

Multi locus analysis of *Pristionchus pacificus* on La Réunion Island reveals an evolutionary history shaped by multiple introductions, constrained dispersal events and rare out-crossing

KATY MORGAN,* ANGELA McGAUGHRAN,* LAURE VILLATE,* MATTHIAS HERRMANN,* HANH WITTE,* GABI BARTELMES,* JACQUES ROCHAT† and RALF J. SOMMER*

*Max Planck Institute for Developmental Biology, Department of Evolutionary Biology, Spemannstr. 37, Tübingen, Germany,

†Insectarium de la Réunion, Pépinière Communale, rue Simon Pernic, 97420 Le Port, Réunion, France

Abstract

Pristionchus pacificus, recently established as a model organism in evolutionary biology, is a cosmopolitan nematode that has a necromenic association with scarab beetles. The diverse array of host beetle species and habitat types occupied by *P. pacificus* make it a good model for investigating local adaptation to novel environments. Presence of *P. pacificus* on La Réunion Island, a young volcanic island with a dynamic geological history and a wide variety of ecozones, facilitates such investigation in an island biogeographic setting. Microsatellite data from 20 markers and 223 strains and mitochondrial sequence data from 272 strains reveal rich genetic diversity among La Réunion *P. pacificus* isolates, shaped by differentially timed introductions from diverse sources and in association with different beetle species. Distinctions between volcanic zones and between arid western and wet eastern climatic zones have likely limited westward dispersal of recently colonized lineages and maintained a genetic distinction between eastern and western clades. The highly selfing lifestyle of *P. pacificus* contributes to the strong fine-scale population structure detected, with each beetle host harbouring strongly differentiated assemblages of strains. Periodic out-crossing generates admixture between genetically diverse lineages, creating a diverse array of allelic combinations likely to increase the evolutionary potential of the species and facilitate adaptation to local environments and beetle hosts.

Keywords: hermaphrodite, Island biogeography, La Réunion Island, Mascareigne, Nematoda, *Pristionchus*

Received 14 July 2011; revision received 23 September 2011; accepted 17 October 2011

Introduction

Volcanic island chains, which have never been connected to other emerged lands, are ideal for investigating colonization and local adaptation to novel environments and as such are commonly referred to as 'laboratories of evolution' (Whittaker & Fernández-Palacios 2007). Local flora and fauna can only have been established through long-distance dispersal and

colonization, after which evolution must have continued in relative isolation from continental influence. This isolation provides an ecological opportunity for successful colonizers, as the variety of available ecological niches facilitates adaptive divergence in association with distinct local environments. Rapid genetic and phenotypic diversification in association with ecological shifts has been demonstrated in a range of taxa within both the Hawaiian and Galapagos Island chains (e.g. Jordan *et al.* 2005; Kleindorfer *et al.* 2006; Lawton-Rauh *et al.* 2007; Mathys & Lockwood 2011); the Mascareigne Island chain has to date received relatively little

Correspondence: Ralf J. Sommer, Fax: ++49 7071 601 498; E-mail: ralf.sommer@tuebingen.mpg.de

attention despite its classification as one of the world's major biodiversity hotspots (Myers *et al.* 2000).

As the youngest island within the chain, La Réunion has a landscape shaped by ongoing volcanic activity and characterized by the steepest altitudinal gradient and greatest variety of habitats of all the Mascareigne Islands. Two volcanoes dominate La Réunion, the extinct Piton des Neiges and the active Piton de la Fournaise, which reach peaks of 3070 and 2631 m, respectively (Strasberg *et al.* 2005). The Piton des Neiges originated as La Réunion emerged from the ocean approximately 2–3 Ma (Gillot & Nativel 1982) and is located in the northwest of the island (Fig. 1a). The three large calderas that dominate the landscape within this region, leading to dramatic changes in altitude across short geographic distances, were created during

the collapse of the centre of the Piton des Neiges following a period of prolonged activity, between 0.35 and 0.13 Ma (Gillot & Nativel 1982). The younger Piton de la Fournaise originated to the southeast of the Piton des Neiges approximately 0.53 Ma (Gillot & Nativel 1989). Thus, the island is characterized by younger southeastern and older northwestern regions (Fig. 1a). Further demarcation among 'ecotypes' is created by prevalent wind patterns across the island. Specifically, the climate on the eastern, windward side of the island is characterized by high rainfall, while the western, leeward side has a substantially drier climate. Together, climatic and geological variables conspire to create a complex of some 19 habitat types or 'ecozones' on La Réunion, within which species diversity and endemism broadly differ (Strasberg *et al.* 2005). The resulting

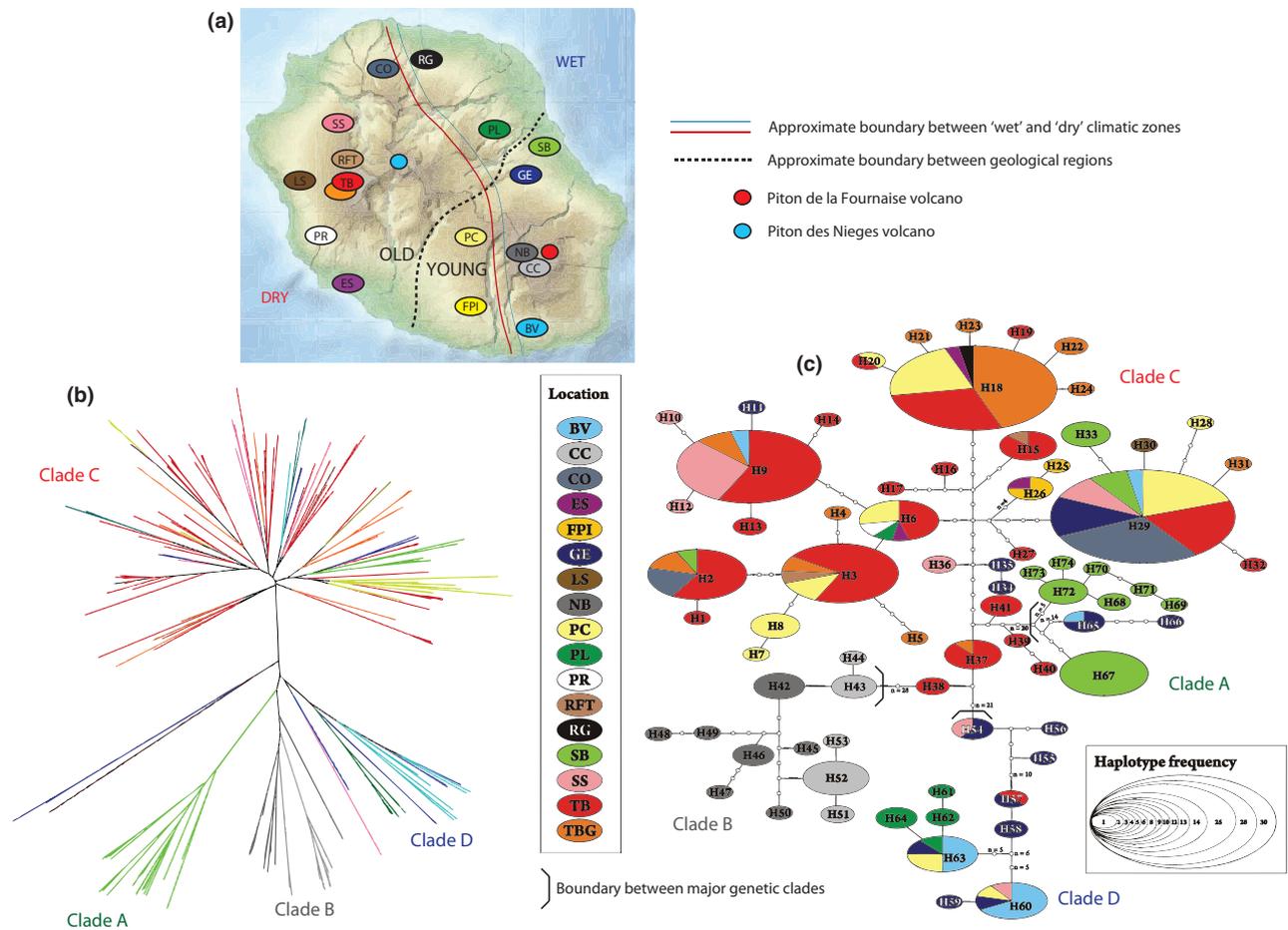


Fig. 1 The geographic distribution of *Pristionchus pacificus* genetic diversity across La Réunion Island; (a) a topographical map of La Réunion Island, indicating the sampling localities, the approximate positions of the island's two volcanoes, the geological divide between older and younger volcanic regions and the division between wetter and drier climatic zones; (b) a neighbour-joining tree showing the genetic relationships between strains, constructed using Single tandem repeat (STR) markers in Phylip; (c) a mt haplotype network, as constructed using *trcs* version 1.21, indicating genetic relationships between strains – the size of the circle is proportional to the frequency of the haplotype, and the length of branches is proportional to the number of mutations inferred between haplotypes; however, this network is not to scale. Boundaries between major clades are indicated on the mt network. Both the STR neighbour-joining tree and the mitochondrial haplotype network are colour-coded according to sampling locality.

environmental gradients and discontinuities across the island offer considerable opportunity for adaptive divergence and thought to have led to enhanced differentiation within taxa such as the bird species *Zosterops borbonicus* (Mila *et al.* 2010) and the mosquito species *Anopheles arabiensis* (Morlais *et al.* 2005) and *Aedes albopictus* (Paupy *et al.* 2001).

As a cosmopolitan species that inhabits a wide variety of habitat types and environments, the nematode species *Pristionchus pacificus* is a good model for investigating colonization dynamics and adaptive divergence in an island setting. *Pristionchus* species have an interesting ecology, being found in necromenic association with scarab beetles. Following the infestation of a suitable beetle host, the nematodes enter an arrested state, resuming development after host death to feed on bacteria growing on the beetle carcass. While the majority of *Pristionchus* species show high specificity to a particular beetle species, *P. pacificus* is unusually plastic in its host association and is known to infest several distantly related beetle species (Herrmann *et al.* 2006, 2010). On La Réunion Island alone, where it is by far the most prevalent *Pristionchus* species, *P. pacificus* has been isolated from more than ten beetle species, including both invasives and endemics (Herrmann *et al.* 2010). Thus, this island system provides an opportunity to investigate local adaptation and divergence not only in association with a range of ecozones but also in association with a diverse array of potential host beetle species.

The initial establishment of *P. pacificus* on La Réunion and subsequent adaptation to novel environments are likely to have been substantially influenced by colonization and population dynamics. For example, multiple introductions from diverse sources may increase adaptive potential through the generation of high genetic diversity (Roman & Darling 2007; Dlugosch & Parker 2008); a previous study of *P. pacificus* found high mitochondrial (mt) diversity within La Réunion populations, suggesting multiple introductions may indeed have contributed to successful colonization (Herrmann *et al.* 2010). The reproductive strategy of *P. pacificus* may also have contributed to its colonization success; this species has a hermaphroditic mating system, characterized by a predominance of self-fertilizing hermaphrodites with out-crossing only possible through the periodic production of males. The ability to self-fertilize and the consequent reproductive assurance may account for the widespread distribution of this species, which has been collected from across four continents, including Africa, Asia, America and Europe (Herrmann *et al.* 2010). Indeed, all 4 of the *Pristionchus* species detected on La Réunion Island to date are hermaphroditic, despite the fact that a gonochoristic system with obligate out-crossing and separate males and females is the predominant

and likely ancestral system within the *Pristionchus* genus (Mayer *et al.* 2007). This clear over-representation of hermaphroditism as a reproductive strategy among La Réunion *Pristionchus* species supports the hypothesis that hermaphroditism confers an advantage in terms of colonization ability. The frequency and role of out-crossing in *P. pacificus* natural populations are unknown (Click *et al.* 2009); however, the presence of periodic out-crossing provides the opportunity to investigate the roles of self-fertilization, out-crossing and recombination in island colonization and local adaptation.

The scope for island biogeographic studies using *P. pacificus* is further enhanced by the recent development of this species as a model organism in evolutionary biology, with a range of readily available genetic, genomic and transgenic tools (Hong & Sommer 2006; Schlager *et al.* 2009). Recent work has focused on fine-scale intraspecific natural variation in both developmental processes and life history traits in *P. pacificus*. By integrating these aspects with ecological and population genetics-based approaches, we can investigate how adaptive divergence in an island setting drives phenotypic change (e.g. Zauner and Sommer, 2007, Mayer & Sommer 2011) and how such change may be initiated at the microevolutionary level (Sommer 2009).

We investigate the evolutionary history of *P. pacificus* on La Réunion Island with a multilocus approach. We use 20 microsatellite [Single tandem repeat, or (STR)] markers designed to cover a representative sample of all six of *P. pacificus*' chromosomes, genotyping a total of 223 strains. In addition, we add a further 196 sequences to an existing ($n = 76$) mtDNA data set (Herrmann *et al.* 2010). Our sampling approach encompasses several beetle species and geographic regions across La Réunion Island to address the following questions: (i) Is the high genetic diversity previously detected using mtDNA (Herrmann *et al.* 2010), potentially generated by multiple independent colonizations of La Réunion, reflected in both the STR and extended mt data sets? (ii) Is there an association between population genetic structure and either geological or ecological divisions? (iii) Is an association between genetic lineages and specific host beetle species apparent, or do lineages show plasticity in their beetle association? (iv) Do genome-wide patterns of linkage disequilibrium (LD) confirm selfing as the predominant reproductive strategy?

Methods

Sample collection and DNA extraction

Beetle samples were collected from multiple locations on La Réunion Island (Supplementary material

Table S1; Fig. 1a) and processed as described in Herrmann *et al.* (2006, 2010). In brief, the carcasses of freshly sacrificed beetles were placed on individual NGM agar plates seeded with the *Escherichia coli* strain OP50. The plates were monitored daily over a period of up to 6 weeks for the emergence of nematodes. Any single adult nematodes emerging from the beetle carcasses were picked individually and placed onto fresh NGM plates and seeded with OP50. This was continued until five adult nematodes had been isolated from the beetle carcass or until eggs or larvae were observed on the plate, at which point we could no longer be confident that any adult nematodes observed were from the original 'P0' generation inhabiting the live beetle host. This protocol was followed in all instances except for three, for which a total of 7, 10 and 11 nematodes were isolated from individual *Oryctes borbonicus* beetles.

Isogenic lines were created from all single adult *Pristionchus pacificus* hermaphrodites by allowing them to reproduce and maintaining their offspring in culture. Throughout this manuscript, the term 'strain' refers to these laboratory-maintained isogenic lines. As the isogenic lines were maintained in the laboratory by selfing for at least ten generations, any heterozygosity present in the original P0 generation is expected to have been eroded. Although there are caveats associated with this protocol (see Discussion), these were unavoidable for this study, as extraction from the single P0 hermaphrodite yielded insufficient DNA for sequencing and STR genotyping. Species identity was confirmed by sequencing the small single-ribosomal RNA (SSU) locus (Mayer *et al.* 2007).

Sequence data from mt *ND6* and *ND4L* genes was available for 76 strains from a previous study (Herrmann *et al.* 2010); in this study, a further 196 strains (isolated from 17 geographic locations and 11 beetle species) were sequenced at these loci (see Fig. 1a and Table S1), which together form a 760-bp fragment, using the primers (HZ11896 and HZ11897), extraction protocol and thermal cycling conditions of Herrmann *et al.* (2010). Thus, in total, 272 lines comprise the mt data set; sequences generated in this study are available from GenBank, with the accession numbers JN812789–JN812965.

Developing and genotyping STR markers

Single tandem repeat markers were developed by scanning the *Pristionchus pacificus* genome using Tandem Repeats Finder (Benson 1999). Loose and strict methods were used, with alignment scores of 20 and 50, respectively, and alignment weights of {2, 7, 7} and {2, 3, 5}, respectively. The scan results were screened to eliminate markers included within gene predictions, as iden-

tified using the web-based resource: <http://Pristionchus.org>. Finally, 20 markers were selected so that all six chromosomes were represented. Pure and long motifs were preferentially selected, as these are likely to be more polymorphic. Primers were designed for the selected markers using PRIMER3 version 0.4.0 (Rozen & Skaletsky 2000). Details of the selected markers and primers are shown in Supplementary material Table S2.

An M13 tail (5'-CACGACGTTGTAACGAC-3') was attached to the 5' end of all forward primers and labelled for genotyping with one of the following dyes: 6-FAM, VIC, NED or PET (Applied Biosystems, Foster City, USA). STR amplification was performed in final volumes of 20 μ L, which included 12.5 mM MgCl₂, 200 μ M dNTPs and primers at concentrations of 0.4 μ M. All PCR were performed with an annealing temperature of 55 °C. Products labelled with different fluorescent dyes were multiplexed and genotyped on an ABI 3730xl, using the internal size standard LIZ-500 and the ABI GENEMAPPER SOFTWARE version 4.0 (Applied Biosystems).

A total of 223 strains (including 45 from Herrmann *et al.* (2010) and 178 new strains), isolated from 15 geographic locations and nine beetle species, were genotyped at their STR loci. Strains for which both mt sequence and STR data were obtained number 190. (see Fig. 1a and Table S1). STR genotype data are available from Dryad, with the accession: doi:10.5061/dryad.n1stov2.

Population genetic analyses

Diversity statistics. Both STR and mt data sets were partitioned in three ways, according to: (a) sampling locality, (b) host beetle species and (c) both host beetle species and sampling locality. Additionally, where multiple strains were isolated from one individual beetle, STR data were divided into beetle-specific partitions (d). STR variability was estimated for each data partition according to allelic richness and gene diversity (H_E), as calculated using ARLEQUIN version 3.0 (Excoffier & Schneider 2005). To control for the effect of sample size on allelic richness, rarefied allelic richness was calculated using ADZE version 1.0 (Szpiech *et al.* 2008). Mt diversity was estimated using Arlequin, according to the following statistics: haplotype diversity, nucleotide diversity (π), mean number of pairwise differences (θ) and the mean number of segregating sites (S).

Networks and differentiation indices. For the STR data set, a matrix of pairwise distances between individuals was constructed using MSANALYSER version 4.05 (Dieringer & Schlötterer 2003), based on Nei's Chord distance (Nei

1972). This distance matrix was used to construct a neighbour-joining (NJ) tree, using the Neighbor program within the PHYLIP version 3.69 (Felsenstein 1989) package. Owing to the potential for recombination between multiple nuclear markers, the imposition of a bifurcating (NJ) topology may be misleading. Thus, the construction of a neighbour-net network, which is likely to better reflect the history of recombination in the STR sample, was performed using SPLITSTREE version 4.11.3 (Huson & Bryant 2006). For the mt data set, TCS version 1.21 (Clement *et al.* 2000) was used to estimate a haplotype network based on the statistical parsimony algorithm of Templeton *et al.* (1992). A connection limit of 95% was applied, and predictions from coalescent theory were used to resolve ambiguous loops (Templeton *et al.* 1992; Crandall & Templeton 1993; Crandall 1994; Posada & Crandall 2001). The resulting network was examined, and the presence of unique haplotypes (singletons) was identified.

Pairwise genetic differentiation between STR data partitions, as defined according to partitioning schemes (a–d), was estimated through F_{ST} and R_{ST} statistics, calculated in Arlequin. Significance was determined using 1000 permutations. F_{ST} -based statistics may give biased and misleading estimates of differentiation, especially when estimated for highly polymorphic markers such as STRs (Jost 2008; Heller & Siegismund 2009; Leng & Zhang 2011). As an additional measure of genetic differentiation for comparison with F_{ST} , Jost's D was calculated using SMOGD version 1.2.5 (Crawford 2010). This latter statistic has been shown to give substantially more robust inferences than F_{ST} -derived statistics under certain conditions (Leng & Zhang 2011); however, all three measures revealed similar patterns of differentiation for the STR data set. Thus, we refer hereafter only to our R_{ST} results and proceeded to mt analyses using only F_{ST} statistics, for which we used Arlequin and evaluated significance based on 1000 permutations. Finally, exact tests for population differentiation based on allele/haplotype frequencies (Raymond & Rousset 1995) were performed for both data sets for each of the partitions (a–c).

The effect of geographic distance on genetic differentiation was estimated through tests of isolation by distance (IBD). For both the mt and STR data sets, the Mantel correlation coefficient (Mantel 1967) between the pairwise geographic and genetic ($F_{ST}/(1-F_{ST})$, and $R_{ST}/(1-R_{ST})$ for the mt and STR data, respectively) distance matrices was calculated using Arlequin, and significance was estimated using 5000 permutations.

Analysis of molecular variance. To examine the level of genetic structure according to sampling locality, host beetle species and both of these factors combined,

hierarchical analysis of molecular variance (AMOVA) was performed in Arlequin with both STR and mt data sets partitioned according to schemes (a–c). Additionally, AMOVA was performed to examine the effect of geological and ecological divisions (Fig. 1a) on levels of genetic variation. As the majority of the 19 'ecozones' identified by Strasberg *et al.* (2005) are represented by only one sampling site, the ecozone definitions were used to group sample locations into broader moist and arid ecological divisions. Thus, for AMOVA examining partition (a), groups were defined according to (i) the division between the older and younger geological regions of the island and (ii) the division between the wetter eastern and drier western sides of the island (Fig. 1a). Additionally, to investigate fine-scale population structuring within a sampling locality, AMOVA was performed on the subset of strains included in STR partitioning scheme (d).

Clustering analysis. The Bayesian clustering algorithms implemented in STRUCTURE version 2.2 (Pritchard *et al.* 2000) and InStruct (Gao *et al.* 2007) were used to estimate the most likely number of genetic clusters (K) in the STR data set and to probabilistically assign individuals to those clusters. InStruct is an extension to STRUCTURE designed to accommodate selfing and, rather than assuming random mating and Hardy-Weinberg equilibrium (HWE) within populations, avoids potential bias introduced by selfing (Falush *et al.* 2003) by allowing selfing or inbreeding rates to vary between clusters. InStruct was used alongside STRUCTURE for K values of 1–20 and the results were compared, as recommended by the authors (Gao *et al.* 2007). Optimal values of K were determined using deviance information criteria. Five independent MCMC chains were run in InStruct, each with 1 000 000 iterations, a burn-in of 500 000 generations and a thinning interval of 100 between retained draws. Burn-in and MCMC lengths of 10 000 and 100 000, respectively, were used in STRUCTURE, and convergence of chains was checked after run completion. For each program, six replicate runs were conducted for each value of K , and these were checked for concordance. Replicate runs of STRUCTURE and InStruct were analysed using CLUMPP version 1.1.2 (Jakobsson & Rosenberg 2007), which handles differences between replicate runs by matching clusters and average Q values across them. Results were graphically displayed using DISTRUCT version 1.1 (Rosenberg 2004).

Inferences of population structure may be influenced by departures from neutrality in molecular markers. Thus, outlier STR loci were identified using BAYESCAN version 1.0 (Foll & Gaggiotti 2008), which uses differences in allele frequencies between populations to

identify loci likely to be under the influence of selection. BayeScan was run with data partitioned by schemes (a) and (c). Neutrality was rejected for those loci at which the \log_{10} (Bayes factor) exceeded 1.5. This included the following loci when data were partitioned according to (a): M55, M59, M65, M66 and M69; and (c): M50, M55, M59, M65 and M66. STRUCTURE and InStruct were therefore run on both the full data set and the reduced data set of 14 neutral loci, from which inferred outlier loci were excluded (see Table S2).

Linkage Disequilibrium. Repeated self-fertilization leads to extensive LD across the genome and, as a result, rapidly reduces the effective recombination rate (Charlesworth 2003). To test the hypothesis of genome-wide LD, the exact test of LD, an extension of the Fisher's exact test based on contingency tables, was performed for all pairs of STR loci using Arlequin. The Markov chain was run for 100 000 steps, following an initial dememorization of 1000 steps. Additionally, the index of association I_A , a measure of multilocus LD, was estimated for the full data set, as well as each of the partitions according to schemes (a–c). This was performed using MULTILOCUS 1.3 (Agapow & Burt 2001). Significance was determined through 1000 randomizations.

Recombination events can also be detected through inconsistencies in the genetic relationships between individuals inferred from different molecular markers, including those that differ in their mode of inheritance (e.g. Haber *et al.* 2005). Thus, genetic relationships inferred using STR markers were compared with those inferred from mt markers for all strains. Additionally, STR data were examined to identify any individual strains with admixed eastern and western clade haplotypes.

Results

Diversity statistics

All data partitions according to schemes (a–c) showed some polymorphism across STR loci, and the majority were also polymorphic at mt loci. For the STR data sets, the number of alleles per locus ranged from 4 to 49, with a mean of 14.75 (Table 1). Gene diversity (H_E) ranged from 0.090 to 0.951, with a mean of 0.699. The number of haplotypes per genetic population ranged from 1 to 78 in the mt data set, depending on the partition analysed, while mt haplotype diversity (for $n > 5$) ranged from 0.409 to 0.953 (Supplementary material, Table S3). A total of 194 STR and 74 mt haplotypes were detected within 223 and 272 individuals, respectively.

Strains from Grand Etang (GE) harboured the greatest diversity at STR markers ($H_E = 0.633$, $Ar(5) = 0.912$), while strains isolated from Le Cratère Commersion (CC) and Plaine des Lianes (PL) harboured the least ($H_E = 0.221$ and 0.222 ; $Ar(5) = 0.301$ and 0.138 , respectively) (Table 1). Mt diversity was also highest for strains isolated from GE ($\pi = 0.036$, $\theta = 26.853$) and was lowest within Nez de Boeuf (NB) and Colorado (CO) strains ($\pi = 0.006$ and 0.006 ; $\theta = 4.692$ and 4.500 , respectively) (Supplementary material Table S3). With data partitioned according to host beetle species, strains isolated from *Adoretus* sp. harboured the greatest STR diversity ($H_E = 0.662$, $Ar(5) = 2.976$), and strains isolated from *Amneidus godefroyi* and *Marronus borbonicus* harboured the least ($H_E = 0.371$ and 0.089 ; $Ar(5) = 1.874$ and 1.231 , respectively) (Table 1). Strains isolated from *Adoretus* sp. also had the greatest mt diversity ($\pi = 0.034$, $\theta = 25.118$); while strains isolated from *Ataenius scutellaris*, *Hoplochelus marginalis* and *Oryctes borbonicus* harboured the least ($\pi = 0.001$, 0.013 , 0.013 ; and $\theta = 0.857$, 9.910 , 9.379 , respectively) (Table S3). Finally, of the 30 individual beetles from which two or more strains were isolated, 28 harboured at least two STR haplotypes. Within these 28 beetles, diversity was low, with strains isolated from the same beetle being monomorphic at the majority of, but not at all, STR loci.

Networks and differentiation indices

There are no obvious differences in the relationships inferred through the STR NJ tree and neighbour-net network. As geographic relationships are easier to visualize in the NJ tree, only the tree is displayed in Fig. 1a; the neighbour-net network is available as supplementary material (Supplementary Fig. S1). All trees/networks for both the STR and mt loci reveal four major clades (Figs. 1b,c and S1), corresponding to the four worldwide clades identified by Herrmann *et al.* (2010). Hereafter, we maintain the nomenclature used by Herrmann *et al.* (2010) and refer to these clades as 'A', 'B', 'C' and 'D' (Fig. 1a,b). These distinct clades are separated by 20 or more mutational steps (>5% uncorrected p-distance sequence divergence) in the mt network (Fig. 2b). The mt clade A includes strains from the eastern localities Saint Benoit (SB), GE and Basse Vallée (BV), whereas the STR clade A is made up exclusively of strains from SB. In both the STR and mt networks, clade B is made up exclusively of strains from the central volcano plateau (NB and CC), while clade D consists primarily of strains from the east (BV, GE and PL). Finally, clade C consists largely of 'western' strains, with a small number of 'eastern' (BV, GE, SB) strains also represented. Thus, both STR and mt data indicate a clear pattern of differentiation between eastern and

Table 1 Population genetic statistics for *Pristionchus pacificus* STR data, as partitioned by: (a) original geographic collection location; (b) original host beetle species; and (c) original geographic collection location and host beetle species

Population	<i>N</i>	<i>N_A</i>	Ar(5)	Ap(5)	<i>H_E</i>	<i>I_A</i>
Full data set						
Total	224	14.750	3.160	3.160	0.699	2.191 (<i>P</i> < 0.001)
(a) Partitioned data set						
Basse Vallée (BV)	9	2.900	2.426	0.428	0.496	4.148 (<i>P</i> < 0.010)
Le Cratère Commerson (CC)	16	1.650	1.588	0.301	0.221	3.003 (<i>P</i> < 0.010)
Colorado (CO)	12	2.350	2.154	0.682	0.410	5.922 (<i>P</i> < 0.010)
Grand Etang (GE)	16	4.900	2.962	0.912	0.633	4.003 (<i>P</i> < 0.010)
Nez de Boeuf (NB)	13	2.250	1.650	0.416	0.318	1.524 (<i>P</i> < 0.010)
Plaines des Cafres (PC)	25	3.850	2.395	0.481	0.503	2.948 (<i>P</i> < 0.010)
Plaine des Lianes (PL)	9	1.650	1.433	0.138	0.222	1.896 (<i>P</i> < 0.010)
Saint Benoit (SB)	29	4.100	2.404	0.742	0.528	4.412 (<i>P</i> < 0.010)
Sans Souci (SS)	10	3.150	2.283	0.334	0.458	3.415 (<i>P</i> < 0.010)
Trois Bassins (TB)	69	7.050	2.414	0.411	0.499	0.806 (<i>P</i> < 0.010)
Trois Bassins Garden (TBG)	7	3.100	2.476	0.464	0.468	-0.190 (<i>P</i> = 0.637)
(b) Partitioned data set						
<i>Adoretus</i> sp.	43	7.450	2.976	1.221	0.662	2.805 (<i>P</i> < 0.010)
<i>Amneidusgodefroyi</i>	24	2.850	1.874	0.811	0.371	2.043 (<i>P</i> < 0.010)
<i>Aphodius</i> sp.	13	2.700	2.375	0.310	0.496	7.819 (<i>P</i> < 0.010)
<i>Hoplia retusa</i>	38	5.350	2.574	0.632	0.556	1.837 (<i>P</i> < 0.010)
<i>Hoplochelus marginalis</i>	23	4.800	2.518	0.683	0.515	1.571 (<i>P</i> < 0.010)
<i>Maladera affinis</i>	18	4.150	2.465	0.518	0.554	4.883 (<i>P</i> < 0.010)
<i>Oryctes barbonicus</i>	58	6.750	2.420	0.576	0.495	1.236 (<i>P</i> < 0.010)
<i>Marronus borbonicus</i>	5	1.150	1.231	0.107	0.089	-0.167 (<i>P</i> = 0.759)
(c) Partitioned data set						
BV <i>Adoretus</i>	9	2.900	2.535	0.406	0.496	4.148 (<i>P</i> < 0.010)
CC <i>Amneidus</i>	16	1.650	1.568	0.232	0.221	3.003 (<i>P</i> < 0.010)
CO <i>Adoretus</i>	5	1.000	1.154	0.478	0.036	-0.250 (<i>P</i> = 1.00)
CO <i>Hoplochelus</i>	7	1.850	2.098	0.398	0.348	6.384 (<i>P</i> < 0.010)
GE <i>Adoretus</i>	15	4.650	3.024	0.913	0.627	4.151 (<i>P</i> < 0.010)
NB <i>Amneidus</i>	8	2.000	1.699	0.294	0.307	1.740 (<i>P</i> < 0.010)
NB <i>Marronus</i>	5	1.150	1.231	0.107	0.089	-0.168 (<i>P</i> = 0.759)
PC <i>Hoplia</i>	25	3.850	2.502	0.403	0.503	2.949 (<i>P</i> < 0.010)
PL <i>Adoretus</i>	9	1.650	1.466	0.089	0.222	1.896 (<i>P</i> < 0.010)
SB <i>Aphodius</i>	14	3.000	2.462	0.195	0.512	7.819 (<i>P</i> < 0.010)
SB <i>Maladera</i>	15	3.100	2.180	0.308	0.438	3.331 (<i>P</i> < 0.010)
SS <i>Oryctes</i>	10	3.150	2.382	0.448	0.458	3.415 (<i>P</i> < 0.010)
TB <i>Hoplia</i>	11	3.000	2.355	0.326	0.482	2.666 (<i>P</i> < 0.010)
TB <i>Hoplochelus</i>	13	3.550	2.400	0.373	0.466	1.598 (<i>P</i> < 0.010)
TB <i>Oryctes</i>	45	5.750	2.446	0.406	0.469	1.138 (<i>P</i> < 0.010)

STR, Single tandem repeat; *n*, sample size; *N_A*, number of alleles.

Rarefied allelic richness, using a maximum standardized sample size of 5 (Ar(5)).

Rarefied private allelic richness, using a maximum standardized sample size of 5 (Ap(5)).

Expected heterozygosity (*H_E*); Standardized multilocus index of association (*I_A*).

western localities, with the majority of strains from eastern localities falling into clades A, B and D, and the majority of strains from western localities falling into clade C (Fig. 1).

The Mantel tests indicated no significant correlation between geographic and either mt or STR genetic distances; thus, we find no evidence for isolation-by-distance effects in our data (*P* = 0.0582 and 0.351 for mt and STR data, respectively). Both STR *R_{ST}* and mt *F_{ST}*

statistics indicated marked and significant differentiation among *Pristionchus pacificus* populations as defined according to all three partitioning schemes (Supplementary material Table S4). With data partitioned according to sampling locality (and where *n* < 5), both STR and mt markers showed localities CC, NB and PL to be among the most strongly differentiated (mean STR *R_{ST}* = 0.685, 0.617, 0.603; and mean mt *F_{ST}* = 0.563, 0.750, 0.630, respectively) and GE to be the least

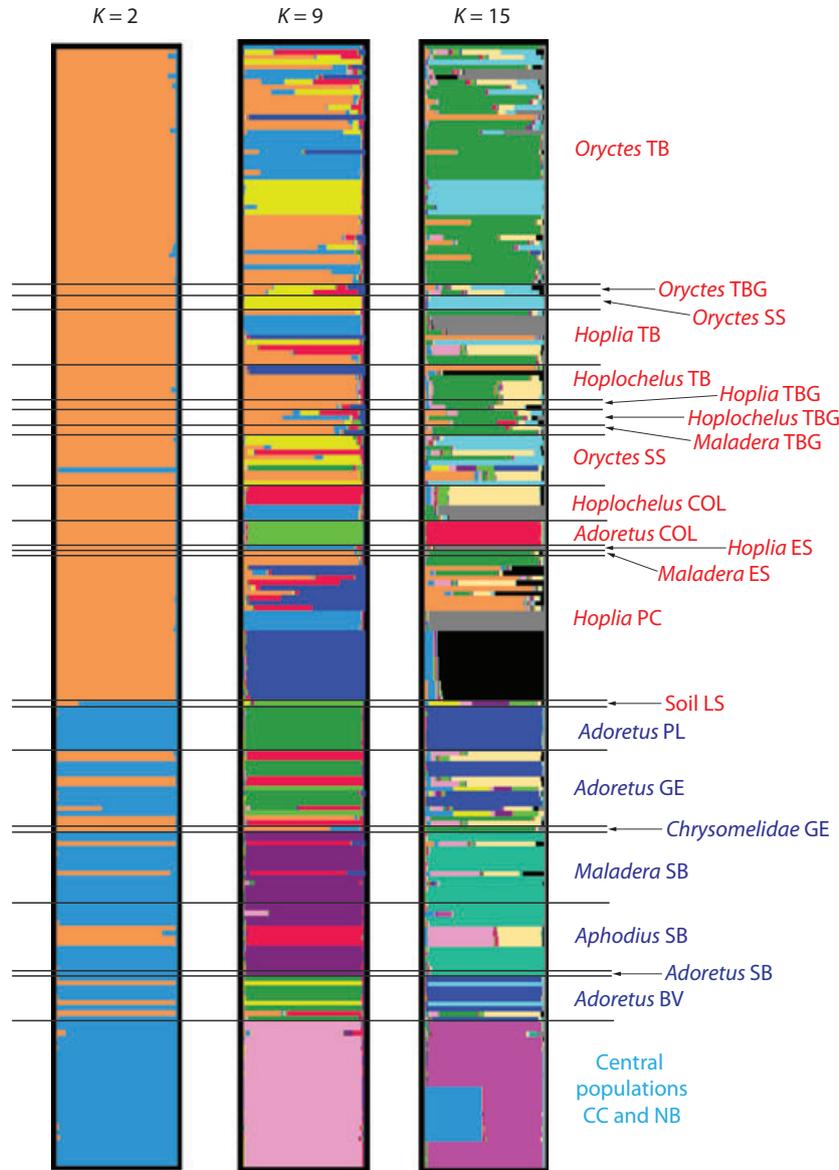


Fig. 2 Population genetic structure as inferred using the Bayesian clustering algorithm in InStruct. Each horizontal bar represents a *Pristionchus pacificus* strain; the coloured segments of each bar represent the proportion of that strain's genetic membership to a particular cluster. The sampling locality and host beetle species from which strains were isolated are indicated. Population genetic structure schemes are presented for $K = 2, 9$ and 15 .

differentiated (STR $R_{ST} = 0.210$, mt $F_{ST} = 0.245$). When data were partitioned according to host beetle species, strains isolated from *A. godefroyi*, *M. borbonicus* and *Maladera affinis* were among the most strongly differentiated for both genetic markers (mean STR $R_{ST} = 0.460, 0.663$ and 0.445 ; mean mt $F_{ST} = 0.566, 0.715$ and 0.459).

When considered individually, differentiation statistics for strains isolated from different beetle species at the same location were low and often non-significant (Table S4). For example, at SB, low but significant differentiation was detected between strains isolated from *M. affinis* and from *Aphodius* sp. at both STR and mt

markers (STR $R_{ST} = 0.105$, mt $F_{ST} = 0.105$). At Trois Bas-sins (TB), STR differentiation was also low but significant between strains isolated from *H. marginalis*, *H. retusa* and *Oryctes borbonicus*. Mt differentiation between strains isolated from different beetle species within TB was generally non-significant, with only *H. retusa* and *H. marginalis* showing significant (low) differentiation ($F_{ST} = 0.087$) (Table S4). Conversely, where strains were isolated from the same host beetle species but from different geographic locations, differentiation indices were mostly significant but low at both markers. For example, strains isolated from *A. godefroyi*

from CC and NB were significantly differentiated at both markers (STR $R_{ST} = 0.342$, mt $F_{ST} = 0.202$). The same was true for strains isolated from *H. retusa* at Plaines des Cafres (PC) and TB (STR $R_{ST} = 0.151$, mt $F_{ST} = 0.083$) and from *O. borbonicus* at Sans Souci (SS) and TB (STR $R_{ST} = 0.269$, mt $F_{ST} = 0.079$) (Table S4).

When considered in the light of geological divides across the island (Fig. 1a), where $n > 5$, differentiation between populations located in younger vs. older regions of the island was higher on average than differentiation between populations located within younger or older regions, and populations located in the younger area were more differentiated than those in the older (mean R_{ST} and $F_{ST} = 0.465$ and 0.588 for STR and mt, respectively, for younger vs. older; mean $R_{ST} = 0.440$ and $F_{ST} = 0.456$ between younger locations; mean $R_{ST} = 0.215$ and $F_{ST} = 0.258$ between older locations). Similarly, populations from locations on the wetter side of the island (Fig. 1a) were more differentiated than those on the drier side (mean $R_{ST} = 0.395$, $F_{ST} = 0.464$ between wet locations; mean $R_{ST} = 0.262$, $F_{ST} = 0.239$ between dry locations), and differentiation between populations from wet and dry locations was higher still (mean $R_{ST} = 0.527$ and $F_{ST} = 0.532$).

Analysis of molecular variance

AMOVA of data sets partitioned according to schemes (a–c) revealed that significant variation was explained by sampling locality (47.65% of STR variation, $P < 0.001$; 50.28% of mt variation, $P < 0.001$), host beetle species (36.18%, $P < 0.001$ for STR; 42.05%, $P < 0.001$ for mt) and both sampling locality and host beetle species combined (49.16%, $P < 0.001$; and 50.18%, $P < 0.001$) (Tables 2 and 3). Thus, sampling locality explained a slightly higher level of genetic variance than host beetle species, for both genetic markers.

The distinction between geologically older and younger localities did not explain a significant amount of either STR (7.52%, $P = 0.182$) or mt variation (9.06%, $P = 0.070$). The classification of localities according to 'wet' or 'dry' (Fig. 1a), however, explained a significant amount of both STR and mt variation (33.60%, $P < 0.001$; 27.38%, $P = 0.005$) (Tables 2 and 3).

When AMOVA was performed on the STR data set as partitioned by scheme (d), where beetles harboured multiple nematode strains, the majority of variation among strains was partitioned across different beetles (82.62%, $P < 0.001$). However, significant variation among strains isolated from the same beetle was also present (17.38%, $P < 0.001$), indicating that divergent strains can infest the same individual host. When strains from individual beetles were further grouped by geographic location, the majority of variation was

distributed between individual beetles within locations (57.13%, $P < 0.001$), but a significant level of variation was also distributed between these locations (26.00%, $P < 0.001$) (Table 2).

Clustering analyses

There were no obvious differences in the assignment of individuals to populations as performed in InStruct and STRUCTURE, using either the full or reduced data sets (Table S2); therefore, only the results of InStruct using the full data set are presented graphically in Fig. 2. At $K = 2$, individuals were confined to two clusters, one of which consisted almost exclusively of individuals from the eastern side of the island, while the other included mostly strains isolated from the west (Fig. 2). These two clusters correspond to mt and STR clades A, B and D (east), and C (west), respectively (Fig. 1). The three geographically defined clusters corresponding to the central plateau localities CC and NB, the eastern localities BV, GE and PL, and the eastern locality SB became apparent as K was increased to six and remained consistent as K was increased up to 15 (Fig. 2). The cluster comprising localities CC and NB corresponds to clade B; that comprising localities BV, GE and PL corresponds to clade D; and that comprising SB corresponds to clade A (Figs 1a,b and 2). Increasing the number of clusters from $K = 6$ to $K = 15$ resulted in increased division among the 'western' (clade C) strains, although the additional clusters showed high admixture and did not clearly correspond to either host beetle species or sampling locality.

When all loci were included, the optimal number of clusters identified by both STRUCTURE and InStruct was $K = 15$, with $\text{LnP}(D)$ remaining high for $K = 16$ – 20 for both algorithms. For the reduced data set excluding the outlier loci identified by BayeScan (Table S2), both STRUCTURE and InStruct identified $K = 14$ as the optimal number of clusters, with $\text{LnP}(D)$ falling off more sharply for $K > 14$. Finally, all clusters identified by InStruct were estimated to have selfing rates of between 0.80 and 0.98.

Evidence for recombination events

The relationships between strains inferred from STR data were largely similar to those inferred using mt markers, with the exception of four discrepancies indicating potential recombination/admixture events (Fig. 3). Within the STR clade C strains, RS5401, isolated from *Adoretus* sp. at GE, falls within the mt clade D. Strains RS5410 and RSA059, also both isolated from *Adoretus* sp. at GE, showed the reverse pattern, falling within STR clade D but mt clade C. Finally, RS5419,

Table 2 Percentage of variation (%) of molecular variance attributed to various levels of hierarchical population structure for *Pristionchus pacificus* STR markers, for analyses characterized according to: (a) geographic sampling locality; (b) host beetle species; (c) both geographic sampling locality and host beetle species; and (d) individual beetle host

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
(a) Test 1 (no structure)				
Among populations	14	102346.305	257.52707 V_a	47.65 ($P < 0.001$)
Within populations	433	122531.776	282.98332 V_b	52.35
Total	447	224878.080	540.51039	
F_{ST} : 0.47645				
Test 2 (structure: geological; young and old)				
Among groups	1	2.4866.879	41.77408 V_a	7.52 ($P = 0.182$)
Among populations within groups	13	75789.758	230.01394 V_b	41.42 ($P < 0.001$)
Within populations	433	122531.776	282.98332 V_b	51.05 ($P < 0.001$)
Total	447	224878.080	615.53066	
F_{ST} : 0.48946, F_{SC} :0.44793, F_{CT} : 0.07523				
Test 3 (structure: ecological; wet and dry)				
Among groups	1	53675.764	211.36664	33.60 ($P < 0.001$)
Among populations within groups	13	46980.874	134.24704	21.34 ($P < 0.001$)
Within populations	433	122531.776	282.98332 V_b	44.99 ($P < 0.001$)
Total	447	224878.080	629.10770	
F_{ST} : 0.54937, F_{SC} :0.32136, F_{CT} : 0.33598				
(b)				
Among populations	7	73426.034	192.33887 V_a	36.18 ($P < 0.001$)
Within populations	436	147929.488	339.28782 V_b	63.82 ($P < 0.001$)
Total	445	221355.523	531.62669	
$F_{ST} = 0.36179$				
(c)				
Among populations	19	109059.794	264.75614 V_a	49.16 ($P < 0.001$)
Within populations	412	112797.836	273.78115 V_b	50.84 ($P < 0.001$)
Total	445	221857.630	538.53730	
$F_{ST} = 0.49162$				
(d) Test 1 (no structure)				
Among populations	20	77483.524	381.92887 V_a	82.62 ($P < 0.001$)
Within populations	189	15186.038	80.34941 V_b	17.38 ($P < 0.001$)
Total	209	92669.562	462.27828	
$F_{ST} = 0.82619$				
Test 2 (structure: by sampling locality)				
Among groups	6	38828.494	123.92170 V_a	26.00 ($P < 0.001$)
Among populations within groups	14	38655.030	272.26470 V_b	57.13 ($P < 0.001$)
Within populations	189	15186.038	80.34941 V_c	16.86 ($P < 0.001$)
Total	209	92669.562	476.53581	
F_{ST} : 0.83139, F_{SC} : 0.77213, F_{CT} : 0.26005				

isolated from *Maladera affinis* at SB, falls within clade A of the STR tree/network and clade C of the mt network. Within the STR data, several strains showed apparently admixed haplotypes, consisting of both predominantly eastern and predominantly western alleles (Fig. 3) and providing additional evidence of potential recombination events.

The mt data support the occurrence of admixture events. Specifically, a total of 74 haplotypes were found from 272 individuals; of these, 41 were singletons and 33 were shared (Fig. 1c; Table S3). Of the 17 geographic locations, just three [La Saline (LS), but $n = 1$; CC and NB] did not share haplotypes with other

locations. The location TB shared the largest number (ten) of haplotypes with other locations, while GE and PC each shared seven haplotypes among locations. Overall, 18 instances of haplotype sharing were within geographic locations. Six instances of sharing across the northeast/southwest divide occurred; these involved the eastern location BV three times, SB twice and GE four times. Sharing of haplotypes among beetle species was seen in 15/33 instances (Table S3). For example, haplotype H9 characterized several strains isolated from the beetles *Adoretus* sp., *Aphodius* sp., *H. retusa* and *Oryctes borbonicus*, while strains with haplotype H29 were found on *Adoretus* sp., *H. retusa*, *M. affinis* and

Table 3 Percentage of variation (%) of molecular variance attributed to various levels of hierarchical population structure for *Pris-tionchus pacificus* mitochondrial sequences (mtDNA *ND6* and *ND4L*), for analyses characterized according to: (a) geographic sampling locality; (b) host beetle species; and (c) both geographic sampling locality and host beetle species

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
(a) Test 1 (no structure)				
Among populations	16	1534.224	6.02604 V_a	50.28 ($P < 0.001$)
Within populations	258	1537.296	5.95851 V_b	49.72 ($P < 0.001$)
Total	274	3071.520	11.98456	
F_{ST} : 0.50282				
Test 2 (structure: geological; young and old)				
Among groups	1	346.957	1.12994 V_a	9.06 ($P = 0.070$)
Among populations within groups	15	1187.267	5.37785 V_b	43.14 ($P < 0.001$)
Within populations	258	1537.296	5.95851 V_c	47.80 ($P < 0.001$)
Total	274	3071.520	12.46631	
F_{SC} : 0.47439; F_{ST} : 0.52203; F_{CT} : 0.09064				
Test 3 (structure: ecological; wet and dry)				
Among groups	1	617.358	3.74534 V_a	27.38 ($P = 0.005$)
Among populations within groups	15	916.865	3.97542 V_b	29.06 ($P < 0.001$)
Within populations	258	1537.296	5.95851 V_c	43.56 ($P < 0.001$)
Total	274	3071.520	13.67927	
F_{SC} : 0.40019; F_{ST} : 0.56441; F_{CT} : 0.27380				
(b)				
Among populations	10	1227.627	5.02870 V_a	42.05 ($P < 0.001$)
Within populations	262	1815.519	6.92946 V_b	57.95 ($P < 0.001$)
Total	272	3043.147	11.95816	
F_{ST} : 0.42052				
(c)				
Among populations	28	1632.979	5.82185 V_a	50.18 ($P < 0.001$)
Within populations	244	1410.160	5.77935 V_b	49.82 ($P < 0.001$)
Total	272	3043.139	11.60119	
F_{ST} : 0.50183				

STR, single tandem repeat.

Strain	Microsatellite clade	M50	M52	M53	M58	M59	M62	M66	Mitochondrial clade
RSB055	C	231	212	242	282	202	192	196	C
RSB057	C	231	212	174	282	202	192	196	C
RSB073	C	231	172	272	252	224	194	196	C
RSA045	C	231	184	270	252	314	192	196	C
RS5409	C	231	200	218	252	280	194	196	D
RS5410	D	231	180	260	246	180	192	200	C
RSA106	D	231	180	174	244	258	196	222	D
RS5421	A	231	172	170	242	212	190	204	A
RSB069	A	229	180	170	234	192	188	200	A
RSB071	A	229	172	170	242	192	194	200	A
RSB086	A	229	172	186	252	214	190	204	A
RSA059	D	229	184	174	246	208	196	200	C
RS5419	A	229	172	170	242	198	190	200	C
RS5406	D	229	184	174	246	222	196	200	D

Fig. 3 Examples of strains for which there are inconsistencies between genetic relationships as inferred using different molecular markers, indicating instances of out-crossing and recombination.

O. borbonicus. Haplotype sharing occurred in 12, 11, 9 and 9 instances for *O. borbonicus*, *H. marginalis*, *Adoretus* sp. and *H. retusa*, respectively, while no instances of haplotype sharing among beetle species were seen for *A. godefroyi* or *M. borbonicus*.

Linkage disequilibrium

Highly significant pairwise LD ($P < 0.001$) was detected between all pairs of STR loci, indicating extensive LD across the genome. Multilocus LD, as estimated by the

standardized index of association I_A , was highly significant for the full data set ($I_A = 2.191$, $P < 0.001$) and for the majority of the data partitions with $n > 5$. Within sampling localities, I_A ranged from 0.806 ($P < 0.010$) at TB to 5.922 ($P < 0.010$) at CO (Table 1).

Discussion

Analyses based on 194 STR and 74 mt haplotypes detected from 223 and 272 strains, respectively, found substantial genetic variation and population structure within La Réunion *Pristionchus pacificus*. This manifested in the presence of highly significant differentiation indices and a low degree of haplotype sharing among populations. Approximately 15 subpopulations characterize the *P. pacificus* 'meta-population', and these are further defined by ecological and/or geological constraints, which have generated contrasting patterns in the east and west of the island. The patterns observed in this small island setting are markedly distinct from the low genetic diversity and lack of structure detected among global populations of *Caenorhabditis elegans* (Rockman & Kruglyak 2009), the hermaphroditic nematode with which, owing to their similar life cycle and status as model organisms in evolutionary biology, *P. pacificus* is commonly compared. Rather, population genetic patterns detected in *P. pacificus* would seem more similar to those reported in the partially selfing hermaphroditic killifish species *Kryptolebias marmoratus*, which shows high genetic diversity and strong population structure over a range of spatial scales (Tatarenkov *et al.* 2007). Their contrasting population genetic patterns suggest that, despite their similar life cycle and reproductive system, *P. pacificus* and *C. elegans* have very different evolutionary histories.

Four major clades or lineages on the island are identified based on the STR/mt networks and InStruct groupings. The high genetic distances between these four major clades (>20 mutational steps/>5% uncorrected p-distance sequence divergence between clades at mt loci) and the concordant patterns of differentiation at unlinked mt and nuclear markers suggest a long period of isolation with consequent divergence through genetic drift. While strong diversifying selection could lead to rapid divergence at nuclear markers linked to those under selection, this would not be reflected by high levels of concordant differentiation at mt loci. Thus, the most likely explanation for the patterns of differentiation observed is that, as proposed by Herrmann *et al.* (2010), multiple invasions of La Réunion Island from genetically distinct global populations of *P. pacificus* have generated the rich genetic diversity detected on the island. The alternative hypothesis of a single early colonization would require a period of isolation on La Réunion Island

sufficient to drastically limit gene flow and lead to extensive differentiation at unlinked neutral markers. Given the relatively young age of La Réunion Island (2–3 Ma), this seems unlikely. Presence of highest genetic diversity and lowest differentiation from other locations, coupled with it being a location for which admixture between mt and STR eastern and western haplotypes occurs, highlights GE (Fig. 1a) as the potential entry point for incoming strains. Alternatively, this location may have served as a refuge in historical times. From GE, high genetic diversity resulting from multiple introductions may have favoured the successful subsequent colonization of *P. pacificus*, through increased evolutionary potential aiding adaptation to the island's wide variety of local environments (Roman & Darling 2007; Crawford & Whitney 2010).

The rich genetic diversity characterizing La Réunion *P. pacificus* is not, however, distributed uniformly across the island; the four major clades differ markedly with respect to their geographic distribution and general pattern of differentiation. Clade C includes strains sampled from locations across the island, but with a higher proportion sampled from the west, whereas clades A, B and D have more limited and primarily eastern distributions. While strains within clades A, B and D each fall into stable InStruct groupings, the strains within clade C are divided into multiple groups at the optimal K -value (Fig. 2), suggesting greater heterogeneity and population structure in clade C. Because the four major clades most likely result from independent colonizations, the relative timing of invasion events may explain the differences in distribution and differentiation. An early colonization of La Réunion would have allowed for substantial dispersal across the older western part of the island and the consequent structure and widespread western distribution of clade C observed today, whereas more recent eastern colonization of the island by the remaining three lineages would have left little time for dispersal to other geographic regions. Clade C was revealed through analysis of a worldwide selection of strains to be almost exclusive to La Réunion (Herrmann *et al.* 2010). Assuming the considerable diversity within clade C has evolved *in situ* on La Réunion, this supports placing the colonization event for this lineage at an earlier date than that of the other clades.

Differential colonization events through time are not the only potential factors that may have limited the dispersal of eastern lineages. The lack of evidence for an association between geographic distance and genetic differentiation suggests that the clear genetic break between eastern and western sampling localities, as evidenced by the InStruct groupings at $K = 2$ (Fig. 2), is caused by some barrier to dispersal between these

geographic regions, rather than by isolation with geographic distance *per se*. The distribution of clade A strains (e.g. at localities BV, PL and GE) suggests a distinct pattern of dispersal across large geographic distances within the east of the island, but a failure to disperse to western localities. The substantial level of variation explained by grouping populations according to 'wet' and 'dry' climatic zones suggests a role for ecology in limiting dispersal and driving differentiation. Local adaptation, together with geological boundaries created by ongoing volcanic activity, may drive rapid phenotypic and genotypic divergence between climatic zones, as shown by the spatially structured patterns of genetic and metabolic rate variation within springtail (*Cryptopygus antarcticus*) populations on the 1-Ma-old Marion Island (McGaughan *et al.* 2010). On La Réunion, local adaptation to the distinct climates on the wet eastern and arid western sides of the island is proposed to have driven differentiation between populations of mosquitoes (Paupy *et al.* 2001; Morlais *et al.* 2005). It is possible that similar processes are influencing genetic structure in La Réunion *P. pacificus*, with eastern lineages being poorly adapted to the arid western climate and clade C strains being better able to tolerate the western ecological conditions.

The distinction between geological volcanic regions on La Réunion provides an additional potential explanation for the observed pattern of east-west differentiation. Although current geological classifications (Fig. 1a) do not explain a significant amount of genetic variation, the precise location of geological barriers is likely to have changed substantially throughout the islands' dynamic volcanic history. Geological barriers originating during the period of intense volcanic activity around 0.53 Ma (Gillot & Nativel 1989a,b) may have limited the dispersal of more recent colonizers but may not have been present during the colonization of La Réunion by clade C haplotypes.

Because the most likely mode of *P. pacificus* dispersal is via the movements of their beetle hosts, any climatic differences/geological boundaries may be shaping nematode distribution indirectly, by limiting the movements of their beetle hosts. Association of certain genetic lineages with specific host beetle species may also explain differences in the distribution of nematode lineages, owing to differential dispersal capacities and distribution ranges of the host beetles. Host specificity in nature would lead to enhanced differentiation between isolated populations and may be maintained via mechanisms such as differential chemoattraction of certain lineages to specific hosts (A McGaughan, K Morgan and R J Sommer, unpublished data). Differential host preferences would lead to differences in the local environments experienced by strains, such as the assemblage of bacteria

feeding on the beetle carcass and the habitats frequented by the hosts and may lead to selection-driven phenotypic differentiation between strains (e.g. Mayer & Sommer 2011). However, the evidence for an association between genetic structure and host beetle species is not entirely clear. The AMOVA results show a significant amount of variation is explained by host beetle species, and some specificity in the beetle host association across genetic clades A, B and D is apparent. For example, clade A and clade B strains are only found in association with *Aphodius* sp. and *Maladera affinis*, and *A. godefroyi* and *M. borbonicus*, respectively, and clade D is primarily made up of strains isolated from *Adoretus* sp. located at multiple, geographically distant sampling localities. Because La Réunion *Adoretus* and *Aphodius* ssp. are known to constitute species complexes, however, it is not known whether all strains isolated from each of these genera are from one or from several species.

The lack of clear clustering of *P. pacificus* haplotypes according to host species within a sampling locality and the generally low or non-significant levels of differentiation between strains isolated from different beetle species contrast with the afore mentioned AMOVA results. This apparent discrepancy may be caused by the interaction between host beetle species and sampling locality, i.e. distinguishing between the effects of host beetle species and geography on population structure is difficult, owing to the differing distribution ranges of beetle species. For example, the high differentiation of strains from localities NB and CC may be a result of specificity of clade B strains to their host beetles, *A. godefroyi* and *M. borbonicus*, or it may be due to local adaptation to the cooler, high-altitude climatic conditions that characterize the central plateau localities (Fig. 1a). Strains in clade B isolated from either of their taxonomically distinct beetle hosts show little differentiation, suggesting a lack of strong specificity with respect to beetle species *per se*. However, all strains from *M. borbonicus* have, to date, been isolated from one beetle specimen; more extensive sampling of this and other potential host beetle species within the region is clearly necessary. A degree of plasticity with respect to host species is supported by the apparently recent association of clade C strains with novel invasive species. *Maladera affinis* and *H. marginalis* are recently invasive species of La Réunion and thought to have colonized the island from Asia and Madagascar, between 800 and 1800 and during the 1970s, respectively (Vercambre *et al.* 1991; Cheke & Hume 2008). Despite their recent invasion, both beetle species harbour strains from clade C, which is also associated with several endemic La Réunion beetle species and has an early inferred colonization date. This suggests the recent association of clade C strains with novel invasive beetle species (in the case of

H. marginalis, within the last 50 years). Association with novel hosts may represent spurious founding events akin to the colonization of new geographic areas in other species. Such founding events may lead to non-adaptive genotype–host associations, as some genotypes may have by chance encountered a new beetle species (with distinct habitat breadth) first.

Patterns of genetic diversity across the island are influenced not only by adaptive and/or extrinsic forces such as dispersal and host beetle ecology but also by the levels of out-crossing within populations. Strong levels of LD across all six chromosomes, resulting from reduced effective recombination following repeated selfing (Charlesworth 2003), clearly support self-fertilization as the predominant mode of reproduction. However, close examination of genotypic data reveals evidence of periodic out-crossing and recombination. The clustering algorithm implemented in InStruct estimated selfing rates of between 80% and 96%, leaving low but significant levels of out-crossing. The observed high genetic diversity within *P. pacificus* would seem surprising given the species' mainly selfing lifestyle. However, even rare instances of out-crossing within predominantly selfing populations can result in the rapid generation of vast genotypic diversity. Extensive population genetic surveys of the plant species *Arabidopsis thaliana*, for example, have revealed that despite the predominance of selfing in this species, levels of polymorphism are not substantially depleted (Nordborg *et al.* 2005). Specifically, out-crossing followed by repeated generations of selfing can create large arrays of novel allelic combinations across the genome, through the generation of sets of naturally recombinant inbred lines (Siol *et al.* 2008). Periodic out-crossing between genetically divergent strains (for example, from eastern and western lineages), as is evidenced from the observed patterns of admixture at location GE (Fig. 3), would seem especially important. Thus, despite their apparently low periodicity, out-crossing events, coupled with the presence of distinct genetic lineages originating from multiple introductions, may play a vital role in adaptation to novel and harsh environments, through the generation of novel allelic combinations on which natural selection can act (Anderson *et al.* 2010).

An estimation of the level of heterozygosity in natural populations would also give an indication of the frequency of out-crossing, because heterozygosity is rapidly eroded by repeated generations of self-fertilization. Because this study used isogenic lines that had been maintained in the laboratory by selfing for several generations, an estimation of heterozygosity in the natural populations was not possible; however, this will be an interesting area for future study. It is possible that the

maintenance of isogenic lines through selfing may have influenced population genetic patterns, through the biased loss of certain alleles in heterozygotes owing to selection against deleterious alleles at linked loci, although such deleterious alleles are likely to be rare in natural populations. The concordance between population genetic patterns at mt loci, which would be unaffected by this bias, and those at autosomal loci indicates that the major conclusions of this study are robust.

Self-fertilization is also likely to be responsible for fine-scale population structure seen at the level of individual hosts; our sampling of multiple strains from individual beetles revealed each beetle to be an 'island', harbouring closely related but non-identical populations of individuals that are strongly differentiated from those on other hosts. This is likely to be the result of repeated generations of self-fertilization, resulting in differential fixation of variants introduced by mutation or migration with occasional out-crossing, hence increasing differentiation between localized populations. The observed fine-scale population structure provides an additional explanation for the high genetic diversity observed within sampling localities; each location comprises sets of highly differentiated units with variation maintained by repeated selfing and reduced effective recombination. With the exception of *Oryctes borbonicus*, which has an infestation rate of over 90% (Herrmann *et al.* 2010), the majority of beetle species sampled for this study show lower than 20% infestation rates at most sampling sites (unpublished data). Despite this, nematode-positive beetles tend to be infested with multiple strains, suggesting either that long-term persistence and reproduction take place on the individual beetle host or that populations of *P. pacificus* aggregate and infest beetle hosts together. Because *P. pacificus* strains maintain an arrested 'dauer' state of development on the beetle host, ceasing reproduction until the host death, the latter explanation seems more likely.

The ability of *P. pacificus* to tolerate a wide variety of environments and its co-dispersal with a variety of beetle species make it a good model species for investigating the complex effects of environmental, ecological and geological factors on local adaptation and genotypic evolution. Here, we demonstrate how differential colonization and host associations have achieved a richly diverse pattern in *P. pacificus*, of isolated populations potentially subjected to independent selective effects. In concert, elements of life history strategy of *P. pacificus*, such as the hermaphroditic mating system with occasional out-crossing, have resulted in a system whereby the effects of geological and climate barriers associated with volcanic island systems can become exacerbated, further contributing to significant genetic differentiation

among isolated nematode populations. Combining an island system with the highly complex nature of genetic partitioning in *P. pacificus*, coupled with the expansive genetic, genomic and transgenic toolkit (Hong & Sommer 2006) available for this species, provides us with a unique opportunity. On La Réunion Island, we can unite ecological, genetic, developmental and population-based approaches to begin to disentangle the intricacies of the evolutionary history of *P. pacificus*.

Acknowledgements

We would like to thank the Office National des Forêts and the Parc National de La Réunion, for supplying collection permits. We also thank Sophie Gasnier, Jacques Poussereau and Dominique Martiré (Insectarium de La Réunion) for help with beetle collection. Thanks to Ruxandra Molnar, for her assistance with primer design and microsatellite genotyping. Finally, we thank Max Planck Society and the Alexander von Humboldt Foundation (Research Fellowship for Postdoctoral Researchers to AM) for funding.

References

- Agapow P-M, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, **1**, 101–102.
- Anderson JL, Morran LT, Phillips PC (2010) Outcrossing and the maintenance of males within *C. elegans* populations. *Journal of Heredity*, **101**(Suppl), S62–S67.
- Benson G (1999) Tandem repeats finder: a program to analyse DNA sequences. *Nucleic Acids Research*, **27**, 573–580.
- Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **358**, 1051–1070.
- Cheke A, Hume J (2008) *Lost Land of the Dodo, an Ecological History of Mauritius, Réunion and Rodrigues*. Yale University Press, New Haven, CT.
- Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.
- Click A, Savaliya CH, Kienle S, Herrmann M, Pires-daSilva A (2009) Natural variation of outcrossing in the hermaphroditic nematode *Pristionchus pacificus*. *BMC Evolutionary Biology*, **9**, 75.
- Crandall K (1994) Interspecific cladogram estimation: accuracy at higher levels of divergence. *Systematic Biology*, **43**, 222–235.
- Crandall K, Templeton AR (1993) Empirical tests of some predictions from coalescent theory. *Genetics*, **134**, 619–633.
- Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.
- Crawford KM, Whitney KD (2010) Population genetic diversity influences colonization success. *Molecular Ecology*, **19**, 1253–1263.
- Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.
- Excoffier LGL, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, **164**, 1567–1587.
- Felsenstein J (1989) PHYLIP – phylogeny inference package (Version 3.2). *Cladistics*, **5**, 164–166.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.
- Gao H, Williamson S, Bustamante CD (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics*, **176**, 1635–1651.
- Gillot PY, Nativel P (1982) K–Ar chronology of the ultimate activity of piton des neiges volcano, reunion island, Indian ocean. *Journal of Volcanology and Geothermal Research*, **13**, 131–146.
- Gillot P, Nativel P (1989) Eruptive history of the Piton de la Fournaise volcano, Reunion Island, Indian Ocean. *Journal of Volcanology and Geothermal Research*, **36**, 53–65.
- Haber M, Schungel M, Putz A, Muller S, Hasert B, Schulenburg H (2005) Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Molecular Biology and Evolution*, **22**, 160–173.
- Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation GST, DEST and G'ST: how wrong have we been? *Molecular Ecology*, **18**, 2080–2083.
- Herrmann M, Mayer WE, Sommer RJ (2006) Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology*, **109**, 96–108.
- Herrmann M, Kienle S, Rochat J, Mayer WE, Sommer RJ (2010) Haplotype diversity of the nematode *Pristionchus pacificus* on Réunion in the Indian Ocean suggests multiple independent invasions. *Biological Journal of the Linnean Society*, **100**, 170–179.
- Hong RL, Sommer RJ (2006) *Pristionchus pacificus* – a well rounded nematode. *Bioessays*, **28**, 651–659.
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.
- Jordan MA, Snell HL, Snell HM, Jordan WC (2005) Phenotypic divergence despite high levels of gene flow in Galápagos lava lizards (*Microlophus albemarlensis*). *Molecular Ecology*, **14**, 859–867.
- Jost LOU (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.
- Kleindorfer S, Chapman TW, Winkler H, Sulloway FJ (2006) Adaptive divergence in contiguous populations of Darwins' small ground finch (*Geospiza fuliginos*). *Evolutionary Ecology Research*, **8**, 357–372.
- Lawton-Rauh A, Robichaux RH, Purugganan MD (2007) Diversity and divergence patterns in regulatory genes

- suggest differential gene flow in recently derived species of the Hawaiian silversword alliance adaptive radiation (Asteraceae). *Molecular Ecology*, **16**, 3995–4013.
- Leng L, Zhang D-X (2011) Measuring population differentiation using GST or D? A simulation study with microsatellite DNA markers under a finite island model and nonequilibrium conditions. *Molecular Ecology*, **20**, 2494–2509.
- Mantel N (1967) The detection of disease clustering and generalized regression approach. *Cancer Research*, **27**, 209–220.
- Mathys BA, Lockwood JL (2011) Contemporary morphological diversification of passerine birds introduced to the Hawaiian archipelago. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 2392–2400.
- Mayer MG, Sommer RJ (2011) Natural variation in *Pristionchus pacificus* dauer formation reveals cross-preference rather than self-preference of nematode dauer pheromones. *Proceedings of the Royal Society B: Biological Sciences*, doi: 10.1098/rspb.2010.2760
- Mayer W, Herrmann M, Sommer R (2007) Phylogeny of the nematode genus *Pristionchus* and implications for biodiversity, biogeography and the evolution of hermaphroditism. *BMC Evolutionary Biology*, **7**, 104.
- McGaughan A, Convey P, Stevens M, Chown S (2010) Metabolic rate, genetic and microclimate variation among springtail populations from sub-Antarctic Marion Island. *Polar Biology*, **33**, 909–918.
- Mila B, Warren B, Heeb P, Thebaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158.
- Morlais I, Girod R, Hunt R, Simard F, Fontenille D (2005) Population structure of *Anopheles arabiensis* on La Réunion Island, Indian Ocean. *The American Journal of Tropical Medicine and Hygiene*, **73**, 1077–1082.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- Nei M (1972) Genetic distance between populations. *The American Naturalist*, **106**, 283–292.
- Nordborg M, Hu TT, Ishino Y *et al.* (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biology*, **3**, e196.
- Paupy C, Girod R, Salvan M, Rodhain F, Failloux A-B (2001) Population structure of *Aedes albopictus* from La Reunion Island (Indian Ocean) with respect to susceptibility to a dengue virus. *Heredity*, **87**, 273–283.
- Posada D, Crandall K (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, **16**, 37–45.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.
- Rockman MV, Kruglyak L (2009) Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genetics*, **5**, e1000419.
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, **22**, 454–464.
- Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
- Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Human Press, Totowa, NJ.
- Schlager B, Wang X, Braach G, Sommer RJ (2009) Molecular cloning of a dominant roller mutant and establishment of DNA-mediated transformation in the nematode *Pristionchus pacificus*. *Genesis*, **47**, 300–304.
- Siol M, Prosperi JM, Bonnin I, Ronfort J (2008) How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity*, **100**, 517–525.
- Sommer RJ (2009) The future of evo-devo: model systems and evolutionary theory. *Nature Reviews Genetics*, **10**, 416–422.
- Strasberg D, Rouget M, Richardson D, Baret S, Dupont J, Cowling R (2005) An assessment of habitat diversity and transformation on La Réunion Island (Mascarene Islands, Indian Ocean) as a basis for identifying broad-scale conservation priorities. *Biodiversity and Conservation*, **14**, 3015–3032.
- Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498–2504.
- Tatarenkov A, Gao H, Mackiewicz M, Taylor DS, Turner BJ, Avise JC (2007) Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (*Kryptolebias marmoratus*). *Molecular Ecology*, **16**, 2701–2711.
- Templeton AR, Crandall K, 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.
- Vercambre B, Goebel O, Morel G, Robert P, Guillon M (1991) Programme de lutte biologique contre *Hoplochelus marginalis* (Coleoptera : Melolonthidae) nouveau ravageur des cultures à l'île de la Réunion. In: *Proceedings Rencontres Caraïbes en Lutte Biologique* (eds Parvis C, Kermarec A), . Guadeloupe, 5–7 Novembre 1990.
- Whittaker RJ, Fernández-Palacios JM (2007) *Island Biogeography: Ecology, Evolution and Conservation*. Oxford University Press Inc., New York, NY.
- Zauner H, Sommer RJ (2007) Evolution of robustness in the signaling network of *Pristionchus vulva* development. *Proceedings of the National Academy*, **104**, 10086–10091.

K.M. and A.M. are population geneticists interested in combining molecular and ecological approaches to investigate evolutionary history and adaptation in natural populations. L.V. is a population geneticist interested in reproductive modes and dispersal. M.H. is interested in nematode biogeography, phylogeny and taxonomy, and in insect-nematode interactions. J.R. is interested in the population ecology and conservation of insects and other terrestrial arthropods. H.W. and G.B. specialise in nematology and molecular genetic techniques. R.J.S. is interested in the diversity of biological form and the developmental and evolutionary mechanisms involved in generating this diversity.

Data accessibility

All GPS, microsatellite genotype and mitochondrial sequence data have been deposited in Dryad, with the accession: doi:10.5061/dryad.ns1stov2.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Neighbour-net network constructed using microsatellite data, in SplitsTree 4.11.3 (Huson & Bryant 2006).

Table S1 Sampling information for strains genotyped for microsatellite markers, and sequenced at the mitochondrial loci *ND4* and *ND6L*.

Table S2 Microsatellite primer information.

Table S3 Population statistics and genetic characteristics of sampled locations on La Réunion Island for *Pristionchus pacificus* mtDNA (*ND6* and *ND4L*) sequences.

Table S4 Pairwise R_{ST} and F_{ST} values between data partitions, based on: (i) 20 nuclear microsatellite markers, and (ii) 736 bp of mitochondrial sequence (*ND6* and *ND4L*), respectively.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.