}$ *EU*	*# NA*	*# EU*	*# total*	R_S *NA*	R_S *EU*	R_T *total*	F_{ST}
Nv-22	26	0.88	0.85	12	11	14	10.47	9.77	10.76	0.04
Nv-23	28	0.79	0.57	14	8	17	12.35	7.63	11.4	0.07
Nv-41	15	0.74	0.89	12	11	15	11.13	12.73	11.94	0.14
Nv-46	21	0.78	0.86	7	12	12	6.89	11.46	8.79	0.07
N1-10	14	0.4	0.4	5	3	6	5.06	5	5.09	0.08
Nv-300[a]	6	0.51	0.47	4	4	6	3.36	3.68	3.56	0.58
Nv-303	7	0.76	0.73	6	6	6	5.73	5.9	5.73	0.09
Nv-313	14	0.89	0.81	13	8	16	11.96	8.9	11.41	0.16
Nv-316	16	0.76	0.75	10	6	11	8.61	6.77	8.39	0.14
Mean		0.72 ± 0.16	0.70 ± 0.18	9.22 ± 3.55	7.67 ± 3.02	11.44 ± 4.48	8.39 ± 3.28	7.98 ± 3.00	8.56 ± 3.12	0.15 ± 0.054

Abbreviations: EU, Europe; NA, North America; R, richness.
Number of repeats are based on sequenced genome, gene diversity ($\hat{H}$), allele number (#), allelic richness per locus and population, as well as the total sample, and Weir and Cockerham's F_{ST} (Weir and Cockerham, 1984) of the microsatellite data. Means over all loci are given with ± s.d. values. Mean F_{ST} is given with ± s.e value after Jackkning over loci.
[a]Nv 300 is a trinucleotide marker, whereas all other markers are dinucleotide repeats.

Table 3 Average nucleotide diversity of North American and European strains based on eleven markers sequenced totaling 7079 base pairs.

	NA strains		*EU strains*	
	π *(silent)*	π *(non-synonymous)*	π *(silent)*	π *(non-synonymous)*
Average	0.0016	0.0005	0.0016	0.0010

Abbreviations: EU, Europe; NA, North America.
Silent sites represent synonymous as well as non-coding sites. All values are corrected with Jukes and Cantor formula (Jukes and Cantor, 1969).

identical (Supplementary online Figure 2). The mean genetic distance estimated with both parameters indicates no clear difference between the continents. The variation within EU population is slightly higher than that within NA population (D_S ± s.e.: EU: 0.36 ± 0.10; NA: 0.35 ± 0.02; between EU and NA: 0.43 ± 0.04; $\delta\mu^2$ ± s.e.: EU: 22.05 ± 8.06; NA: 13.92 ± 2.30; between EU and NA: 17.18 ± 2.54).

Nuclear sequence variation

As no distinct differences in microsatellites could be detected between the two populations, a survey of nuclear sequence diversity was initiated. The expectation was that the greater resolving power of these sequences would shed light on the level of nuclear diversity of the strains from these two areas. Furthermore, because of the lower mutation rate of single-copy nuclear DNA relative to microsatellites, these single-copy sequences are likely to clarify whether NA populations underwent a nuclear genetic bottleneck relative to EU populations.

The 11 different nuclear markers together yielded a data set of 7079 base pairs for each of the 18 different strains. Two measures of nucleotide diversity were enumerated: diversity in silent sites (synonymous and non-coding) and in non-synonymous sites (summarized in Table 3). Unlike the mitochondrial data, there was no greater variation in the EU strains and in fact the levels of silent site diversity were the same in both (0.16%), and thus, not significantly different (Mann–Whitney *U*-test: $U = 61$, DF = 11, $P = 0.5$). Diversity in non-synonymous sites for NA and EU strains were 0.05 and 0.1%, respectively, and again there was no significant difference in level of variation between the two (Mann–Whitney *U*-test: $U = 44$, DF = 11, $P = 0.86$). Although there are unique mutations within the two populations, they are mainly strain-specific and there are no fixed polymorphisms separating the two populations. The EU population showed six unique single nucleotide polymorphisms (SNPs), whereas the NA population had three. All the other SNPs were shared polymorphisms between the two populations. Thus, there is no significant difference between the levels of nuclear sequence diversity between the two regions. These shared polymorphisms could be due to retention of ancestral polymorphisms or due to gene flow between NA and EU populations, and we currently cannot distinguish between these alternatives. This is also true for the shared polymorphisms in microsatellites.

Wolbachia variation

A key feature of a *Wolbachia*-induced sweep is the hitchhiking of the associated cytotype with the invading infection (Turelli and Hoffmann, 1991; Werren, 1997; Ballard and Whitlock, 2004; Hurst and Jiggins, 2005). As the *N. vitripennis* strains clearly differ in the levels of mitochondrial diversity, an obvious hypothesis was that the causal force was a different *Wolbachia* infection sweeping through the NA population. Previously, Raychoudhury *et al*. (2009) showed, with a smaller data set of *N. vitripennis* samples, that there is no *Wolbachia* variation in the five housekeeping MLST genes in NA and EU populations. Using the larger data set in this study, we screened additional *Wolbachia* from representative samples from each of the major mitochondrial haplotypes. A total of five different genes from the MLST scheme and an additional marker, *wsp*, were sequenced from each of the 22 different strains for both the two *Wolbachia* strains found in *N. vitripennis* (*w*NvitA and *w*NvitB). We found three to be uninfected, three to be infected with type *w*NvitA only, and 15 to be doubly infected with *w*NvitA and *w*NvitB (Supplementary Table 1). One sample, HV-736, could not be sequenced completely because of poor DNA quality. In all the sequenced individuals, we did not find any variation within *Wolbachia* strains across NA and EU

populations. It should be emphasized that failure to detect sequence variation in these six genes does not rule it out in other portions of the genome, particularly in hypervariable and rapidly evolving regions or regions under directional selection (Riegler *et al*., 2005). A previous study shows that *Wolbachia* have an estimated mutation rate 120-fold lower than mitochondria in *Nasonia* (Raychoudhury *et al*., 2009), and therefore the absence of variation in *Wolbachia* compared with mitochondria is not surprising.

Previous studies have shown that nearly 100% of *N. vitripennis* collected from nature are double-infected with both A and B super group *Wolbachia* (van Opijnen *et al*., 2005; Raychoudhury *et al*., 2009). All the single-infected and uninfected strains in this study are laboratory lines that have been in culture for at least a few years (Supplementary Table 1, indicated with *), and it has been shown that laboratory culturing can result in loss of *Wolbachia* infections (Perrot-Minnot *et al*., 1996), particularly during long-term refrigeration storage of diapausing larvae, routinely used for maintenance of strains. Thus the most parsimonious explanation of the infection status of these strains is the secondary loss of infections in the laboratory rather than actual infection status in the field.

Comparing mitochondrial, nuclear and *Wolbachia* variation

On the basis of nuclear sequence and microsatellite variation between NA and EU populations, it can be concluded that the two populations on either side of the Atlantic did not experience a reduction in the diversity of the nuclear genes relative to each other. Thus, these data are inconsistent with a founder effect in NA *N. vitripennis.* The other phenomenon that can explain the observed reduction in mitochondrial diversity in NA population is a mitochondrial sweep. This would show a reduction in the mitochondrial diversity of the NA strains but the nuclear variation would remain similar to that of the EU strains. We compared the number of polymorphic positions and number of mutations across the mitochondria and the nuclear markers from the strains of the two populations. As seen in Table 4, there are significantly lower numbers of both polymorphisms and polymorphic sites in NA mitochondria relative to nuclear variation. The NA population shows a reduced number of polymorphisms in its mitochondria compared with EU population (8 versus 33, Fisher's Exact test, $P = 18.28 \times 10^{-4}$). This contrasts with the number of polymorphisms in the nuclear genes (28 in each). The second measure (Table 4b) enumerates the total polymorphic positions in the two populations. There is a significant reduction in the number of mitochondrial polymorphic positions in the NA samples compared with EU samples (7 versus 32 positions, Fisher's Exact test, $P = 6.51 \times 10^{-4}$). This is in contrast to very similar levels of polymorphic nuclear positions (30 and 28 in NA and EU population, respectively). The mitochondria of the EU strains show 8% sites to be polymorphic, whereas, in contrast, the NA populations show only 1.8% polymorphic positions. Although *Nasonia* mitochondria are known to have an approximately 35–40-fold higher substitution rate than the nucleus (Oliveira *et al*., 2008; Raychoudhury *et al*., 2009), there is actually lower variation in mitochondrial genes relative to nuclear genes in NA population. This clearly shows that the NA population has undergone a significant reduction in its mitochondrial diversity but not in its nuclear genes.

Table 4 The number of polymorphic sites and mutations in the NA and EU populations

	Mitochondria	*Nuclear*
(a)	*Within-population mutations*	
	(399 bp)	(7079 bp)
NA	8 (2%)	28 (0.4%)
EU	33 (8.2%)	28 (0.4%)
(b)	*Total Polymorphic Positions*	
NA	7 (1.8%)	30 (0.4%)
EU	32 (8%)	28 (0.4%)
NV total	35 (8.8%)	41 (0.6%)

Abbreviations: EU, European; NA, North American.

No variation was found in either the A or B super group *Wolbachia* for the five MLST gene set, which are housekeeping genes generally under stabilizing selection. This reflects the lower estimated mutation rate for *Wolbachia* relative to that for the nuclear genome of approximately 1/3 (Raychoudhury *et al*., 2009), and is consistent with a *Wolbachia* sweep in the NA population and insufficient time for *Wolbachia* synonymous variation to accumulate.

Discussion

We found that the genetic diversity of *N. vitripennis* at the mitochondrial level is much higher in Europe than in North America, even though the latter specimens were spread over a larger geographical area (nucleotide diversity π: NA = 0.31%, EU = 2.36%, see Figure 1). As a consequence, the phylogenetic analysis shows a much weaker structuring in North America (Figure 2). There is one predominant haplotype (# 2) in North America that is also present in Europe but is infrequent. The negative Tajima's *D* value suggests that there has been a rapid expansion of *N. vitripennis* in North America, whereas there is no evidence for such an event in Europe.

In the case of a founder effect during the colonization of North America, one would expect a drastic reduction in genetic diversity. However, the differences detected in the microsatellite variation are only subtle, and therefore do not indicate a recent bottleneck. Moreover, the microsatellite data do not reveal a split between the EU and NA populations (Supplementary online Figure 2). Therefore, microsatellite, nuclear SNP and mitochondrial sequence data all show shared polymorphisms between NA and EU populations. Presently, we cannot distinguish between the retention of ancestral polymorphisms and gene flow to explain this pattern. Although microsatellites show high mutation rates that can lead to convergence (homoplasy) over time (Nauta and Weissing, 1996), this is unlikely for SNPs and mitochondrial sequences given the observed divergence levels. Moreover, although levels of SNP variation are low, there are both continent-specific unique SNPs and shared polymorphisms, but no fixed polymorphism between the two continents, making it difficult to resolve between gene flow and retention of ancestral polymorphisms.

Although mitochondrial variation is severely reduced in NA population, nuclear sequence and microsatellite variation is similar in NA and EU populations, supporting the hypothesis that there has been a recent cytoplasmic sweep in NA populations, and ruling out a genetic bottleneck during a founder event of NA *N. vitripennis* derived from EU populations. The likely cause of a mitochondrial sweep is either an advantageous mitochondrial mutation or an advantageous *Wolbachia* mutation. A selective sweep of a *Wolbachia* type would result in hitchhiking of the associated mitochondrial haplotype due to their maternal coinheritance (Turelli and Hoffmann, 1991). As pointed out by Turelli (1994), advantageous mutations in *Wolbachia*, which makes it unidirectionally incompatible with the resident infection (that is, males are incompatible with resident bacteria-infected females) or reduce the fecundity cost of the infection, can sweep through the population and replace the resident *Wolbachia* strain. This is accompanied by the sweep of the linked mitochondrial haplotype in the process. There is some evidence that such *Wolbachia* sweeps have happened subsequent to establishment of an infection in a species, both in *D. melanogaster* (Riegler *et al*., 2005) and in *D. simulans* (Weeks *et al*., 2007). Furthermore, it would be expected to occur as *Wolbachia* endosymbionts adapt to a new host environment after horizontal transmission from a different insect, as we know has happened for several *Wolbachia* endosymbionts in the *Nasonia* system (Raychoudhury *et al*., 2009). The absence of variation in the five *Wolbachia* MLST loci in NA and EU populations is not problematic for the hypothesis of a *Wolbachia*-induced sweep, as beneficial mutation could occur anywhere in the $\sim$1 Mb genome and can result in a selective sweep. However, given that both the mitochondria-induced and the *Wolbachia*-induced sweep will create the same patterns of reduced mitochondrial variation, we cannot distinguish between these two causes with the current data. Fitness tests of the relevant *Wolbachia* and mitochondrial haplotypes together and separated from each other could help resolve this question.

The star-like pattern of the NA portion of the mitochondrial haplotype network can be interpreted as the result of a recently expanding population (Avise, 2000) or as incomplete mitochondrial sweep. Such historical patterns of mitochondrial variation have been interpreted to reflect genetic bottlenecks due to founder populations. Scheffer and Grissell (2003) found no mitochondrial genetic variation in the introduced populations of South American seed-feeding wasps (*Megastigmus transvaalensis*), but found a high variability in the African source population. Aoki *et al*. (2008) investigated the consequences of a glacial bottleneck on the mitochondrial genetic variation of a seed parasitic weevil (*Cuculio hilgendorfi*) in Japan and found a similar pattern. In contrast to these studies, in *N. vitripennis* the reduced variation is mirrored only in the mitochondria and not in the nuclear variation. The nuclear markers (both microsatellites and multi-gene sequences) show equal levels of variation. So, unlike the instances mentioned, there is no evidence that *N. vitripennis* in North America was introduced from Europe.

Our findings emphasize the importance of examining both mitochondrial and nuclear variation to infer genetic bottlenecks versus mitochondrial sweeps. The issue is particularly relevant to arthropods due to the high incidence of *Wolbachia* (Hilgenboecker *et al*., 2008), which can promote the incidences of mitochondrial sweeps.

Nasonia has an extremely high rate of synonymous mutation in its mitochondrial genes that are estimated to evolve at a rate nearly 35–40 times faster than the nuclear genes (Oliveira *et al*., 2008). However, the *Wolbachia* genome evolves at a rate that is approximately 1/3 that of the nuclear genes (Raychoudhury *et al*., 2009). Extrapolating, there are roughly 105–120 mutations in the *Nasonia* mitochondria for every mutation in *Wolbachia*. This can explain the mitochondrial diversity with respect to the lack of diversity in the *Wolbachia* infections of *N. vitripennis*. Furthermore, given the high mutation rate observed in *Nasonia* mitochondria, the high similarity between the common haplotype in North America and an uncommon variant in Europe implies that the sweep in North America would have been very recent. Thus, comparing nuclear, mitochondrial and *Wolbachia* variation in NA and EU populations, we conclude that NA populations have gone through a recent mitochondrial sweep, possibly due to a selective sweep of their resident *Wolbachia* infections.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank H Beran (Landesbund für Vogelschutz Bayern) and R Peters (University of Hamburg) for sending us samples; and D Winkler (Cornell University) and J Rogers for providing access to field populations; and R Sen for helpful comments on the manuscript. We also thank D Garrigan for guiding us through the simulations. This study was supported by a NWO Pioneer grant (ALW833.02.003) to LWB; a field grant of the Schure-Beijerinck-Popping Fonds of the Royal Netherlands Academy of Sciences to BKG; US NSF EF-0328363 and US NSF DEB-0821936 to JHW.

References

Aoki K, Kato M, Murakami N (2008). Glacial bottleneck and postglacial recolonization of a seed parasitic weevil, *Curculio hilgendorfi*, inferred from mitochondrial DNA variation. *Mol Ecol* **17**: 3276–3289.

Avise JC (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press: Cambridge, MA, USA.

Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR *et al.* (2006). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol* **72**: 7098–7110.

Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. *Mol Ecol* **13**: 729–744.

Bandelt HJ, Forster P, Rohl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**: 37–48.

Beukeboom L, Desplan C (2003). *Nasonia*. *Curr Biol* **13**: R860.

Bordenstein SR, O'Hara FP, Werren JH (2001). *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* **409**: 707–710.

Breeuwer JAJ, Werren JH (1990). Microorganisms associated with chromosome destruction and reproductive isolation between 2 insect species. *Nature* **346**: 558–560.

Breeuwer JAJ, Werren JH (1995). Hybrid breakdown between 2 haplodiploid species—the role of nuclear and cytoplasmic genes. *Evolution* **49**: 705–717.

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.

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.

Garrigan D, Lewontin RC, Wakeley J: Detecting natural selection from DNA sequence data. *Mol Biol Evol* (in press).

Goldstein DB, Pollock DD (1994). Least squares estimation of molecular distance–noise abatement in phylogenetic reconstruction. *Theor Popul Biol* **45**: 219–226.

Goudet J (2001). FSTAT, a program to estimate and test gene diversities and fixation indices. [2.9.3].

Grillenberger BK, Koevoets T, Burton-Chellew MN, Sykes EM, Shuker DM, van de Zande L *et al*. (2008). Genetic structure of natural *Nasonia vitripennis* populations: validating assumptions of sex ratio theory. *Mol Ecol* **17**: 2854–2864.

Haag-Liautard C, Coffey N, Houle D, Lynch M, Charlesworth B, Keithley PD (2008). Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biol* **6**: 1706–1714.

Hall TA (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

Hedrick PW (2005). A standardized genetic differentiation measure. *Evolution* **59**: 1633–1638.

Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008). How many species are infected with *Wolbachia*? A statistical analysis of current data. *FEMS Microbiol Lett* **281**: 215–220.

Hurst GDD, Jiggins FM (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc R Soc Lond B Biol Sci* **272**: 1525–1534.

Maniatis T, Fritsch EF, Sambrook J (1982). *Molecular Cloning (a Laboratory Manual)* 11th edn Cold Spring Harbour Laboratory Press: New York.

Nauta MJ, Weissing FJ (1996). Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics* **143**: 1021–1032.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.

Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH (2008). Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera : Pteromalidae). *Mol Biol Evol* **25**: 2167–2180.

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

Perrot-Minnot MJ, Guo LR, Werren JH (1996). Single and double infections with *Wolbachia* in the parasitic wasp *Nasonia vitripennis*: effects on compatibility. *Genetics* **143**: 961–972.

Pietsch C, Rütten K, Gadau J (2004). Eleven microsatellite markers in *Nasonia*, Ashmead 1904 (Hymenoptera; Pteromalidae). *Mol Ecol Notes* **4**: 43–45.

Pultz M, Leaf DS (2003). The jewel wasp *Nasonia*: querying the genome with haplo-diploid genetics. *Genesis* **35**: 185–191.

Raychoudhury R, Baldo L, Oliveira DCSG, Werren JH (2009). Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression and co-divergence in the *Nasonia* species complex. *Evolution* **63**: 165–183.

Raychoudhury R, Desjardins CA, Buellesbach J, Loehlin DW, Grillenberger BK, Beukeboom LW *et al*. (2010). Behavioral and genetic characteristics of a new species of *Nasonia*. *Heredity* (doi:10.1038/hdy.2009.147).

Riegler M, Sidhu M, Miller WJ, O'Neil SL (2005). Evidence for a global *Wolbachia* replacement in *Drosophila melanogaster*. *Curr Biol* **15**: 1428–1433.

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

Scheffer SJ, Grissell EE (2003). Tracing the geographical origin of *Megastigmus transvaalensis* (Hymenoptera: Torymidae): an African wasp feeding on a South American plant in North America. *Mol Ecol* **12**: 415–421.

Shuker DM, Lynch J, Peire Morais A (2003). Moving from model to non-model organisms? Lessons from *Nasonia* wasps. *Bioessays* **25**: 1247–1248.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Templeton AR, Crandall KA, Sing CF (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.

Turelli M, Hoffmann AA (1991). Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**: 440–442.

Turelli M (1994). Evolution of incompatibility-inducing microbes and their hosts. *Evolution* **48**: 1500–1513.

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**: 997–1005.

Werren JH (1980). Sex ratio adaptations to local mate competition in a parasitic wasp. *Science* **208**: 1157–1159.

Werren JH (1997). Biology of *Wolbachia*. *Annu Rev Entomol* **42**: 587–609.

Werren JH, Loehlin DW (2009). *Nasonia*:an emerging model system with haplodiploid genetics. *Emerging Model Organisms 2*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York.

Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK *et al*. (2010). Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* **327**: 343.

Whiting AR (1967). The biology of the parasitic wasp *Mormoniella vitripennis* [=*Nasonia brevicornis*] (Walker). *Q Rev Biol* **42**: 333–406.

Zhou W, Rousset F, O'Neil S (1998). Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc R Soc Lond B Biol Sci* **265**: 509–515.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)