

Genetic structure of populations of the human hookworm, *Necator americanus*, in China

J. M. HAWDON,* T. LI,† B. ZHAN* and M. S. BLOUIN‡

*Department of Microbiology and Tropical Medicine, George Washington University Medical Center, 725 Ross Hall, 2300 Eye St. NW, Washington DC 20037, USA, †Institute of Parasitic Diseases, Chinese Academy of Preventative Medicine, Shanghai, China,

‡Department of Zoology, Oregon State University, Corvallis, USA

Abstract

Twenty-one to 58 individual *Necator americanus* were sampled from each of four villages in south-western China. Each nematode was sequenced for 588 bp of the mitochondrial cytochrome oxidase I gene. Allelic and nucleotide diversity varied two-fold among villages. Overall F_{ST} among populations was ≈ 0.28 , but this large value resulted from one low-diversity population that had a large genetic distance to the other three populations ($F_{ST} = 0.10$ without that population). There was no correlation between geographical and genetic distance among sites. Thus, the genetic structure of this species in China may be characterized by variable effective sizes and uneven movement among sites. We discuss the implications of this genetic structure for vaccine development and the spread of drug resistance in human hookworms, and compare the genetic structure of hookworms with that of other nematodes.

Keywords: Ancylostomidae, gene flow, mitochondrial DNA, *Necator americanus*, nematode

Received 27 August 2000; revision received 17 January 2001; accepted 17 January 2001

Introduction

Infection with the human hookworms, *Ancylostoma duodenale* and *Necator americanus*, remains a burden on the health and economies of many developing nations in the tropics (Hotez *et al.* 1997). Anthelmintic drugs are effective, but rapid re-infection and the high cost of treatment hamper control efforts (Pawlowski 1990; Albonico *et al.* 1995). Furthermore, the widespread development of resistance to benzimidazole anthelmintics in trichostrongyle nematodes of veterinary importance (Conder & Campbell 1995) suggests that a similar situation could arise in human hookworms as anthelmintic treatment programmes intensify. Indeed, recent reports of mebendazole treatment failures in Mali suggest that anthelmintic resistance in hookworms may already have developed in Africa (De Clercq *et al.* 1997). An effective vaccine against hookworms would obviate the need for the widespread use of anthelmintics, but vaccine development efforts are still underway (Hotez *et al.* 1999).

An important factor that may influence both the utility of new vaccines and the development and spread of anthelmintic resistance in hookworms is the population genetic

structure of the parasite. Because genetic structure is largely controlled by the effective sizes of populations (N_e) and the rates of gene flow among them (m), knowledge of genetic structure gives insight into the rate and geographical scale over which both adaptation and random differentiation are likely to occur. For example, in the case of vaccine development, it is important to know whether populations are likely to differ at the loci targeted by highly specific recombinant vaccines, and thus to predict the geographical scale over which new vaccines need to be tested (Gupta *et al.* 1994; Anderson *et al.* 1998). Knowledge of genetic structure is even more important for controlling the development of anthelmintic resistance (Roush & Daly 1971; Anderson *et al.* 1998). High migration rates between populations retard local adaptation when selection pressures differ among populations. However, high migration helps to spread favourable alleles from their population of origin when selection favours the same phenotype in all populations. Also, selection is most efficient in large populations (more initial genetic variation to act on and less random drift, Hartl & Clark 1997). For example, large effective sizes and high gene flow may partly explain the rapid development and spread of resistance to benzimidazoles in trichostrongyles infecting livestock (Blouin *et al.* 1995; Conder & Campbell 1995). Whether widespread anthelmintic resistance is

Correspondence: John M. Hawdon. Fax: 1-202-994-2913; E-mail: mtmjmh@gwumc.edu.

equally likely to develop in human hookworms remains unclear.

For these reasons it is important to understand the genetic structure of human hookworm populations. Although people from developed countries are very mobile, in many developing nations the movement of people and their parasites among villages may be infrequent. Cultural practices may enhance or decrease the probability of transmission within and between villages. For example, in some situations transmission may cycle primarily within families owing to use of night soil in family gardens (Chang 1949). Thus, it is difficult to predict what effective population sizes and migration rates should be for human hookworms, and it is likely that these parameters vary from site to site. Consequently, it is not obvious a priori whether hookworms will show a genetic structure most similar to that seen in, for example, trichostrongylid nematodes at one extreme (high gene flow, very high genetic diversity) or to that seen in soil-dwelling nematodes at the other extreme (strongly structured, low diversity; Blouin 1998; Blouin *et al.* 1999).

Approximately 194 million people are infected with hookworms in rural areas of China (Xu *et al.* 1995; Hotez *et al.* 1997). The Chinese government has begun large-scale de-worming programmes in several areas using benzimidazole anthelmintics (Xia *et al.* 1991; Wen *et al.* 1998). In order to delay the development of drug resistance in these populations, it is important to understand the factors that will influence the evolution of resistance in China, including the genetic structure of the parasite. We are also developing recombinant vaccines to be tested in China (Hotez *et al.* 1999), therefore a knowledge of the extent to which the species is subdivided into genetically distinct populations will be important for choosing test sites. Therefore, we describe the genetic structure of *N. americanus* populations sampled from rural villages in China. We used mitochondrial DNA (mtDNA) sequence data in order to make our results comparable with previous studies on genetic structure in other nematodes.

Materials and methods

Hookworm specimens were collected from four study sites (villages) in southern China (Fig. 1A, Table 1), two in Sichuan province (sites 1 and 2), one in Yunnan province (site 3) and one on the island of Hainan off the southern mainland coast (site 4). The sites are separated by distances ranging from 230 to 1600 km.

Adult hookworms were obtained by de-worming ≈ 30 patients in each village, and were stored in 100% methanol. Single worm lysates were prepared following Higuchi (1992). One microlitre of template was used as a polymerase chain reaction (PCR) containing 2 mM $MgCl_2$, 1 U *Taq* DNA polymerase (Promega, Madison, WI, USA), and 100 ng each of Folmer *et al.*'s (1994) conserved cytochrome

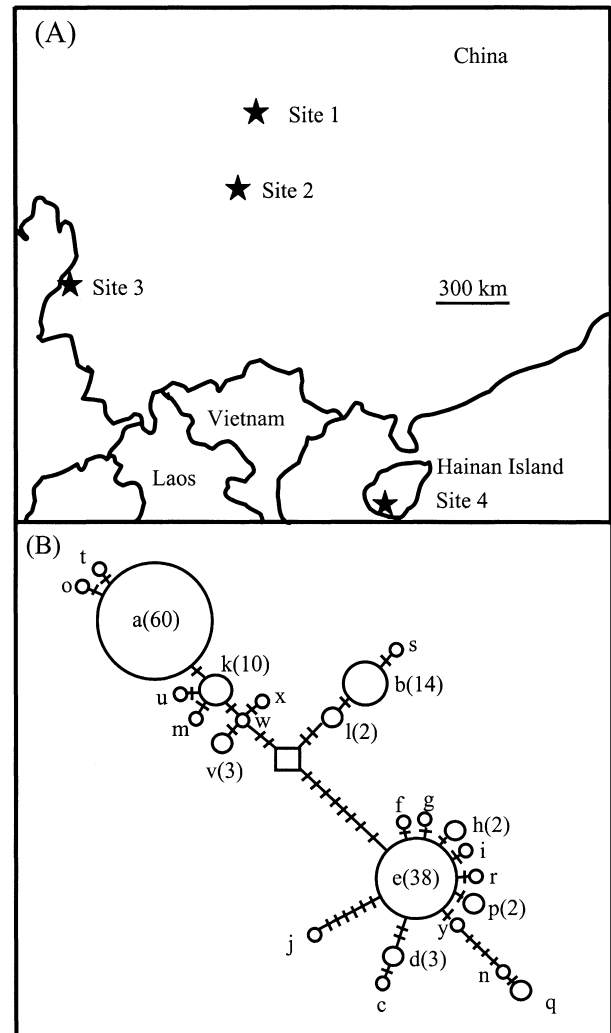


Fig. 1 (A) Map of south-western China showing the geographical arrangement of the study sites (indicated by stars). (B) Relationships among the 25 haplotypes. A minimum spanning tree of the haplotypes was obtained using ARELQUIN 2.0 (Schneider *et al.* 2000) and modified by hand. The number of copies of each haplotype in the dataset is indicated in parentheses for haplotypes occurring more than once. The area of circle is roughly proportional to relative abundance. Hash marks indicate nucleotide substitutions separating adjacent haplotypes, and the triangle represents an unseen haplotype.

oxidase I primers COR722 and COF14, and subjected to 35 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min. Amplicons were sequenced in both directions.

Five hundred and eighty-eight bases of the *CO I* gene were analysed, corresponding to positions 39–626 of the *Ancylostoma caninum CO I* gene sequence (Accession no. U57030). Sample sizes per site are listed in Table 1. In some cases up to five worms were sequenced from the same host. ARELQUIN 2.0 was also used to obtain AMOVA (Excoffier *et al.* 1992) estimates of *F*-statistics (Φ_{ST}) from the uncorrected

Table 1 Sample sizes, number of unique haplotypes observed, nucleotide diversity (π), haplotype diversity (h), Fu's F_s and Tajima's D statistics, and frequencies of the four most common alleles (a, b, e and k) in each population. P -values for Fu's and Tajima's tests are in parentheses. N = number of worms sequenced. Haplotype diversity is defined as in Nei (1987)

Population*	N	No. hap.	π	h	F_s	D	Frequencies of common alleles			
							a	b	e	k
1	38	13	0.011	0.83	0.56 (0.64)	0.79 (0.22)	0.18	0.11	0.37	0.08
2	58	7	0.006	0.34	2.81 (0.88)	-0.76 (0.26)	0.81	0.03	0.07	0.02
3	34	14	0.012	0.89	-0.54 (0.44)	0.60 (0.22)	0.06	0.24	0.24	0.15
4	21	6	0.009	0.65	2.99 (0.90)	0.34 (0.36)	0.19	0	0.57	0.05
Entire dataset	151	25	0.012	0.77	-1.04 (0.43)	0.20 (0.67)	—	—	—	—

*Study site locations: 1 = lat. 31°02', long. 104°45'; 2 = lat. 28°49', lon. 105°59'; 3 = lat. 25°51', lon. 98°51'; 4 = lat. 18°52', lon. 109°88'.

number of nucleotide differences between sequences, to obtain traditional F -statistics (i.e. based on allele frequencies), and to conduct neutrality tests using Fu's F (Fu 1997) and Tajima's D (Tajima 1989) statistics.

Anderson *et al.* (1993) found that *Ascaris* within human hosts in Guatemalan villages shared identical mtDNA haplotypes more often than expected by chance alone, a result which suggests that transmission is not random within villages. A similar situation may occur with hookworms because the use of night soil on family gardens could cause infection and reinfection to largely cycle within families (Chang 1949). Two or more hookworm sequences from the same host were available in at least 10 hosts per site from three of the four sites (10, 12 and 15 hosts in sites 2, 1 and 3, respectively). In order to test whether hookworms sampled within the same host had identical haplotypes more often than expected by chance alone, we used these samples in a hierarchical F -statistics analysis on allele identity, and tested the significance of the variance component among hosts within a site.

Results

Twenty-five unique haplotypes were observed in the dataset (Fig. 1B). Fu's and Tajima's tests indicate that the tree does not deviate substantially from that expected in a single population under mutation–drift equilibrium (Table 1). Thus, there is no evidence that the sampled haplotypes represent a recent mix of previously isolated lineages.

In this study only 5.8% of the total variation in allelic identity was distributed among hosts within sites, and this value was not significantly different from zero ($P = 0.06$). Thus, there is no strong evidence that nematode genotypes are nonrandomly distributed among hosts within villages.

Haplotype diversity and nucleotide diversity were similar among sites 1, 3 and 4, but substantially lower in site 2 (Table 1). Overall $\Phi_{ST} = 0.28$ among the set of four populations. However, this large Φ_{ST} results because population 2 is very different from the other three (Table 2). If one

Table 2 Pairwise estimates of Φ_{ST} (distance based; below diagonal) and traditional F_{ST} (based on allele frequencies; above diagonal) between populations

Population	1	2	3	4
1	—	0.309*	0.015	0.012
2	0.377*	—	0.374*	0.426*
3	0.081*	0.256*	—	0.093*
4	0.015	0.542*	0.223*	—

*Significantly different from zero at $P < 0.01$.

excludes population 2 then the overall $\Phi_{ST} = 0.10$. The large effect of population 2 results because it has a high frequency of allele a while the other populations tend to have few copies of allele a and more copies of allele e (Table 1, Fig. 1B). Traditional F_{ST} gives overall and pairwise values very similar to the Φ_{ST} estimates (overall $F_{ST} = 0.24$; pairwise in Table 2). On a regional scale there is no correlation between genetic and geographical distances among populations (Fig. 1A). For example, population 1 is closest to the very distinct population 2 (230 km; pairwise $\Phi_{ST} = 0.38$; $P < 0.001$), whereas it is indistinguishable from population 4 which is 1600 km away on the island of Hainan (pairwise $\Phi_{ST} = 0.015$; $P = 0.23$). A recent bottleneck in site 2 is consistent with its low diversity and large genetic distance from the other populations (Paetkau *et al.* 1997; Charlesworth 1998), but the cause of such a bottleneck is not obvious.

Discussion

Implications for hookworm control in China

The lack of correlation between geographical and genetic distances, and the large variation in genetic diversity among populations, suggests that one should not assume that these populations conform to a simple model of genetic structure, such as a stepping stone in drift–migration

equilibrium. As a working hypothesis, we suggest that *Necator americanus* in China may be characterized by: (i) sporadic bouts of gene flow over short or long distances, owing to nonrandom movement by infected humans; and (ii) by fluctuating population sizes. This genetic structure probably results from varying cultural practices that influence transmission and infection rates. For example, farmers in urban areas often collect faeces from community latrines for use on their farms, and so might be exposed to infective stages from other townspeople and visitors, such as passing truck drivers. Such practices are uncommon in more rural areas.

There are several practical consequences of the hypothesized genetic structure. First, there is unlikely to be a simple relationship between genetic similarity and geographical distance, at least over the geographical scale sampled in this study. Thus, predicting which populations are likely to differ substantially at loci of interest, such as vaccine antigens, will not be easy without actually genotyping each population. Second, the rate of development of anthelmintic resistance in any given population may be difficult to predict as it will depend on the particular history of effective size fluctuations at that site, and on the intensities of selection in that population and in the populations from which it receives migrants (which may not be the closest geographically).

Comparison with other nematodes

Data on mtDNA diversity and genetic structure are now available for several nematode species (Anderson *et al.* 1998; Blouin 1998). *Necator americanus* does not show the huge mtDNA diversities and high rates of gene flow among populations seen in trichostrongylid parasites of livestock (Blouin *et al.* 1995), or the low diversities and highly structured populations typical of some plant and insect parasitic nematodes (Hugall *et al.* 1994; Blouin *et al.* 1999). In terms of levels of diversity within populations, and F_{ST} among populations, the population genetic structure of *Necator* is most similar to that of *Ascaris*, another parasite of humans (Anderson *et al.* 1995; Anderson & Jaenike 1997). That *Necator* and *Ascaris* have the most similar genetic structures is perhaps not surprising given they use the same host, both have obligate sexual life cycles, and they are transmitted via faecal contamination. Thus, the opportunities for gene flow are similar in the two species. As for effective size, the average number of worms per infected host is higher in *Necator* (mean of 26 per person from these four sites) than in *Ascaris* (one per person) (Anderson *et al.* 1997). However, *Ascaris* rely on tough, long-lived infective eggs for transmission, whereas infective *Necator* are relatively short-lived and prone to desiccation. So there is no long-term reservoir of *Necator* larvae, whereas the resistant eggs of *Ascaris* may act like a 'seed bank' to dampen large

fluctuations in effective size that may be more typical of *Necator* populations. Thus effective sizes probably average out to be similar in the two species, giving similar genetic structures.

This is one of only a handful of studies on genetic structure in nematodes, and the only one on hookworms. More sites from China and from other countries need to be sampled in order to fully describe the genetic structure of *N. americanus*, and to predict what environmental and cultural factors control gene flow and effective size in this species. A comparative study of *Necator* and other nematode parasites of humans (e.g. *Ascaris*, *Ancylostoma*, *Strongyloides*), that used the same markers and host populations would go a long way towards revealing how key features of nematode life cycles interact with human behaviours to control the movement, transmission and persistence of these parasites in human populations.

Acknowledgements

The authors thank the people of the villages for their cooperation and contributions to this study. We also thank Liu Changhua, Zhang Xiaorong and Qiu Dongchuan of the Sichuan Institute of Parasitic Diseases, Chen Jizhang, Xing Fuying and Li Shanwen of the Hainan Institute of Parasitic Diseases, and Zhang Lili, Peng Yankun and Li Fuhua of the Yunnan Provincial Anti-endemic Disease Station. Finally, we thank Xiao Shuhua and Feng Zheng of the Institute of Parasitic Diseases in Shanghai, and Peter Hotez for help and support. This research was supported by TMRC grant number 1-P50A139461 from the NIH.

References

- Albonico M, Smith PG, Ercole E *et al.* (1995) Rate of reinfection with intestinal nematodes after treatment of children with mebendazole or albendazole in a highly endemic area. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **89**, 538–541.
- Anderson TJ, Blouin MS, Beech RN (1998) Population biology of parasitic nematodes: applications of genetic markers. *Advances in Parasitology*, **41**, 219–283.
- Anderson TJC, Jaenike J (1997) Host specificity, evolutionary relationships and macrogeographic differentiation among *Ascaris* populations from humans and pigs. *Parasitology*, **115**, 325–342.
- Anderson TJC, Romero Abal ME, Jaenike J (1993) Genetic structure and epidemiology of *Ascaris* populations: patterns of host affiliation in Guatemala. *Parasitology*, **107**, 319–334.
- Anderson TJC, Romero-Abal ME, Jaenike J (1995) Mitochondrial DNA and *Ascaris* microepidemiology: the composition of parasite populations from individual hosts, families and villages. *Parasitology*, **110**, 221–229.
- Blouin MS (1998) Mitochondrial DNA diversity in nematodes. *Journal of Helminthology*, **72**, 285–289.
- Blouin MS, Liu J, Berry RE (1999) Life cycle variation and the genetic structure of nematode populations. *Heredity*, **83**, 253–259.
- Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, **141**, 1007–1014.

- Chang K (1949) *Studies on Hookworm Disease in Szechwan Province, West China*. Johns Hopkins Press, Baltimore.
- Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, **15**, 538–543.
- Conder GA, Campbell WC (1995) Chemotherapy of nematode infections of veterinary importance, with special reference to drug resistance. *Advances in Parasitology*, **35**, 1–84.
- De Clercq D, Sacko M, Behnke J, Gilbert F, Dorny P, Vercruyse J (1997) Failure of mebendazole in treatment of human hookworm infections in the southern region of Mali. *American Journal of Tropical Medicine and Hygiene*, **57**, 25–30.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Gupta S, Trenholme K, Anderson RM, Day KP (1994) Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science*, **263**, 961–963.
- Hartl DL, Clark AG (1997) *Principles of Population Genetics*. Sinauer, Sunderland, MA.
- Higuchi R (1992) Simple and rapid preparation of samples for PCR. In: *PCR Technology: Principles and Applications for DNA Amplification* (ed. Erlich HA), pp. 31–38. Freeman, New York.
- Hotez PJ, Ghosh K, Hawdon JM *et al.* (1999) Experimental approaches to the development of a recombinant hookworm vaccine. *Immunological Reviews*, **171**, 163–171.
- Hotez PJ, Zheng F, Long Qi X *et al.* (1997) Emerging and reemerging helminthiases and the public health of China. *Emerging Infectious Diseases*, **3**, 303–310.
- Hugall A, Moritz C, Stanton J, Wolstenholme DR (1994) Low, but strongly structured mitochondrial DNA diversity in root knot nematodes (*Meloidogyne*). *Genetics*, **136**, 903–912.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.
- Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943–1957.
- Pawlowski ZS (1990) Strategies for hookworm control. In: *Hookworm Disease: Current Status and New Directions* (eds Schad GA, Warren KS), pp. 210–217. Taylor & Francis, London.
- Roush RT, Daly JC (1971) The role of population genetics in resistance research and management. In: *Pesticide Resistance in Arthropods* (eds Roush RT, Tabashnik BE), pp. 97–152. Chapman & Hall, New York.
- Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN: A Software for Population Genetics Data Analysis*. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Switzerland.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–596.
- Wen L, Xia Z, Yao S, Su Y, Ren S, Song C (1998) A ten year longitudinal observation on the control of hookworm and other soil-transmitted nematodiosis in the study site [In Chinese]. *Chinese Journal of Parasitic Disease Control*, **11**, 165–167.
- Xia ZH, Yao SY, Su YL, Wen LY, Song CC (1991) Studies on the control of hookworm and other soil-transmitted helminthiases in farmers in Zhejiang Province China. *Southeast Asian Journal of Tropical Medicine and Public Health*, **22**, 618–622.
- Xu LQ, Yu SH, Jiang ZX, Yang JL, Lai CQ, Zhang XJ, Zheng CQ (1995) Soil-transmitted helminthiases. Nationwide survey in China. *Bulletin of the World Health Organization*, **73**, 507–513.

This project is a component of the NIH-supported Tropical Medicine Research Center research project entitled 'Emerging Helminthiases of China' in collaboration with the Institute of Parasitic Diseases in Shanghai, China. The overall goal of the hookworm component is to determine whether genetic variation will have an impact upon the development and implementation of a recombinant hookworm vaccine. John M. Hawdon's research interests include genetic variation in hookworms, and nematode growth and development. He and his colleague Bin Zhan of the Department of Microbiology and Tropical Medicine at George Washington University Medical Center are also engaged in the development of a recombinant hookworm vaccine. Tiehua Li is faculty at the Institute of Parasitic Diseases in Shanghai, where she studies genetic variation in the human hookworms. Michael S. Blouin's research at Oregon State University focuses on the effects of life cycle variation on genetic structure in parasites and other taxa.
