PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2007) **274**, 2669–2677
doi:10.1098/rspb.2007.0877
Published online 28 August 2007

Disentangling hybridization and host colonization in parasitic roundworms of humans and pigs

Charles D. Criscione[1,*], Joel D. Anderson[2], Dan Sudimack[1], Weidong Peng[3], Bharat Jha[4], Sarah Williams-Blangero[1] and Timothy J. C. Anderson[1]

[1]*Department of Genetics, Southwest Foundation for Biomedical Research, PO Box 760549, San Antonio, TX 78245, USA*
[2]*Perry R. Bass Marine Fisheries Research Station, Coastal Fisheries Division, Texas Parks and Wildlife Department, Palacios, TX 77465, USA*
[3]*Jiangxi Medical Science Research Institute, Nanchang University, 461 Ba Yi Road, Nanchang, Jiangxi 330006, People's Republic of China*
[4]*Tribhuvan University, Institute of Medicine, Majarajgung, PO Box 1524, Kathmandu, Nepal*

Knowledge of cross-transmission and hybridization between parasites of humans and reservoir hosts is critical for understanding the evolution of the parasite and for implementing control programmes. There is now a consensus that populations of pig and human *Ascaris* (roundworms) show significant genetic subdivision. However, it is unclear whether this has resulted from a single or multiple host shift(s). Furthermore, previous molecular data have not been sufficient to determine whether sympatric populations of human and pig *Ascaris* can exchange genes. To disentangle patterns of host colonization and hybridization, we used 23 microsatellite loci to conduct Bayesian clustering analyses of individual worms collected from pigs and humans. We observed strong differentiation between populations which was primarily driven by geography, with secondary differentiation resulting from host affiliation within locations. This pattern is consistent with multiple host colonization events. However, there is low support for the short internal branches of the dendrograms. In part, the relationships among clusters may result from current hybridization among sympatric human and pig roundworms. Indeed, congruence in three Bayesian methods indicated that 4 and 7% of roundworms sampled from Guatemala and China, respectively, were hybrids. These results indicate that there is contemporary cross-transmission between populations of human and pig *Ascaris*.

Keywords: *Ascaris* spp.; molecular epidemiology; hybridization; microsatellites; nematode; parasite

1. INTRODUCTION

Hybridization is recognized as a significant process in the evolution of free-living organisms (Arnold 2006), but the role and extent that this process plays in the evolution of parasitic organisms remains insufficiently studied (Arnold 2004). As with other organisms, hybridization among parasite lineages could result in the introgression of novel genes (e.g. host infectivity or drug-resistant genes), promote divergence via reinforcement, lead to homogenization across the genomes of the interbreeding populations or promote rapid adaptive diversification (Barton 2001; Olden *et al.* 2004; Seehausen 2004). Genetic exchange between parasite lineages may also lead to the evolution of more virulent genotypes or genotypes with reduced host specificity (Arnold 2004). Thus, knowledge of hybridization is critical in studies of parasite evolution and epidemiology.

The long-standing question of whether *Ascaris* roundworms of humans and pigs can cross-transmit among host species or even interbreed (Macko & Dubinsky 1997; Anderson 2001; Crompton 2001) is a clear example of where hybridization studies could benefit epidemiology. Over one billion people worldwide are infected with *Ascaris*, which results in major health and economic burdens in developing countries (O'Lorcain & Holland 2000; WHO 2002). Furthermore, infections of *Ascaris* in pigs are a source of substantial economic loss (Nejsum *et al.* 2005*a* and references therein). Therefore, data on the potential for cross-transmission or hybridization between *Ascaris* of humans and pigs are critical not only for understanding the evolution of the parasite, but also for implementing effective control programmes.

Molecular markers and population genetics analyses are useful tools for elucidating parasite transmission dynamics (Criscione *et al.* 2005; de Meeus *et al.* 2007) and have begun to shed light on *Ascaris* transmission between humans and pigs. In China and Guatemala, where there is endemic transmission of both human and pig *Ascaris*, molecular studies have shown significant genetic subdivision between sympatric populations of pig and human *Ascaris* (Anderson *et al.* 1993; Anderson & Jaenike 1997; Peng *et al.* 1998, 2003, 2005). These studies conclude that there is little contemporary cross-transmission between host species. However, the studies in China and Guatemala were conducted in isolation, with different molecular markers. Thus, it is not known whether the observed genetic subdivision is the result of a single host shift between pigs and humans or multiple host colonizations in allopatry

*Author and address for correspondence: Address starting from August 2008: Department of Biology, 3258 TAMU, Texas A&M University, College Station, TX 77843, USA. (ccriscio@sfbrgenetics.org).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2007.0877 or via http://www.journals.royalsoc.ac.uk.

Received 29 June 2007
Accepted 8 August 2007

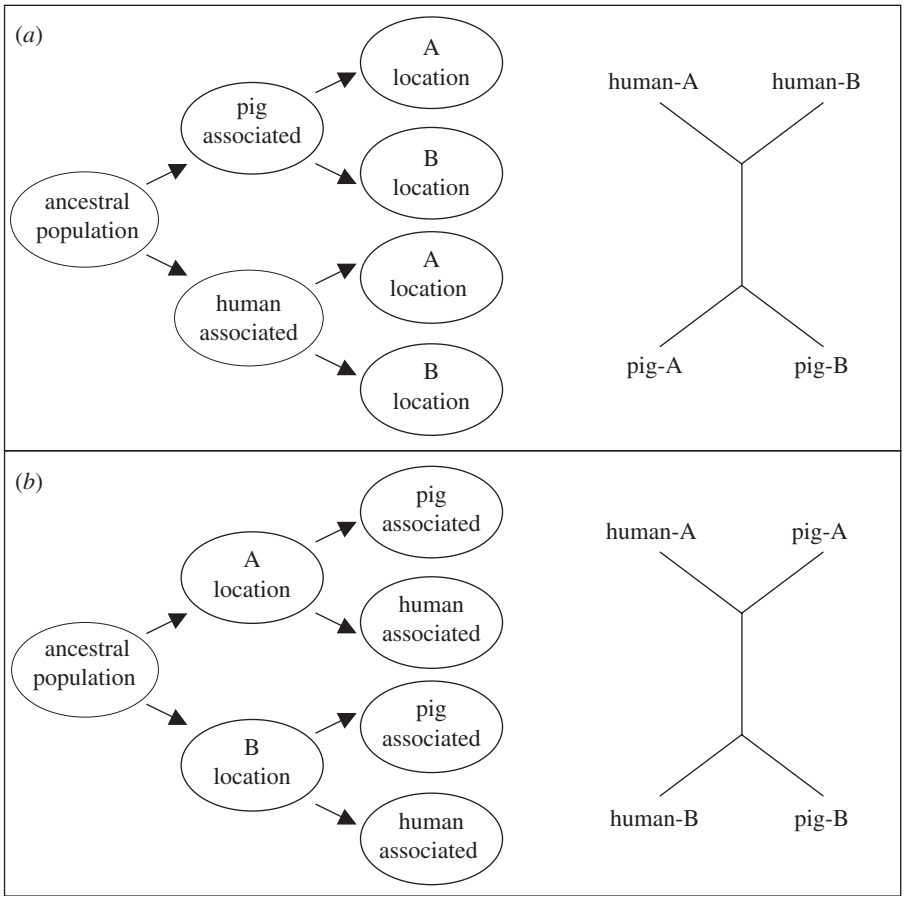

Figure 1. Models of geographical or host colonization. (*a*) Genetic isolation first occurs among hosts and then among geographical locations. This model implies that there was a single host shift and leads to the clustering of host-derived populations (dendrogram shown on the right). (*b*) The ancestral population is first subdivided by geography. Following geographical isolation, host-associated populations form. This model would result in a geographical clustering pattern (dendrogram shown on the right). These models represent simplified patterns of host or geographical colonization. More complex branching patterns are possible as more locations and host-derived populations are added. However, any deviation from a pattern that strictly clusters host-derived populations would suggest multiple host shifts.

followed by subsequent host isolation (figure 1). Observations of human or chimpanzee infections with pig-derived parasites in Denmark (Nejsum *et al.* 2005*b*, 2006) and USA (Anderson 1995; neither country has endemic human transmission) raise the possibility of frequent host colonization events. Furthermore, the existence of shared mitochondrial or nuclear polymorphisms in these studies suggests that there is potential for hybridization between host-associated populations (Anderson *et al.* 1993; Anderson & Jaenike 1997; Peng *et al.* 1998, 2003, 2005). Unfortunately, the above molecular ecology studies used a single marker (i.e. the internal transcribed spacer of the ribosomal or mitochondrial DNA) or nuclear makers with low polymorphism (Anderson & Jaenike 1997), and therefore lacked power to disentangle the processes of incomplete lineage sorting, historical introgression or contemporary hybridization as causes of the shared polymorphisms. Thus, whether human and pig populations of *Ascaris* can hybridize remains an open question.

Here we used 23 microsatellite loci and Bayesian clustering analyses of individual worms collected from pigs (China and Guatemala) and humans (China, Guatemala and Nepal) to examine the patterns of historical host colonization and contemporary hybridization. We observed that host-associated populations clustered together within geographical regions. This pattern could be caused by either multiple host colonization events or homogenization due to recurrent hybridization. Consistent with the latter explanation, we detected strong evidence for both cross-infection and hybridization in sympatric populations of human and pig roundworms. We discuss the implications that host colonization and hybridization will have for the evolution and epidemiology of *Ascaris* in humans and pigs.

2. MATERIAL AND METHODS

(a) *Sampling*

A total of 129 roundworms (male or female) were sampled from China, Guatemala and Nepal. In China, 21 human- and 21 pig-derived worms (from 15 and 17 individual hosts, respectively) were collected from Changjiang county (Hainan province) in 1999 (see Peng *et al.* (2003) for collection details). Worms from Guatemala were sampled from the villages of Chiamal (21 human-derived from 4 hosts) or Santa Cruz Naranjo (13 human- and 11 pig-derived each from a different host) in 1991 (collection details given in Anderson *et al.* 1993, 1995). Worms collected from the Jiri, Nepal villages of Upper Sikri, the Ahale hamlet (21 human-derived from 13 hosts) and Kharayoban (21 human-derived from 13 hosts) were sampled during 2002–2003. All worms from Nepal were obtained using the methods described by Williams-Blangero *et al.* (1999, 2002). Even though we were unable to obtain pig-derived worms from Nepal, we included

the human-derived worms from Nepal to determine whether the clustering analyses (discussed in §2*c*) were robust to additional samples from another geographical location. Qualitatively, the results are similar if we remove the samples collected from Nepal.

(b) *Genotyping*

All worms were genotyped at 23 autosomal microsatellite loci (GenBank accession numbers: DQ988845, DQ988847–DQ988849, DQ988853, DQ988855–DQ988857, DQ988859, DQ988860, DQ988862–DQ988867, DQ988869–DQ988872, CB101754, BQ380931, BQ835581). Methods for DNA extraction, polymerase chain reaction, genotyping and assessment of genotyping error were previously given by Criscione *et al.* (2007). There were complete multilocus genotypes for all worms except for five that had missing data for a single locus.

(c) *Broad-scale analyses of host colonization patterns*

Three methods were used to examine broad-scale relationships and genetic structure among the two pig- and five human-derived population samples. First, we calculated pairwise F_{ST} (Weir & Cockerham 1984) among populations and tested (1000 permutations of individuals) for differentiation between pairs of populations using FSTAT v. 2.9.3 (Goudet 1995). Sequential Bonferroni correction was applied. F_{ST} values were standardized according to Meirmans (2006). Second, we constructed a neighbour-joining tree based on the Cavalli-Sforza & Edwards (1967) chord distance, the distance measure that provides the best estimation of tree topology (Takezaki & Nei 1996), using PHYLIP v. 3.66 (Felsenstein 2005). Stability of the topology was assessed by 1000 bootstraps over loci. The above traditional methods, however, rely on *a priori* delimitation of populations (i.e. the two pig- and five human-derived population samples). Thus, as a third means of assessing population structure, we used Bayesian clustering methods of individuals to identify the populations for comparison. Two software programs were used for the clustering analyses: STRUCTURE v. 2.1 (Falush *et al.* 2003) and BAPS v. 4.14 (Corander & Marttinen 2006; Corander *et al.* 2006). In STRUCTURE, we ran 20 replications of k populations ($k=1$–14) with a Markov chain Monte Carlo (MCMC) burn-in of 50 000 steps and 100 000 steps after burn-in. The admixture and correlated allele frequency models were used. We determined the optimal number of clusters (k) using the posterior probability (pp) of the data for a given k (i.e. ln $P(D)$ in the STRUCTURE output). To examine the relationships among the clusters from the optimal k, we created a neighbour-joining tree as described above by using the posterior allele frequencies for each cluster (given in the STRUCTURE output). In BAPS, we ran 20 replicates of k ranging from 2 to 14 and then constructed a neighbour-joining tree with the Kullback–Leibler divergence matrix provided as output with BAPS. This matrix can be used as a measure of relative genetic distance between the BAPS-identified clusters (BAPS v. 4.14 manual distributed with the program).

(d) *Analyses for detecting hybrids within sympatric populations*

To search for hybrids within sympatric populations of human and pig parasites, we used three software programs (STRUCTURE, BAPS and NEWHYBRIDS v. 1.1; Anderson & Thompson 2002) that implement model-based Bayesian methods to infer hybridization. The methods used in these programs differ in approach and statistical treatment of variables (see Anderson & Thompson 2002; Corander & Marttinen 2006; Corander *et al.* 2006). For example, a value (Q) of genome admixture (i.e. the proportion of an individual's genome originating from the parental populations) is provided in STRUCTURE and BAPS, whereas a pp of being a pure-bred or a hybrid is estimated for each individual in NEWHYBRIDS (Vähä & Primmer 2006). All three programs are similar in that they assume Hardy–Weinberg and linkage equilibrium among loci. Furthermore, these programs have two advantages for allowing one to detect hybrids. First, analyses can be conducted when no taxa-specific markers exist, as is currently the case with *Ascaris* of humans and pigs (Anderson & Thompson 2002; Vähä & Primmer 2006). Second, pure samples of parental taxa are not required (Anderson & Thompson 2002). The latter is also a problem with *Ascaris* as the collection of a pure host-derived sample would probably be confounded with geographical isolation.

We used sympatric samples of roundworms from humans and pigs in two locations: Santa Cruz Naranjo, Guatemala and Changjiang county, Hainan province, China. Analyses of hybridization were conducted separately for each location. We were conservative in classifying individuals as hybrids in order to prevent the overestimation of cross-transmission between human and pig *Ascaris*. For example, since STRUCTURE, BAPS and NEWHYBRIDS use different model-based Bayesian methods for the detection of hybrids, we looked for congruence of results among the three programs as suggested by Vähä & Primmer (2006).

We used uniform priors in NEWHYBRIDS with a 100 000 MCMC burn-in and 1 000 000 steps for sampling. We ran three independent runs to test the stability of the analysis. A pp value of 50% was used as a threshold for assigning an individual to either a human or pig pure-bred category or the hybrid category. The hybrid category was the sum of the probabilities for the categories of F_1, F_2 and backcross of F_1 to parental (Vähä & Primmer 2006). In STRUCTURE, k was set to 2 and was run using the admixture and independent allele frequency models (100 000 MCMC burn-in and 1 000 000 steps for sampling). The STRUCTURE analysis was run five times to examine stability. An individual was classified as a hybrid if its Q-values were between 0.2 and 0.8. This threshold was used to give an optimal balance between the efficiency and accuracy for categorizing individuals as parental or hybrid (Vähä & Primmer 2006).

To further evaluate the robustness of our results, we used simulations to test whether putative hybrids could be the result of random chance alone (i.e. be false positives). Our simulations are similar to those conducted by Paetkau *et al.* (2004). Allele frequencies from the set of human- or pig-derived worms were used to generate 100 datasets of sample sizes equivalent to the original dataset (i.e. 21 human- and 21 pig-derived for the China data and 13 human- and 11 pig-derived for the Guatemala data). Each dataset was run in STRUCTURE as described above, but using 30 000 MCMC burn-in and 50 000 steps for sampling. We used WHICHLOCI v. 1.0 to create the datasets (Banks *et al.* 2003). Populations were simulated with random mating within and no mating between host-associated populations. Putative hybrid genotypes were kept within their original host-sampled population and thus included in the allele frequencies used for simulation. By simulating populations in this manner, analyses of the simulated datasets should be biased to find

Table 1. Multilocus estimates of F_{ST} between pairs of populations. (All comparisons except for one (shown in italics) was significant ($p<0.001$) after Bonferroni correction. Standardized F_{ST} is below and raw F_{ST} is above the diagonal. In the population abbreviations, the first letter H or P indicates host species. The following two letters show geographical location: CH, China Hainan; GC, Guatemala Chiamal; GS, Guatemala Santa Cruz Naranjo; NA, Nepal Ahale; NK, Nepal Kharayoban.)

	HCH	HGC	HGS	HNA	HNK	PCH	PGS
HCH	—	0.087	0.095	0.135	0.111	0.043	0.042
HGC	0.315	—	*0.009*	0.187	0.172	0.120	0.091
HGS	0.339	*0.027*	—	0.187	0.179	0.134	0.094
HNA	0.502	0.580	0.566	—	0.017	0.159	0.164
HNK	0.438	0.561	0.573	0.056	—	0.138	0.143
PCH	0.186	0.425	0.466	0.579	0.534	—	0.058
PGS	0.173	0.302	0.307	0.565	0.526	0.238	—

more false positives. Therefore, these simulations should provide a more conservative determination of hybrid status for the individuals in the original dataset.

The settings we used for both NEWHYBRIDS and STRUCTURE provide simultaneous estimates of cluster origin and admixture of each individual. BAPS, on the other hand, first estimates cluster membership for each individual. We set $k=2$ and ran 50 replicates to determine cluster membership. BAPS then performs an analysis to determine each individual's admixture and provides a simulation analysis to test the significance of the admixture. Following recommendations in the BAPS manual, we used 200 iterations to estimate the admixture coefficients and simulated 200 reference individuals with 200 iterations each for significance testing. In all programs, no prior information on an individual worm's host sampling origin was provided.

3. RESULTS

(a) *Relationships among human- and pig-derived roundworms*

Pairwise genetic differentiation was significant ($p<0.001$ after Bonferroni correction) among all comparisons except for the two human-derived populations in Guatemala (table 1). Pairwise F_{ST} estimates did not show any clear host or geographical pattern of affiliation among samples (table 1). For example, F_{ST} was similar between the human-derived sample from China and the pig-derived samples from China and Guatemala (0.186 and 0.173, respectively), whereas F_{ST} between the two pig-derived populations was 0.238. The human-derived samples from Guatemala were as differentiated from the human-derived China population as they were from the pig-derived Guatemala population ($F_{ST}=0.302$–0.339). F_{ST} comparisons with the human-derived samples from Nepal were the greatest and ranged from 0.438 to 0.579.

In contrast, all clustering analyses showed that differentiation among worms was primarily driven by geography, with secondary differentiation resulting from host affiliation within locations (figure 2). The neighbour-joining tree based on *a priori* population designations shows that human- and pig-derived worms from Guatemala cluster and likewise for the human- and pig-derived worms from China. However, there is low bootstrap support for these internal branches (figure 2*a*).

The STRUCTURE analysis provided an optimal $k=5$. As there was a clear peak for the ln $P(D)$ at $k=5$ (electronic supplementary material, figure 1), there was no need to use alternative criteria to determine the optimal k (e.g. Evanno *et al.* 2005). All individual roundworms showed strong membership to their respective genetic cluster (i.e. all had a Q-value >0.8 except for three, which had a Q-value >0.7). All human-derived worms from Guatemala formed a single cluster and likewise for the human-derived worms from Nepal. The remaining three clusters were exactly the same as the original sampling units (e.g. the 21 pig-derived worms from China formed a cluster; figure 2*b*). Relationships among these five clusters (figure 2*b*) were concordant with the dendrogram in figure 2*a*, but, again, bootstrap support was low for the internal branches.

The optimal k from BAPS was 6, but one cluster had a single human-derived worm from China (figure 2*c*). The individual clustering and the relationships among clusters were largely concordant with the STRUCTURE analysis with the exception that four human-derived worms from China belonged to the same cluster as the pig-derived worms from China (figure 2*c*). Assessment of branch support is not possible in this analysis; however, the internal branches were short. This result is consistent with the dendrograms shown in figure 2*a*,*b*.

(b) *Detection of hybrids*

Table 2 and figure 3 show the individual roundworms from Guatemala or China, which were identified by one or more programs as being hybrids. The multiple runs in NEWHYBRIDS and STRUCTURE produced consistent results for both datasets. In Guatemala, one pig-derived worm (14-PGS) was classified as a hybrid by all three programs, including the statistical significance detected by BAPS (table 2; figure 3). Furthermore, our simulations with STRUCTURE showed only 2 out of 100 datasets, where a single individual had a Q-value between 0.2 and 0.8. This result indicates that the individual in the observed dataset is probably not a false positive ($p=0.02$). We chose to assess the probability of false positives over the number of simulated datasets versus the number of simulated individuals (see Berthier *et al.* 2006) to provide a more conservative classification of putative hybrids.

In China, five human-derived roundworms each from a different host individual were classified by one or more of the programs as being a hybrid (table 2; figure 3). Interestingly, four of these worms (3-HCH, 8-HCH, 11-HCH and 19-HCH) were the same that BAPS clustered with the pig-derived worms in the broad-scale analysis (figure 2*c*). Furthermore, worm 18-HCH was the single individual that BAPS placed in its own cluster (figure 2*c*). Inconsistent results among the programs were found for individuals 3-HCH and 18-HCH (figure 3), thus their status as hybrids

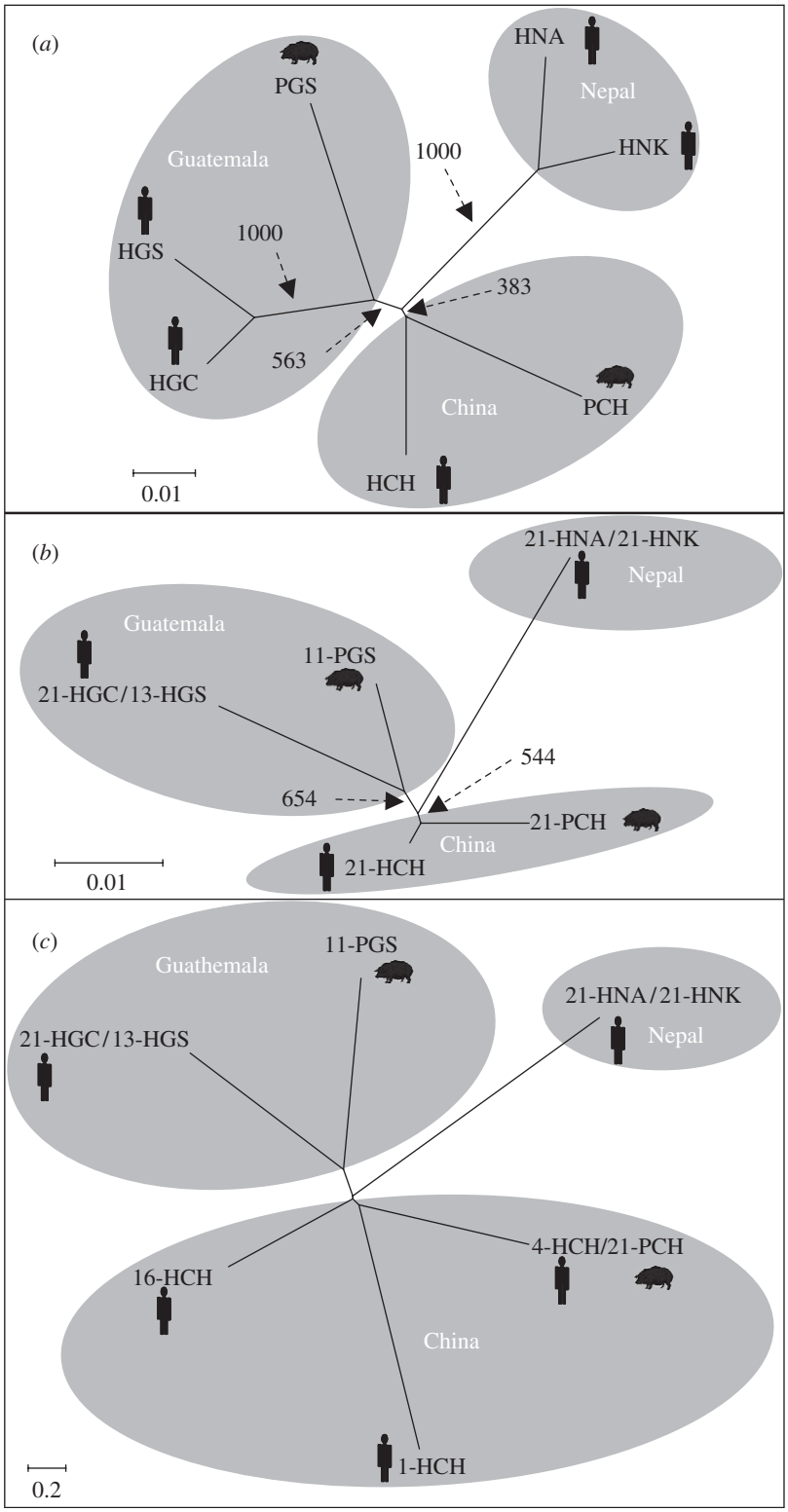

Figure 2. Neighbour-joining dendrograms showing the relationships among pig and human *Ascaris* samples from different geographical locations. (*a*) Clustering of populations based on *a priori* delimitation of populations and the Cavalli-Sforza & Edwards (1967) chord distance. (*b*) STRUCTURE-based clustering and relationships among clusters that were estimated from the Cavalli-Sforza & Edwards (1967) chord distance of the posterior allele frequencies. (*c*) BAPS-based clustering and relationships among clusters that were based on the Kullback–Leibler divergence matrix provided as output with BAPS. Population or sample abbreviations are the same as given in table 1. The number of individuals in each cluster is shown before the abbreviations in (*b*,*c*). *Note*. All three trees show that the primary determinant of clustering among populations is driven by geography, with secondary differentiation resulting from host affiliation within locations. However, internal branches are short and have low bootstrap support (shown with arrows in (*a*,*b*)).

is questionable as they may represent false positives. Indeed, the STRUCTURE simulations demonstrated that 17 out of 100 datasets produced one or more false positives (only one dataset had two false positives). Thus, there was at least a 17% chance of producing a single false positive and only a 1% chance of producing two false positives. On the other

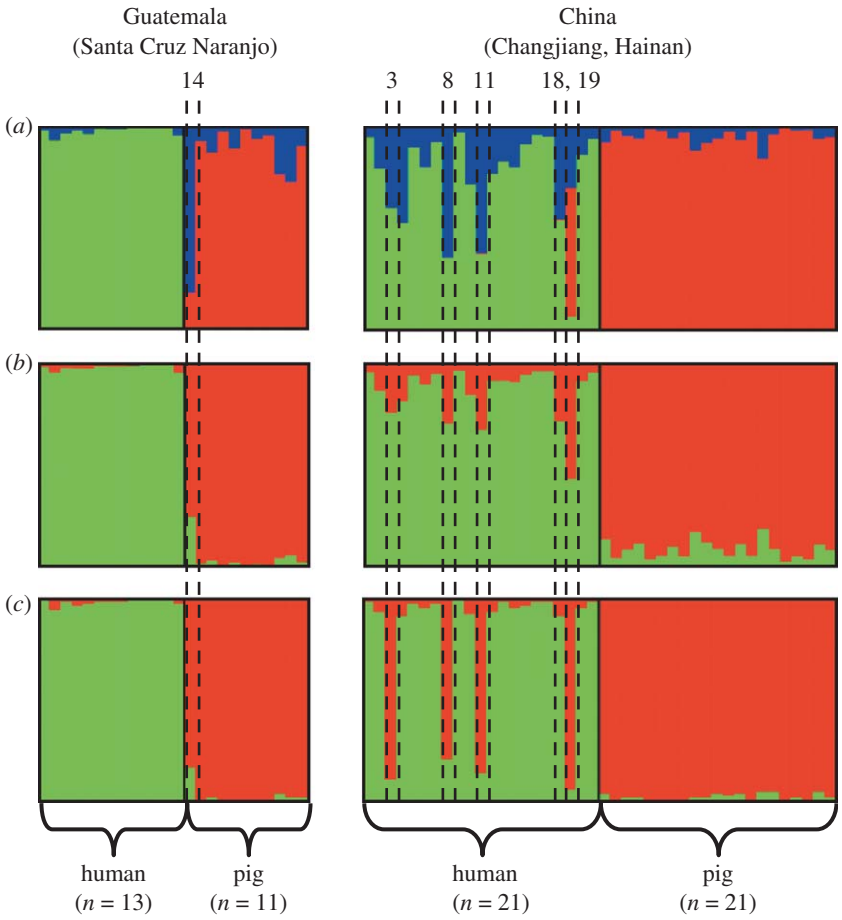

Figure 3. Hybrid analysis results within Santa Cruz Naranjo, Guatemala and Changjiang county, Hainan province, China. The results from (*a*) NEWHYBRIDS, (*b*) STRUCTURE and (*c*) BAPS are presented in three rows. Graphs were created with DISTRUCT (Rosenberg 2004). One vertical bar represents one individual. The bottom of the figure shows the number of worms that were sampled from each host species. (*a*) For NEWHYBRIDS, the *y*-axis is the posterior probability (pp) of a roundworm being a pure-bred human (green), a pure-bred pig (red) or a hybrid (blue). (*b*,*c*) For STRUCTURE and BAPS, the *y*-axis is the *Q*-value, the proportion of an individual's genome belonging to the human cluster (green) or pig cluster (red). Note that there is a good delineation between the human- and pig-derived samples, thus indicating a clear genetic subdivision among the host-associated populations in sympatry. However, the six numbered individuals at the top of the graph (highlighted with the dashed lines) were identified by one or more of the programs as being a putative hybrid (table 2).

Table 2. Individual roundworms classified as hybrids by STRUCTURE (S), BAPS (B) or NEWHYBRIDS (N). (*Q*-values, the proportion of an individual's genome belonging to a cluster, are shown for STRUCTURE and BAPS. A 90% posterior interval (PI) is shown for the *Q*-values from STRUCTURE. The *p*-value for the test of significant admixture is shown for the BAPS results. The NEWHYBRIDS results are illustrated as the pp of belonging to either a pure-bred (pig or human) or hybrid category. Threshold values or significance tests in each program, which were used to classify individuals, are described in the text. A conservative conclusion of hybrid status was based on the strict consensus of the three programs and simulations described in the text. Abbreviations for individuals are the same as given in table legend 1.)

individual	STRUCTURE *Q*-pig (90% PI) *Q*-human (90% PI)	BAPS *Q*-pig (*p*-value) *Q*-human	NEWHYBRIDS pp—category	method classification	conclusion
14-PGS	0.752 (0.441,1.000) 0.248 (0.000,0.559)	0.83 ($p=0.05$) 0.17	82%—hybrid	all methods agree hybrid	hybrid
3-HCH	0.239 (0.000,0.561) 0.761 (0.439,1.000)	0.89 ($p=0.195$) 0.11	60%—human	hybrid (S), pure-bred pig (B), pure-bred human (N)	undetermined
8-HCH	0.290 (0.009,0.564) 0.710 (0.437,0.991)	0.79 ($p=0.045$) 0.21	65%—hybrid	all methods agree hybrid	hybrid
11-HCH	0.326 (0.007,0.655) 0.674 (0.345,0.993)	0.86 ($p=0.11$) 0.14	62%—hybrid	hybrid (S, N), pure-bred pig (B)	hybrid
18-HCH	0.279 (0.000,0.613) 0.721 (0.387,1.000)	0.08 ($p=0.36$) 0.92	54%—human	hybrid (S), pure-bred human (B, N)	undetermined
19-HCH	0.567 (0.013,1.000) 0.433 (0.000,0.987)	0.94 ($p=0.395$) 0.06	64%—pig	hybrid (S), pure-bred pig (B, N)	hybrid

hand, there was complete agreement among the programs in whether worms 8-HCH, 11-HCH and 19-HCH were hybrids. We note that BAPS identified worms 11-HCH and 19-HCH as being pure-bred pig and that NEWHYBRIDS also classified 19-HCH as a pure-bred pig worm. A pure-bred pig classification of worms sampled from humans could potentially be described as first-generation migrants. However, taking the results of all three programs into consideration, we suggest that these worms are more likely to be hybrids than migrants. Furthermore, the conclusion of cross-transmission between human and pig hosts is still reached whether these worms are migrants or hybrids.

4. DISCUSSION

(a) *Host colonization*

We found no evidence for a single host shift between pigs and humans in the broad-scale analyses of host colonization (figure 1*a*). Rather, the population dendrograms (figure 2) indicate that sympatric populations (figure 2*a*) or individuals (figure 2*b*,*c*) of human- and pig-derived *Ascaris* cluster together within geographical locations. The three different methods of clustering yielded nearly identical results (figure 2). The pairwise F_{ST} estimates also do not show a clear human–pig delineation, as the human-derived population from China was more similar to the two pig-derived populations than it was to the human-derived populations from Guatemala (table 1). Taken at face value, these results support the model in figure 1*b* where host-associated populations emerged after geographical isolation. Thus, one would infer that multiple host colonization events have occurred in the evolutionary history of *Ascaris*. In conjunction with previous reports of mature pig-derived parasites (i.e. gravid females) in humans from areas of non-endemic human transmission (Anderson 1995; Nejsum *et al.* 2005*b*, 2006), multiple host colonization events seem plausible. However, there is low bootstrap support for the internal branches (figure 2*a*,*b*). Thus, we realize that there are three other potential explanations for the observed relationships. (i) The high mutation rates of microsatellites may lead to low bootstrap support (see Irion *et al.* 2003 and references therein). (ii) The temporal spacing of the geographical samples may have introduced some genetic heterogeneity in the dataset due to drift. (iii) Contemporary hybridization may partially homogenize sympatric populations of human and pig *Ascaris*, thus leading to the patterns observed in figure 2. Although we cannot rule out explanations (i) and (ii), we did detect hybrids in the datasets from both China and Guatemala (see §4*b*). Thus, low levels of gene flow between sympatric populations of human- and pig-associated *Ascaris* may in part explain the results in figure 2. Shared mitochondrial haplotypes (identical at 43 polymorphic sites) and rDNA sequences (i.e. the internal transcribed spacer subunit 1) between sympatric host-associated populations are also consistent with low levels of introgression (Anderson *et al.* 1993; Anderson 2001; Peng *et al.* 2003). Therefore, we cannot definitively conclude that the dendrograms in figure 2 represent the 'true' relationships among the samples. Additional molecular studies are warranted to more accurately determine the colonization history of human and pig populations of *Ascaris*.

(b) *Hybridization*

Our results shed new light on the potential for cross-transmission and interbreeding between sympatric human- and pig-associated *Ascaris* populations. It is clear from previous studies in areas of non-endemic human transmission that pig-derived worms can cross-transmit to humans (Anderson 1995; Nejsum *et al.* 2005*b*). However, in areas with endemic transmission in both pigs and humans, molecular datasets have lacked power (e.g. used a single marker such as mitochondrial DNA) to accurately determine whether there is cross-transmission and/or interbreeding between human and pig *Ascaris* (Anderson *et al.* 1993; Anderson & Jaenike 1997; Peng *et al.* 1998, 2003, 2005; Anderson 2001). Using polymorphic, multilocus genotypes and a conservative approach to identify putative hybrids, we find evidence for hybridization in both Guatemala and China (table 2; figure 3). The proportion of hybrids in the samples from Guatemala and China was 4% (1 out of 24) and 7% (3 out of 42), respectively. These results indicate that there must have been contemporary interbreeding and thus, necessarily recent cross-transmission, between sympatric human and pig *Ascaris*, as the methods we employed can only detect hybrids going back two generations (Anderson & Thompson 2002). It is noteworthy that all identified hybrids in China (table 2) were sampled from human hosts. However, larger sample sizes are needed to determine whether cross-transmission and interbreeding are unidirectional from pig to human hosts.

Our analyses and previous studies that used mitochondrial DNA markers in sympatric populations found significant genetic structure among human- and pig-associated populations (Anderson *et al.* 1993; Peng *et al.* 2005). Given the evidence of interbreeding between human and pig roundworm populations, it is interesting that there is significant genetic subdivision between the host-associated populations (table 1). There are two factors that may act alone or in concert to increase subdivision between host-associated populations (see McCoy 2003). (i) Ecological separation between hosts inhibits cross-transmission or (ii) host-selective factors result in decreased hybrid fitness (F_1 or subsequent hybrid generations). Since we find 4–7% of hybrids in our small samples, parasite dispersal does not appear to be extremely limited between hosts. If the proportion of hybrid individuals reflects gene flow between host-associated populations, we might expect that sympatric populations of *Ascaris* in humans and pigs would become genetically indistinguishable. However, the significant genetic structure indicates that complete homogenization is not the case, which suggests that post-reproductive barriers limit effective gene flow. Thus, ecological separation among hosts is unlikely the sole cause of genetic subdivision between sympatric human and pig populations of *Ascaris*. Indeed, recent work by Peng *et al.* (2006) suggests that host selection may play a role in the genetic subdivision between sympatric populations of human and pig *Ascaris*. Experimental infections of pigs with eggs of human- and pig-derived *Ascaris* showed that eggs of human-derived *Ascaris* had a very low success rate of establishing in pigs relative to pig-derived eggs (Peng *et al.* 2006). However, larger sample sizes and additional experimental infections are needed to address

the role that host selection plays in maintaining genetic subdivision between sympatric populations of human and pig *Ascaris*.

(c) *Epidemiological implications*

A major goal of molecular epidemiology is to determine whether there is cross-transmission of parasites between humans and other potential hosts (Criscione *et al*. 2005). Such knowledge allows parasite control measures to be designed for maximal impact. Our results are consistent with previous studies that show significant genetic structure between sympatric human- and pig-associated populations of *Ascaris* (Anderson *et al*. 1993; Anderson & Jaenike 1997; Peng *et al*. 1998, 2003, 2005; Anderson 2001). In a review of molecular epidemiological studies on *Ascaris*, Anderson (2001) concluded that because the current genetic data indicated rare or absent cross-transmission in endemic areas, control programmes could focus on human infection and be less concerned with infections in pigs. However, Anderson (2001) warned that rare hybridization events can have serious epidemiological implications. For example, drug-resistant alleles that evolve in one host-associated population could spread to the other host-associated population. The introgression of an allele with a selective advantage can happen when hybrids themselves are at a selective disadvantage (Barton 2001). Furthermore, hybridization itself may produce new combinations of parasite genotypes that increase host range via host immune evasion or parasite virulence (Arnold 2004). Such recombined parasite genotypes may escape from hybrid zones (Barton 2001).

The significance of our results is that we find clear evidence that hybridization occurs between sympatric populations of human- and pig-associated *Ascaris*. Thus, while short-term transmission between human and pig roundworms may be limited, the long-term ability to cross-transmit between host species remains possible. In light of the data presented here, long-term control measures should be re-evaluated with regard to the potential that hybridization may play a role in the evolutionary dynamics of *Ascaris* populations.

In Guatemala and Nepal, ethical approval for the collection of nematodes came from the University of Texas Health Science Center Institutional Review Board in San Antonio, Texas, USA, and/or by the review boards of the respective countries. In China, local governments and clinics assisted with the collections.

We thank Maria-Eugenia Romero-Abal, Noel Solomons and the members of the Center for Studies of Sensory Impairment, Aging and Metabolism (CeSSIAM; Guatemala) and the Jiri Helminth Project (Nepal) for their assistance in organizing collections of nematodes. The molecular work at the Southwest Foundation for Biomedical Research was conducted in facilities constructed with support from Research Facilities Improvement Program grant C06 RR013556 from the National Center for Research Resources, National Institutes of Health. This work was funded by NIH grant R01 AI37091 to S.W.B.

REFERENCES

Anderson, T. J. C. 1995 *Ascaris* infections in humans from North America: molecular evidence for cross infection. *Parasitology* **110**, 215–219.

Anderson, T. J. C. 2001 The dangers of using single locus markers in parasite epidemiology: *Ascaris* as a case study. *Trends Parasitol.* **17**, 183–188. (doi:10.1016/S1471-4922(00)01944-9)

Anderson, T. J. C. & Jaenike, J. 1997 Host specificity, evolutionary relationships and macrogeographic differentiation among *Ascaris* populations from humans and pigs. *Parasitology* **115**, 325–342. (doi:10.1017/S0031182097001339)

Anderson, E. C. & Thompson, E. A. 2002 A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217–1229.

Anderson, T. J. C., Romero-Abal, M. E. & Jaenike, J. 1993 Genetic structure and epidemiology of *Ascaris* populations: patterns of host affiliation in Guatemala. *Parasitology* **107**, 319–334.

Anderson, T. J. C., Romero-Abal, M. E. & Jaenike, J. 1995 Mitochondrial DNA and *Ascaris* microepidemiology: the composition of parasite populations from individual hosts, families and villages. *Parasitology* **110**, 221–229.

Arnold, M. L. 2004 Natural hybridization and the evolution of domesticated, pest and disease organisms. *Mol. Ecol.* **13**, 997–1007. (doi:10.1111/j.1365-294X.2004.02145.x)

Arnold, M. L. 2006 *Evolution through genetic exchange*. Oxford, UK: Oxford University Press.

Banks, M. A., Eichert, W. & Olsen, J. B. 2003 Which genetic loci have greater population assignment power? *Bioinformatics* **19**, 1436–1438. (doi:10.1093/bioinformatics/btg172)

Barton, N. H. 2001 The role of hybridization in evolution. *Mol. Ecol.* **10**, 551–568. (doi:10.1046/j.1365-294x.2001.01216.x)

Berthier, P., Excoffier, L. & Ruedi, M. 2006 Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. *Proc. R. Soc. B* **273**, 3101–3109. (doi:10.1098/rspb.2006.3680)

Cavalli-Sforza, L. L. & Edwards, A. W. F. 1967 Phylogenetic analysis: models and estimation procedures. *Evolution* **21**, 550–570. (doi:10.2307/2406616)

Corander, J. & Marttinen, P. 2006 Bayesian identification of admixture events using multilocus molecular markers. *Mol. Ecol.* **15**, 2833–2843.

Corander, J., Marttinen, P. & Mantyniemi, S. 2006 A Bayesian method for identification of stock mixtures from molecular marker data. *Fish. Bull.* **104**, 550–558.

Criscione, C. D., Poulin, R. & Blouin, M. S. 2005 Molecular ecology of parasites: elucidating ecological and micro-evolutionary processes. *Mol. Ecol.* **14**, 2247–2257. (doi:10.1111/j.1365-294X.2005.02587.x)

Criscione, C. D. *et al*. 2007 Microsatellite markers for the human nematode parasite *Ascaris lumbricoides*: development and assessment of utility. *J. Parasitol.* **93**, 704–708. (doi:10.1645/GE-1058R.1)

Crompton, D. W. T. 2001 *Ascaris* and ascariasis. *Adv. Parasitol.* **48**, 285–375.

de Meeûs, T., McCoy, K. D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Boussès, S. & Renaud, F. 2007 Population genetics and molecular epidemiology or how to "debusquer la bete". *Infect. Genet. Evol.* **7**, 308–332. (doi:10.1016/j.meegid.2006.07.003)

Evanno, G., Regnaut, S. & Goudet, J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620. (doi:10.1111/j.1365-294X.2005.02553.x)

Falush, D., Stephens, M. & Pritchard, J. K. 2003 Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587.

Felsenstein, J. 2005 PHYLIP *(Phylogeny Inference Package) version 3.65*. Distributed by the author. Department of Genome Sciences, University of Washington Seattle.

Goudet, J. 1995 FSTAT (version 1.2): a computer program to calculate *F*-statistics. *J. Hered.* **86**, 485–486.

Irion, D. N., Schaffer, A. L., Famula, T. R., Eggleston, M. L., Hughes, S. S. & Pedersen, N. C. 2003 Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J. Hered.* **94**, 81–87. (doi:10.1093/jhered/esg004)

Macko, J. K. & Dubinsky, P. 1997 Taxonomic deliberations on human and pig ascarids. *Helminthologia* **34**, 167–171.

McCoy, K. D. 2003 Sympatric speciation in parasites—what is sympatry? *Trends Parasitol.* **19**, 400–404. (doi:10.1016/S1471-4922(03)00194-6)

Meirmans, P. G. 2006 Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399–2402.

Nejsum, P., Frydenberg, J., Roepstorff, A. & Parker Jr, E. D. 2005*a* Population structure in *Ascaris suum* (Nematoda) among domestic swine in Denmark as measured by whole genome DNA fingerprinting. *Hereditas* **142**, 7–14. (doi:10.1111/j.1601-5223.2005.01864.x)

Nejsum, P., Parker, E. D., Frydenberg, J., Roepstorff, A., Boes, J., Haque, R., Astrup, I., Prag, J. & Sorensen, U. B. S. 2005*b* Ascariasis is a zoonosis in Denmark. *J. Clin. Microbiol.* **43**, 1142–1148. (doi:10.1128/JCM.43.3.1142-1148.2005)

Nejsum, P., Grondahl, C. & Murrell, K. D. 2006 Molecular evidence for the infection of zoo chimpanzees by pig *Ascaris*. *Vet. Parasitol.* **139**, 203–210. (doi:10.1016/j.vetpar.2006.02.025)

Olden, J. D., Poff, N. L., Douglas, M. R., Douglas, M. E. & Fausch, K. D. 2004 Ecological and evolutionary consequences of biotic homogenization. *Trends Ecol. Evol.* **19**, 18–24. (doi:10.1016/j.tree.2003.09.010)

O'Lorcain, P. & Holland, C. V. 2000 The public health importance of *Acaris lumbricoides*. *Parasitology* **121**, S51–S71. (doi:10.1017/S0031182000006442)

Paetkau, D., Slade, R., Burdens, M. & Estoup, A. 2004 Genetic assignment methods for direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* **13**, 55–65. (doi:10.1046/j.1365-294X.2004.02008.x)

Peng, W., Anderson, T. J. C., Zhou, X. & Kennedy, M. W. 1998 Genetic variation in sympatric *Ascaris* populations from humans and pigs in China. *Parasitology* **117**, 355–361. (doi:10.1017/S0031182098003102)

Peng, W. D., Yuan, K., Zhou, X. M., Hu, M., El-Osta, Y. G. A. & Gasser, R. B. 2003 Molecular epidemiological investigation of *Ascaris* genotypes in China based on single-strand conformation polymorphism analysis of ribosomal DNA. *Electrophoresis* **24**, 2308–2315. (doi:10.1002/elps.200305455)

Peng, W. D., Yuan, K., Hu, M., Zhou, X. M. & Gasser, R. B. 2005 Mutation scanning-coupled analysis of haplotypic variability in mitochondrial DNA regions reveals low gene flow between human and porcine *Ascaris* in endemic regions of China. *Electrophoresis* **26**, 4317–4326. (doi:10.1002/elps.200500276)

Peng, W., Yuan, K., Hu, M., Peng, G., Zhou, X., Hu, N. & Gasser, R. B. 2006 Experimental infections of pigs and mice with selected genotypes of *Ascaris*. *Parasitology* **133**, 651–657. (doi:10.1017/S0031182006000643)

Rosenberg, N. A. 2004 DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137–138. (doi:10.1046/j.1471-8286.2003.00566.x)

Seehausen, O. 2004 Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207. (doi:10.1016/j.tree.2004.01.003)

Takezaki, N. & Nei, M. 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389–399.

Vähä, J. P. & Primmer, C. R. 2006 Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**, 63–72. (doi:10.1111/j.1365-294X.2005.02773.x)

Weir, B. S. & Cockerham, C. C. 1984 Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370. (doi:10.2307/2408641)

WHO 2002 *Prevention and control of schistosomiasis and soil transmitted helminthiasis; report of a WHO expert committee*. Geneva, Switzerland: World Health Organization.

Williams-Blangero, S., Subedi, J., Upadhayay, R. P., Manral, D. B., Rai, D. R., Jha, B., Robinson, E. S. & Blangero, J. 1999 Genetic analysis of susceptibility to infection with *Ascaris lumbricoides*. *Am. J. Trop. Med. Hyg.* **60**, 921–926.

Williams-Blangero, S., VandeBerg, J. L., Subedi, J., Aivaliotis, M. J., Rai, D. R., Upadhayay, R. P., Jha, B. & Blangero, J. 2002 Genes on chromosomes 1 and 13 have significant effects on *Ascaris* infection. *Proc. Natl Acad. Sci. USA* **99**, 5533–5538. (doi:10.1073/pnas.082115999)