Whipworms in humans and pigs: Origins and demography

Mohamed B.F. Hawash[1, 2], Martha Betson[3,4], Azmi Al-Gabury[1], Jennifer Ketzis[5], Arve Lee Willingham[5], Mads F. Bertelsen[6], Philip J. Cooper[7], D. Tim J. Littlewood[8], Xing-Quan Zhu[9], Peter Nejsum[1,*]

1. Department of Veterinary Disease Biology, University of Copenhagen, Frederiksberg, Denmark.
2. Zoology Department, Faculty of Science, Cairo University, Cairo, Egypt.
3. Department of Production and Population Health, Royal Veterinary College, Hatfield, Hertfordshire, UK
4. School of Veterinary Medicine, University of Surrey, Guildford, Surrey, UK
5. Ross University School of Veterinary Medicine, Basseterre, St Kitts, West Indies
6. Centre for Zoo and Wild Animal Health, Copenhagen Zoo, Frederiksberg, Copenhagen, Denmark
7. Institute of Infection and Immunity, St George's University of London, London, UK
8. Department of Life Sciences, Natural History Museum, London, UK
9. State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu Province 730046, PR China

Corresponding author: Peter Nejsum, Department of Veterinary Disease Biology, Faculty of Health Sciences, Copenhagen University, Dyrlaegevej 100, DK-1870, Frederiksberg C, Copenhagen, Denmark. Phone: +45 5054 1392. E-mail: pn@sund.ku.dk

Abstract

Trichuris suis and *T. trichiura* are two different whipworm species that infect pigs and humans, respectively. *T. suis* is found in pigs worldwide while *T. trichiura* is responsible for nearly 460 million infections in people mainly in areas of poor sanitation in tropical and subtropical areas. In this study, we aimed to reconstruct the demographic history of *Trichuris* in humans and pigs, the evolutionary origin of *Trichuris* in these hosts and factors responsible for parasite dispersal globally. Population genetic and phylogenetic analyses were applied to populations of *Trichuris* recovered from humans, pigs and non-human primates in different countries on different continents, namely Denmark, USA, Uganda, Ecuador, China and St. Kitts (Caribbean). We found no differentiation between human-derived *Trichuris* in Uganda and the majority of the *Trichuris* samples from non-human primates suggesting a common African origin of the parasite which then was transmitted to Asia and further to South America. Moreover, the demographic history of pig *Trichuris* underlies the major role played by human activity in transporting pigs and their parasites through colonisation.

Key words: Whipworms, *Trichuris*, humans, pigs, demographic history, evolution

Introduction

A range of mammalian hosts is infected with parasitic whipworms belonging to the genus *Trichuris*. Around 460 million humans are infected with *T. trichiura* mainly in developing countries in South and South-East Asia, Sub-Saharan Africa and Latin America [1]. The prevalence of whipworms in non-human primates is generally high and despite the fact that the taxonomic status is unsettled, they are historically designated as *T. trichuris* [2]. *T. suis* infection in pigs is globally widespread with the highest prevalence in young pigs reared in outdoor production systems [3,4]. However, the evolutionary relationship between *Trichuris* from pigs and primates is poorly understood and in particular the anthropogenic and environmental factors responsible for their global distribution.

Population genetic tools provide a valuable opportunity to investigate the epidemiological history and transmission of parasites and have been adopted to study many parasitic nematodes [5,6]. For instance, population genetic approaches were applied recently to investigate the pattern of transmission of *Ascaris suum* and *A. lumbricoides* between pigs and humans across the globe [7]. Moreover, population genetics can be used to reconstruct the epidemiological and demographic history of micro- and macro- parasites through coalescent analysis coupled with Bayesian approaches [5]. Reconstruction of the epidemiological and demographic history gives us a window into the past to see which factors facilitated spread or the introduction of the parasites to new regions [5]. For instance, the demographic history of *Wuchereria bancrofti* suggests that this parasite was introduced to India 60,000-70,000 years ago through human migration out of Africa. In contrast, the introduction of *W. bancrofti* to Papua New Guinea cannot be explained by ancient human migration but must have been introduced through more recent human migration [8].

66

67 During the evolution of genus *Homo* nearly 4 million years ago, there has been continuous
68 contact with many parasites. Parasites which were transmitted to humans through primate
69 common ancestors are referred to as "heirloom" while the parasites which were acquired more
70 recently through contact with animals e.g. during animal domestication, are termed
71 "souvenirs" [9]. The human whipworm is generally considered heirloom as it is found in the
72 African non-human primates, and parasite eggs were found in human coprolites in
73 archaeological sites before animal domestication and in the New World before the Columbian
74 colonization [10-12]. However, there is no rigorous genetic evidence for this assumption
75 especially with recent studies suggesting that several *Trichuris* species can be found in non-
76 human primates [13,14]. Hence, the genetic and evolutionary relationship between *Trichuris*
77 from humans and non-human primates is poorly understood.

78

79 Herein, we investigated the genetic and evolutionary relationships between populations of
80 *Trichuris* parasites derived from humans, non-human primates and pigs from different
81 continents. The mitochondrial *nad*1 and *rrn*L genes of 140 worms were sequenced and
82 coalescent simulations applied to infer the demographic history of the different human and
83 pig *Trichuris* populations and their evolutionary origin.

84

85 **Results**

86 Phylogenetic and genetic differentiation analyses

87 We used two mitochondrial markers to infer the genetic relationships and the evolutionary
88 history between different *Trichuris* populations obtained from humans, nom-human primates

and pigs from various geographical regions (Table 1). Partial sequences of the large ribosomal subunit (*rrn*L) and NADH dehydrogenase subunit 1 (*nad*1) were genereated for all worms except for *Trichuris* from African green monkey, for which only the *rrn*L gene was sequenced (GenBank accession numbers XXX-YYY). The phylogenetic relationship was found to be identical when inferred using neighbor joining (NJ) and maximum likelihood (ML) clustering methods and for both genes, and hence only the NJ tree for the *rrn*L gene is shown in Figure 1. For the ML tree, Hasegawa-Kishino-Yano with gamma distribution (HKY+G) model was used for the *nad*1 gene and Tamura-Nei with gamma distribution (TrN+G) model for the *rrn*L gene as the best-to-fit substitution models. It is noteworthy that two samples (two *T. trichiura* from humans, one in Uganda and one in China) showed double peaks in the chromatogram of the *nad*1 gene and gave conflicting signal for the *rrn*L gene (clustered in the pig clades). This may indicate co-amplification of nuclear mitochondrial pseudogenes (numts) or a heteroplasmy. Hence, these samples were excluded from further analyses.

Phylogeographic distribution among *T. suis* populations was observed as worms from Uganda and from China were found in separate clades whereas *T. suis* from Denmark and USA clustered together. The Ecuadorian *T. suis* were found in two clades, namely the Denmark & USA clade (4 samples) and the China clade (3 samples). In addition, two pig worms from Uganda clustered with *T. suis* from Denmark & USA. Phylogeographic structure was also observed for worms recovered from humans as *T. trichiura* from Uganda, China and Ecuador were found in separate clades. The majority of the baboon and green monkey worms clustered with the human *T. trichiura* from Uganda. Meanwhile, a few baboon samples (n=7) and a single green monkey worm grouped together in a different clade (*Trichuris* non-human

primates). The Neighbor-net network identified splits that correspond to the clades in the NJ tree (Figure 2).

Analysis of Molecular Variance (AMOVA) was used to analyse the degree of genetic differentiation (Fixation index, F_{st}) between *Trichuris* populations identified in the phylogenetic analysis and is summarised in Table 2 and is given for each of the two markers, but excluding *T. suis* from Ecuador as they cluster both with worms from China and Denmark. In general all the populations were highly differentiated as F_{st} values are above 0.25. *Trichuris* from baboons and green monkey in the two different clades (*Trichuris* non-human primates and *T. trichiura* Uganda) were highly differentiated with F_{st} values of 0.363 for *rrn*L ($P<0.01$) and 0.471 for *nad*1 ($P<0.01$). In contrast, human, baboon and green monkey *Trichuris* in the *T. trichiura* Uganda clade represented undifferentiated populations ($F_{st} < 0.05$, $P > 0.05$). The genetic distances (p-distance) between and within each clade are given in Supplementary Table 1 for the *nad*1 and *rrn*L genes, respectively. In Supplementary Table 2 the nucleotide diversity (π) and haplotype diversity for worms belonging to the different clades are given.

TMRCA and the divergence time of the human and pig *Trichuris* populations

The estimated Θ was 20.7 ± 6.68 for *T. suis* and 146.49 ± 50.1 for *T. trichiura* from humans and the Genetree for their populations is provided in Supplementary Figure S1. Hence, the TMRCA in generations for *T. suis* equals 80,000 generations (upper estimate is 110,000 and lower estimate 60,000) and 560,000 generations for *T. trichiura* (upper estimate is 760,000 and lower estimate is 390,000). The time of divergence between the different populations as estimated by BEAST is given at each node (Figure 3) in number of generations and by IMa2

in Supplementary Figure S2. BEAST, IMa2 and Genetree gave similar results for the time of divergence of the human *Trichuris* populations. However, the estimated time of divergence for *T. suis* populations was nearly three times older for BEAST and IMa2 than Genetree. Genetree relies on importance sampling (IS) algorithm for the coalescent simulations while BEAST and IMa2 each rely on correlated sampling (CS). As the IS algorithm is more suitable for data of low polymorphism [15], its estimate may be the most reliable in our case. There were two main divergence events in *T. suis* and *T. trichiura* populations. For *T. suis* there was an ancient split between the USA/Denmark populations and the China/Uganda populations (80,000-240,000 generations) and a more recent split between the populations from China and Uganda (32,000-90,000 generations). For *T. trichiura*, the first split is between the Uganda population and China/Ecuador populations (500,000 generations) and a more recent split between the China and Ecuador populations (120,000 generations).

Discussion

Herein, we investigated the evolutionary and genetic relationships between populations of *Trichuris* from primates and pigs from different geographical areas in order to infer the evolutionary and demographic history of the parasites, and to identify possible environmental and anthropological factors driving their spread across the globe. *Trichuris* from primates and pigs were genetically very distinct with independent demographic histories as summarised in Figure 4 and discussed below.

The coalescent analysis identified two divergence events in *T. suis* populations. First an ancient split between the DK/USA and the Chinese/Ugandan populations around 80,000 generations ago with a second more recent divergence between the China and Uganda

populations 32,000 generations ago. Interestingly, this latter split between *T. suis* from China and Uganda is in line with that of their host [16]. Common alleles between the domesticated pigs of Far East and East African origin have been identified suggesting close evolutionary relationships between pigs in these regions [16,17]. First, this may reflect transport of pigs from the Far East by the European trading routes to Africa a few hundreds of years ago. Secondly, domesticated pigs may have been introduced to Africa by trading between the ancient civilizations in Africa and Far East or the settlement of Austronesian peoples in East Africa nearly seven thousand years ago [16]. Given the high genetic differentiation between the *T. suis* populations of China and Uganda and assuming one to three generations for *T. suis* per year, the divergence between Chinese and Uganda *T. suis* population happened 32,000-10,700 years ago. Hence, it is unlikely that *T. suis* was introduced by an European intermediary only a few hundred years ago but may be traced back to the introduction of domesticated pigs from the Far East thousands of years ago. However, this does not exclude other waves of recent introduction of *T. suis* in domesticated pigs from the Far East. A recent study found that *A. lumbricoides* in humans on Zanzibar were closely related to worms from Bangladesh [7] suggesting parasite transportation between the Far East and the Indian subcontinent and East Africa.

Intriguingly, the *T. suis* population in Uganda was found to be monomorphic for both markers. This may relate to either a founder effect (the establishment of a new population from few individuals derived from a much larger population) or a selective sweep (strong positive natural selection of some few genotypes) or a combination of both factors. Such a selective sweep might have occurred due to new adaptations to host physiology in the new environment or, more likely, due to a recent bottleneck in the pig populations: e.g. African

185 swine fever outbreaks in Uganda [18] resulted in subsequent reductions in molecular variation
186 among the parasites [19].

187

188 Several studies have reported high genetic distinctiveness between European and Chinese
189 pigs [20,21] in line with our observation for *T. suis* from Denmark and China. However, the
190 divergence time between the pigs in China and Europe has been estimated to be roughly a
191 million years ago [21] which is much older than the divergence time estimated for the Chinese
192 and Danish *T. suis* populations in our study (80,000 years assuming one generation/year). In
193 principle, when populations suffer from many bottlenecks, the coalescent will date back to the
194 most recent bottleneck rather than the most recent common ancestor [5]. Considering the
195 bottlenecks that the pig populations have been through during their migration and transport
196 across Eurasia, with the most recent being nearly 20,000 years ago during the last glacial
197 maxima [21], this may also have resulted in bottlenecks of the associated parasites. In addition,
198 as *T. suis* can survive in the environment for at least 11 years [22] the number of generations per
199 year is very hard to estimate and could be as low as 0.09 generations/year.

200

201 *T. suis* from Denmark and USA were found to cluster together (Figure 1) with their
202 populations being undifferentiated. This is consistent with other parasites of domestic pigs
203 such as *Trichinella spiralis*, which was introduced to the Americas by Europeans [23]. However,
204 two *T. suis* from Uganda clustered with worms from Denmark and USA suggesting recent
205 transport of pigs between these continents. The clustering of *T. suis* from Ecuador with the
206 populations from China and DK/USA may reflect introgression between pigs from Europe
207 and China during the industrial revolution in Europe in the 20th century as found in a
208 previous study [24].

As for pig worms two divergence events were also observed for human derived *Trichuris*, one ancient (~500,000 generation) divergence between Uganda and China and a more recent one between China and Ecuador (~120,000 generations). Assuming the highest and lowest number of generations to be 1 and 3, respectively the divergence times were 500,000-160,000 generations and 120,000-40,000 generations for the Ugandan/Chinese and Chinese/Ecuadorian populations, respectively. The split between the China and Uganda populations preceded the modern human (*Homo sapiens*) migration out of Africa to South East Asia around 60,000-100,000 years ago [25,26] and the human settlement in Latin America 14,000-15,000 years ago [27]. There are two possibilities for this discrepancy. Firstly, one of the early human ancestors (e.g. *H. erectus*) may have transmitted *T. trichuris* when migrating out of Africa [28] but this would not explain how the parasite was then introduced into Latin America. Secondly, the mutation rate of the free living nematode (*C. elegans*) may not be applicable to parasitic nematodes as the mutation rate normally is higher for the latter [29]. The oldest record of *T. trichiura* eggs is in Brazil and dates back to 6000-7000 BP [30] suggesting that the parasite was introduced to the New World with the human migration much earlier than the Columbian colonization, which is concordant with our findings.

Green monkeys, olive, yellow and hamadryas baboons are indigenous species in central and western Africa and the introduction of the green monkey into Saint Kitts by the French in the late Seventeenth century [31] may explain why the majority of the baboon and green monkey worms clustered with human worms from Uganda. However, seven worms from baboons from both Denmark and USA and one from green monkey from St. Kitts were found in a separate clade (*Trichuris* non-human primate), suggesting that different populations are

circulating among these hosts' species although they were sampled from the same habitat. Since many of these worms were sampled from unnatural habitats (zoos), the causes for this differentiation could not be investigated. However, in a sympatric natural transmission area in Uganda some *Trichuris* genotypes were found among all seven investigated non-human primates, whereas other genotypes seem to be more host specific [13]. Hence, in the past, *Trichuris* in primates in Africa may have been isolated either by geography or by host species leading to population differentiation followed by a more recent secondary sympatry [32]. As the *Trichuris* population between humans from Uganda and the non-human primates are undifferentiated this suggests continuous or recent gene flow between the host species and suggests Africa as the origin of *Trichuris* in primates.

Unlike whipworms, it is expected that host shift of the giant roundworm *Ascaris* in humans and pigs took place during animal domestication in the Neolithic period 10,000 years ago [7,12]. This is supported by their very close genetic relationship [7,33] and that *Ascaris* among non-human primates is observed rarely [34] suggesting that this parasite represents a 'souvenir parasite' [9]. Their different evolutionary history is interesting as *Trichuris* and *Ascaris* share several parasitic traits such as mode of transmission and infection and these two parasite species therefore usually co-occur today [35].

In conclusion, we inferred the possible demographic history of *Trichuris* from humans and pigs and their potential evolutionary epicenter, namely in primates in Africa. We suggest that *Trichuris* was dispersed to Asia with human ancestors and that host switching to pigs occurred in China where pigs evolved [21]. *T. suis* was then spread across the globe mainly by anthropogenic factors. We found that *T. trichiura* in humans in Africa is genetically similar to

Trichuris in non-human primates of African origin suggesting that *Trichuris* in humans represents a heirloom parasite. Further studies should investigate the genetic relationships of whipworms from different primates and other host species living in natural habitats in order to explore their demographic history and evolutionary origins and to identify the possible switching events between host species.

Materials and methods

Parasite isolates, DNA extraction and typing of the worms

A total of 140 worms were collected from humans, pigs and non-human primates from different regions (Table 1). All worms were rinsed with tap water and stored in 70% ethanol at 5°C until DNA extraction. The MasterPure DNA Purifications Kit (Epicentre Biotechnologies) was used to extract DNA according to manufacturer's protocol after homogenization in 300 µl of lysis solution and overnight incubation at 56°C except for the populations from Ecuador for which the DNA was extracted previously [24]. Worms were initially typed to confirm worm species by Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) on internal transcribed spacer-2 (ITS-2) following the protocol described by [36]. A negative water control was included in all runs. PCR products and RFLP fragments were stained by GelRed (Biotium) and visualized under UV light in 1.5% agarose gel. All *Trichuris* from baboons and humans showed banding pattern characteristic of *Trichuris* from primates, while all *T. suis* worms showed the banding pattern characteristic of *T. suis*.

Amplification of genetic markers and sequencing

Partial sequences of the two mitochondrial genes *rrn*L *nad*1 were obtained for all samples except for *Trichuris* from African green monkey, for which only the *rrn*L gene was sequenced. 562 bp of the *nad*1 gene was amplified using forward SuiND1_F (5′-CGAGCTTATATAGGTATTTCTCAACG-3′) and reverse SuiND1_R (5′-CGTTGTAGCCTCTTACTAATTCTCTTT-3′) primers while 422 bp of the *rrn*L gene was amplified using primers, forward TrirrnL_F (5'-TGTAAWTCTCCTGCCCAATGA) and reverse TrirrnL_R (5'-CGGTTTAAACTCAAATCACGTA). The PCR conditions were identical for both markers and were conducted in a total volume of 20 µl using 1 µl DNA as template. PCR ingredients were: 1X PCR buffer, 0.2 mM of each dNTP, 0.4 mM of each primer pair, 2.0 mM $MgCl_2$, and 1 U of Hot Start DNA-polymerase (Ampliqon). PCR conditions were initial denaturation at 95°C for 15 min followed by 35 cycles consisting of 95°C for 30 s, 55°C for 30 s and 72°C for 1 min and a final extension at 72°C for 10 min. Agarose gel electrophoresis (1.5%) was used to verify amplification of a single fragment of the expected size. PCR products were enzymatically cleaned prior to sequencing using 10 µl of PCR product, 1 µl Exonuclease I and 2 µl FastAP Thermosensitive Alkaline Phosphotase (1 U/µl) (Fermentas). The samples were incubated for 15 min at 37°C followed by 15 min at 85°C. Finally, cleaned amplicons were sequenced in both directions using same primers used for PCR by Macrogen Inc. in Seoul, South Korea.

Genetic variation, differentiation and phylogenetic relationships

Forward and reverse sequences from each sample were checked, edited manually and assembled using vector NTI [37] and then trimmed using BioEdit [38]. Genetic relatedness and evolutionary relationship were analysed for each of the two markers using 410 bp and 397 bp of the *nad*1 and *rrn*L genes, respectively. Two sequences of the *nad*1 and *rrn*L from the Chinese human and pig *Trichuris* mitochondrial genomes (Accession No: GU385218 and

GU070737, respectively) were included in the dataset. Also, *T. trichiura rrn*L gene sequences from humans in China were included in the dataset (Accession no. AM993017-AM993023). Phylogenetic relationship was inferred using NJ and ML phylogenetic trees in MEGA v6.1 [39]. The best-to-fit substitution model was identified using jModelTest0.1.1 [40] under Akaike information criterion (AIC) [41]. *Trichinella spiralis* was used as outgroup (Accession No: AF293969). The most parsimonious network was inferred by the Neighbor-net method using SplitsTree v.4.13.1 [42]. Neighbor-net network can reveal ambiguous and incompatible sites which usually appear as a reticulate structure in the network.

The nucleotide diversity (π) and haplotype diversity within the different major clades (splits) identified in the phylogenetic analysis were estimated using the software DnaSP v.5 [43]. AMOVA was used to estimate the Fixation index, F_{st} between *Trichuris* populations using Arlequin v.3.5.1.2 [44]. 10,000 permutations were used to test for differentiation between pairs of populations. As the origin of *Trichuris* collected from baboons was unknown due to previous transport between zoological gardens these populations were omitted from F_{st} calculations. Nevertheless, pairwise F_{st} between different clades of baboon *Trichuris* identified in the phylogenetic analysis was calculated. The p-distance between the distinct clades identified in the phylogenetic analysis was calculated using MEGA v.6.1. [39].

Demography, time of divergence and TMRCA

For uniparentally inherited DNA, the time to the most recent common ancestor (TMRCA) in generations is equal to the population size (N). The effective population sizes for the populations of *T. trichiura* and *T. suis* were calculated using the formula $\Theta=2N_{eff}\mu$ where Θ (theta) is the genetic diversity of a population, μ is the mutation rate per gene and N_{eff} is the

329 effective population size. Theta (Θ) and the ancestral history were estimated from Genetree [45]
330 using a concatenated dataset of the two markers. First, sequences were aligned and imported
331 to Map modules in the SNAP workbench [46] to collapse the sequences to haplotypes excluding
332 sites which are indels and infinite site violations. Then, compatibility analysis using CladeEx
333 revealed incompatible sites which were removed [46]. The simulations were repeated 5 times
334 with 10 million runs with different random seeds to ensure convergence of the genealogies.
335 The mutation rate of *Caenorhabditis elegans* was used which is 1.6×10^{-7} per site per
336 generation [47] and found not to be significantly different from other free living nematodes [48].
337 To obtain the mutation rate per gene, the mutation rate per site was multiplied by the number
338 of nucleotides used (807 nt for both markers) giving 1.29×10^{-4} per gene per generation.

339

340 BEAST v.1.6.1 [49] was used to infer the phylogeny and the divergence time using the Bayesian
341 statistical framework and the concatenated dataset with *Trichuris* from pigs and humans as
342 monophyletic groups. Different mutation models were used and the final analysis was done
343 using strict molecular clock with a normal distributed substitution rate of 1.6×10^{-7} ($\pm 0.3 \times$
344 10^{-7}) based on the value of *C. elegans* [47]. The substitution model used here was Hasegawa-
345 Kishino-Yano (HKY) with gamma distribution as best to fit model based on AIC [41] in
346 jModelTest0.1.1 [40]. Yule prior, which is suitable for datasets that combine different species,
347 was used as a tree prior with a random starting tree. Markov chain Monte Carlo (MCMC)
348 chains were run for iterations with a burn in value of 1000. Tracer v.1.6 was used to analyse
349 log files and to check whether the MCMC chains were sufficient by recording effective
350 sample size values to be above 200, which was the case for all the parameters. The three log
351 files of the three independent runs were combined using log combiner v1.6.1 [49]. Tree
352 Annotater v1.6.1. [49] was used to summarize samples from the posterior on maximum

353 credibility tree and the posterior probability limit set to 0.5. Figtree v1.3.1 [49] was used to
354 depict the tree.

355

356 Divergence time was also estimated for the human and pig *Trichuris* populations using IMa2
357 based on an isolation and migration model [50] using the concatenated dataset. Priors used for
358 these data sets were: for *T. trichiura*, t (the upper bound of splitting time) = 295, q (upper
359 bound of population size) = 750, while for *T. suis* t = 40, q = 100. For both data sets, HKY
360 substitution model was used, no migration between populations after splitting was assumed
361 (m = 0), 20 Markov chains with geometric heating scheme (the first and second heating
362 parameters were 0.96 and 0.90, respectively) and 10^6 burn-in steps with 10^5 sampling
363 genealogies were used. Three independent runs were conducted with different seed numbers
364 to assess the convergence.

365

366 **Acknowledgements**

367 We gratefully acknowledge the following for providing worm specimens: J. Urban, United
368 States Department of Agriculture, USA; H. Namwanje, Vector Control Division, Ministry of
369 Health, Kampala, Uganda; S. Nissen, I-H. Poulsen, A. Andreassen, H.H. Petersen University
370 of Copenhagen, Denmark; T. J. C. Anderson and the Southwest National Primate Research
371 Center at Texas Biomedical Research Institute, San Antonio, Texas; Amy Beierschmitt at the
372 Behavioural Science Foundation, St. Kitts. P.J.C. was supported by Wellcome Trust grant
373 088862/Z/09/Z. This work is supported by a Danish Agency for Science, Technology and
374 Innovation grant awarded to P.N. (FSS 09-066344).

375

376 **Author Contributions**

P.N. and M.B.F.H. conceived and designed the study with inputs from M.B and D.T.J.L; J.K., A.L.W., M.F.B., P.J.C., X.Q.Z. provided essential material, M.B.F.H. and A.A.G conducted the molecular work. M.B.F.H. and P.N. analysed the data. M.B.F.H. and P.N. wrote the paper with inputs from M.B and D.T.J.L. All authors critically read and approved the final manuscript.

Competing financial interests: The authors declare no competing financial interests.

References

1 Pullan, R. L., Smith, J. L., Jasrasaria, R. & Brooker, S. J. Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. *Parasit. Vectors* **7**, 37 (2014).

2 Abee C. R., M., K., Tardif, S.D., Morris T. *Nonhuman Primates in Biomedical Research Diseases*. 2nd Edition edn, 243-244 (Academic Press, 2012).

3 Nansen, P. & Roepstorff, A. Parasitic helminths of the pig: factors influencing transmission and infection levels. *Int. J. Par.* **29**, 877-891 (1999).

4 Roepstorff, A., Mejer, H., Nejsum, P. & Thamsborg, S. M. Helminth parasites in pigs: new challenges in pig production and current research highlights. *Vet. Par.* **180**, 72-81 (2011).

5 Archie, E. A., Luikart, G. & Ezenwa, V. O. Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends Ecol. Evol.* **24**, 21-30, (2009).

6 Gilabert, A. & Wasmuth, J. D. Unravelling parasitic nematode natural history using population genetics. *Trends Parasitol.* **29**, 438-448, (2013).

7 Betson, M., Nejsum, P., Bendall, R. P., Deb, R. M. & Stothard, J. R. Molecular epidemiology of ascariasis: a global perspective on the transmission dynamics of *Ascaris* in people and pigs. *J. Inf. Dis.* **210**, 932-941, (2014).

8 Ramesh, A. *et al.* The complete mitochondrial genome sequence of the filarial nematode *Wuchereria bancrofti* from three geographic isolates provides evidence of complex demographic history. *Mol. Biochem. Parasitol.* **183**, 32-41, (2012).

9 Cox, F. E. History of human parasitology. *Cli. Microbiol. Rev.* **15**, 595-612 (2002).

10 Goncalves, M. L., Araujo, A. & Ferreira, L. F. Human intestinal parasites in the past: new findings and a review. *Mem. Inst. Oswaldo Cruz* **98 Suppl 1**, 103-118 (2003).

11 Mitchell D, P. The origins of human parasites: Exploring the evidence forendoparasitism throughout human evolution. *Int. J. Paleopathol.* **3**, 191–198 (2013).

12 Buikstra, J. E., C.A. Roberts. *A global history of paleopathology: pioneers and prospects*. (Oxford University Press, 2012).

13 Ghai, R. R. *et al.* Hidden population structure and cross-species transmission of whipworms (*Trichuris* sp.) in humans and non-human primates in Uganda. *PLoS Neg. Trop. Dis.* **8**, e3256 (2014).

14 Liu, G. H. *et al.* Mitochondrial and nuclear ribosomal DNA evidence supports the existence of a new *Trichuris* species in the endangered francois' leaf-monkey. *PloS one* **8**, e66249 (2013).

15 Kuhner, M. K. Coalescent genealogy samplers: windows into population history. *Trends Ecol. Evol.* **24**, 86-93 (2009).

16 Amills, M., Ramírez, O., Galman-Omitogun, O. & Clop, A. Domestic pigs in Africa. *Afr. Archaeol. Rev.* **30**, 73-82, doi:10.1007/s10437-012-9111-2 (2012).

17 Ramirez, O. *et al.* Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig breeds. *Mol. Biol. Evol.* **26**, 2061-2072, doi:10.1093/molbev/msp118 (2009).

18 Atuhaire, D. K. *et al.* Molecular characterization and phylogenetic study of African swine fever virus isolates from recent outbreaks in Uganda (2010-2013). *Virol. J.* **10**, (2013).

19 Barrett, L. G., Thrall, P. H., Burdon, J. J. & Linde, C. C. Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends Ecol. Evol.* **23** (2008).

20 Giuffra, E. *et al.* The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* **154**, 1785-1791 (2000).

21 Groenen, M. A. *et al.* Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393-398 (2012).

22 Burden, D. J. & Ginnivan, M. J. The destruction of pig helminth ova and larvae in a slurry treatment process. *Vet. Rec.* **103**, 373-375 (1978).

23 Rosenthal, B. M. *et al.* Human dispersal of *Trichinella spiralis* in domesticated pigs. *Infect. Genet. Evol.* **8**, 799-805 (2008).

24 Meekums, H. *et al.* A genetic analysis of *Trichuris trichiura* and *Trichuris suis* from Ecuador. *Parasit. Vectors* **8**, 168 (2015).

25 Liu, W. *et al.* Human remains from Zhirendong, South China, and modern human emergence in East Asia. *Proc. Natl. Acad. Sci.* **107**, 19201-19206 (2010).

26 Beyin, A. Upper Pleistocene human dispersals out of Africa: A Review of the current state of the debate. *Int. J. Evol. Biol.* **2011**, 615094 (2011).

27 Dillehay, T. D. The late Pleistocene cultures of South America. *Evol. Anthr.* **7**, 206-216 (1999).

28 Finlayson, C. Biogeography and evolution of the genus *Homo*. *Trends Ecol. Evol.* **20**, 457-463 (2005).

29 Holterman, M. *et al.* Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown Clades. *Mol. Biol. Evol.* **23**, 1792-1800 (2006).

30 Brinkkemper, O. & van Haaster, H. Eggs of intestinal parasites whipworm (*Trichuris*) and mawworm (*Ascaris*): Non-pollen palynomorphs in archaeological samples. *Rev. Palaeobot. Palynol.* **186**, 16-21 (2012).

456 31 Cawthon Lang KA. *Primate Factsheets: Vervet (Chlorocebus) Taxonomy,*
457 *Morphology, & Ecology* (2006). Available from
458 http://pin.primate.wisc.edu/factsheets/entry/vervet (Accessed 30th June 2015).
459 32 Ricklefs, R. E. Host-pathogen coevolution, secondary sympatry and species
460 diversification. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 1139-1147 (2010).
461 33 Cavallero, S., Snabel, V., Pacella, F., Perrone, V. & D'Amelio, S. Phylogeographical
462 studies of *Ascaris* spp. based on ribosomal and mitochondrial DNA sequences. *PLoS*
463 *Neg. Trop. Dis.* **7**, e2170 (2013).
464 34 Nejsum, P., Grondahl, C. & Murrell, K. D. Molecular evidence for the infection of zoo
465 chimpanzees by pig *Ascaris*. *Vet. Par.***139**, 203-210 (2006).
466 35 Leles D, R. K., Fugassa M, Ferreira LF, Iniguez AM, Araújo A. A parasitological
467 paradox: why is ascarid infection so rare in the prehistoric Americas? *J. Archaeol. Sci.*
468 **37**, 1510-1520 (2010).
469 36 Nissen, S. *et al.* Genetic analysis of *Trichuris suis* and *Trichuris trichiura* recovered
470 from humans and pigs in a sympatric setting in Uganda. *Vet. Par.* **188**, 68-77 (2012).
471 37 Lu, G. & Moriyama, E. N. Vector NTI, a balanced all-in-one sequence analysis suite.
472 *Brief. Bioinform.* **5**, 378-388 (2004).
473 38 Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis
474 program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**, 95-98 (1999).
475 39 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular
476 Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725-2729 (2013).
477 40 Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253-1256
478 (2008).
479 41 Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automat.*
480 *Contr.* **19**, 716-723 (1974).

42 Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254-267 (2006).

43 Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452 (2009).

44 Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10** (2010).

45 Griffiths, R. C. & Tavare, S. Sampling theory for neutral alleles in a varying environment. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **344**, 403-410 (1994).

46 Aylor, D. L., Price, E. W. & Carbone, I. SNAP: Combine and Map modules for multilocus population genetic analysis. *Bioinformatics* **22**, 1399-1401 (2006).

47 Denver, D. R., Morris, K., Lynch, M., Vassilieva, L. L. & Thomas, W. K. High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. *Science* **289**, 2342-2344 (2000).

48 Molnar, R. I., Bartelmes, G., Dinkelacker, I., Witte, H. & Sommer, R. J. Mutation rates and intraspecific divergence of the mitochondrial genome of *Pristionchus pacificus*. *Mol. Biol. Evol.* **28**, 2317-2326 (2011).

49 Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

50 Hey, J. & Nielsen, R. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747-760 (2004).

51 Hansen, T. V., Thamsborg, S. M., Olsen, A., Prichard, R. K. & Nejsum, P. Genetic variations in the beta-tubulin gene and the internal transcribed spacer 2 region of *Trichuris* species from man and baboons. *Parasit. Vectors* **6**, 236 (2013).

507 **Table 1**. A summary of the number of *Trichuris* isolates, the host from which samples were

508 recovered, the country of origin and sampling location(s).

Host (host numbers)	Country (number of samples)	Sampled localities in each country (number of samples)	Reference
Domesticated pigs, *Sus domesticus* (10)	Uganda (18)	Villages ranged 30 Km apart in south west Kabale district (18)	[36]
Domesticated pigs, *Sus domesticus* (5)	China (14)	Guangdong Province (3), Fujian Province (3), Chongqing Municipality (4), Hunan Province (4)	This study
Domesticated pigs, *Sus domesticus* (2)	Denmark (10)	Experimentally infected pigs with local strains of the parasite (10)	[36]
Domesticated pigs, *Sus domesticus* (2)	USA (10)	Experimentally infected pigs with local strains of the parasite (10)	[36]
Domesticated pigs, *Sus domesticus* (1)	Ecuador (7)	Quinidé and Súa Districts, Esmeraldas Province (7)	[24]
Humans (12)	Uganda (17)	Villages ranged in south west of Kabale district (17)	[36]
Human (1)	China (2)	Zhanjiang, Guangdong Province (2)	[51]
Humans (3)	Ecuador (12)	Quinidé and Súa Districts, Esmeraldas Province (12)	[24]
Baboons, *Papio hamadryas* (5)	Denmark (25)	Copenhagen Zoo (12), Knuthenborg Park (13)	[51]
Baboons, *Papio anubis* (2)	USA (9)	Southwest National	[51]
Baboon, *Papio anubis/P. cynocephalus* (1)	USA (2)	Primate Research Center	
Baboon, *Papio anubis/ P. hamadryas* (1)	USA (1)	(SNPRC) Texas (13)	
African Green Monkey, *Chlorocebus sabaeus* (4)	Saint Kitts (11)	Feral population	This Study

509

510 Table 2. Pairwise estimations of population differentiation (F_{st}) between populations of *T. suis*

511 and *T. trichiura* for the *nad*1 gene (below the diagonal) and the *rrn*L gene (above the

512 diagonal). Level of significance is based on 10,000 permutations.

	T. suis DK	*T. suis* Uganda	*T. suis* China	*T. suis* USA
T. suis DK		0. 916***	0.542***	0.000
T. suis Uganda	0.997***		0. 542***	0.916***
T. suis China	0.942***	0.876***		0. 542*
T. suis USA	0.016	0.995***	0.939***	
	T. trichiura Uganda	*T. trichiura* China	*T. trichiura* Ecuador	
T. trichiura Uganda		0.979**	0.932***	
T. trichiura China	0.984**		0.551*	
T. trichiura Ecuador	0.967***	0.778*		

513

514 *P<0.05, **P<0.01, ***P<0.001

515

Figure 1. Phylogenetic relationship between different *Trichuris* populations inferred by Neighbor Joining (NJ) tree based on the *rrn*L gene and the Tamura-Nei with gamma distribution model. Seven major clades were identified and are indicated by different colors. *T. trichiura* from humans from Uganda clustered in one clade together with most *Trichuris* from baboons and African green monkey and are indicated by the maroon color (■). Seven *Trichuris* from baboons and one from African green monkey clustered in a distinct clade and are indicated by the red color (■). *T. trichiura* from China were distinct and are indicated by the green color (■) while worms from Ecuador are indicated by light green (■). The other three clades include *Trichuris* from pigs. *T. suis* populations from USA and Denmark clustered together and are indicated by the blue color (■), whereas *T. suis* from China and Uganda are indicated by pink (■) and purple (■), respectively. Sample key are: B: Baboon, H: Human, P: Pigs, Gm: African green monkey; US, USA; Ch, China; UG, Uganda; DK, Denmark (C for Copenhagen Zoo and K for Knuthenborg).

Figure 2. Neighbour-net network based on concatenated sequences of the *nad*1 and *rrn*L genes. The colors of the different populations are given in Figure 1. *T. suis* from Ecuador cluster with worms from China, USA and Denmark and most *Trichuris* from non-human primates cluster with *T. trichiura* from Uganda.

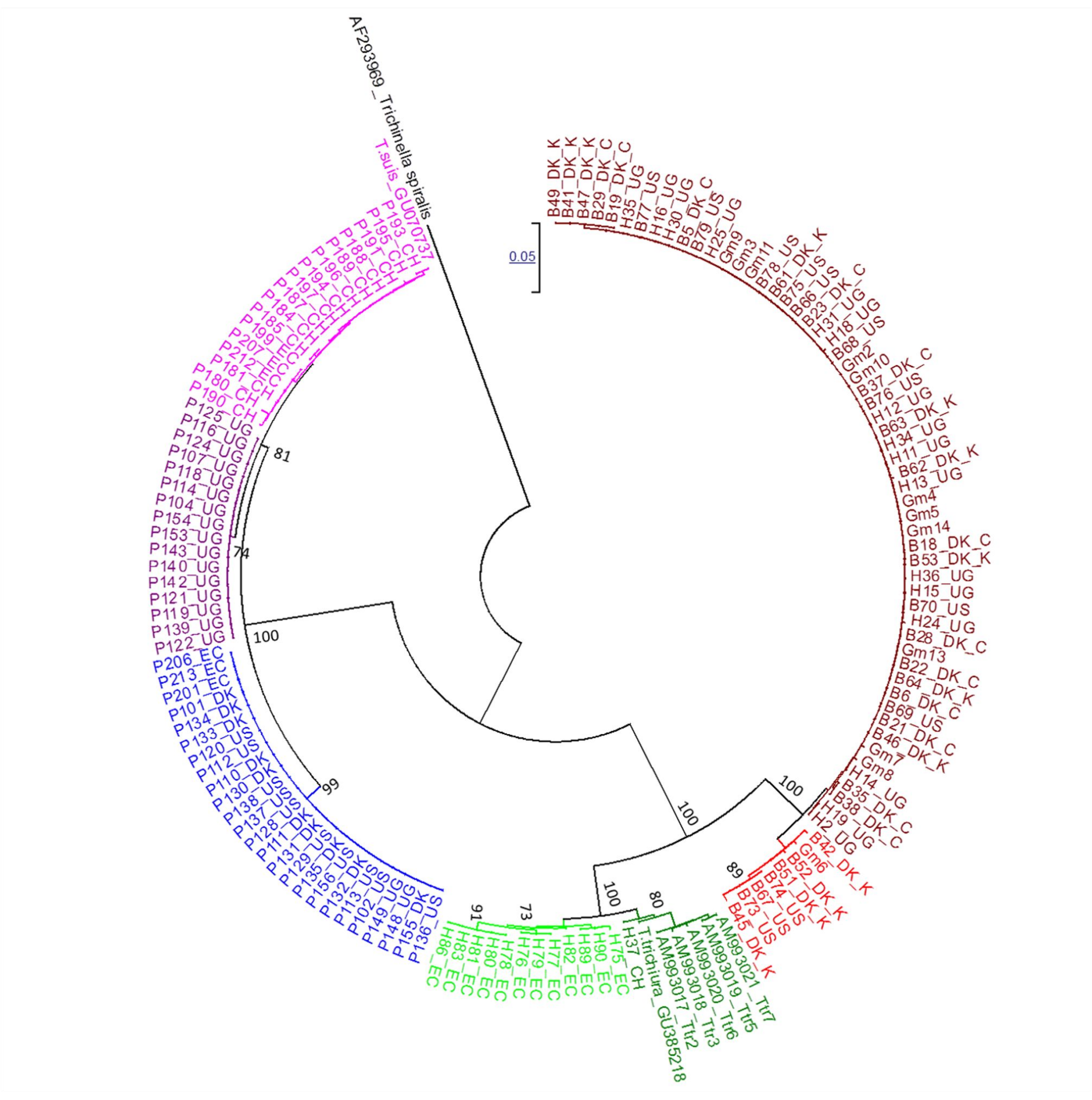
Figure 3. Bayesian phylogeny of the different primates and pig *Trichuris* populations. The different clades are indicated with the same colors used in the phylogenetic tree in Figure 1. All nodes are supported by >99% posterior support. Branch lengths are scaled in number of generations with the scale axis representing 200,000 generations. Median estimates of the divergence time are given at each node by number of generation.

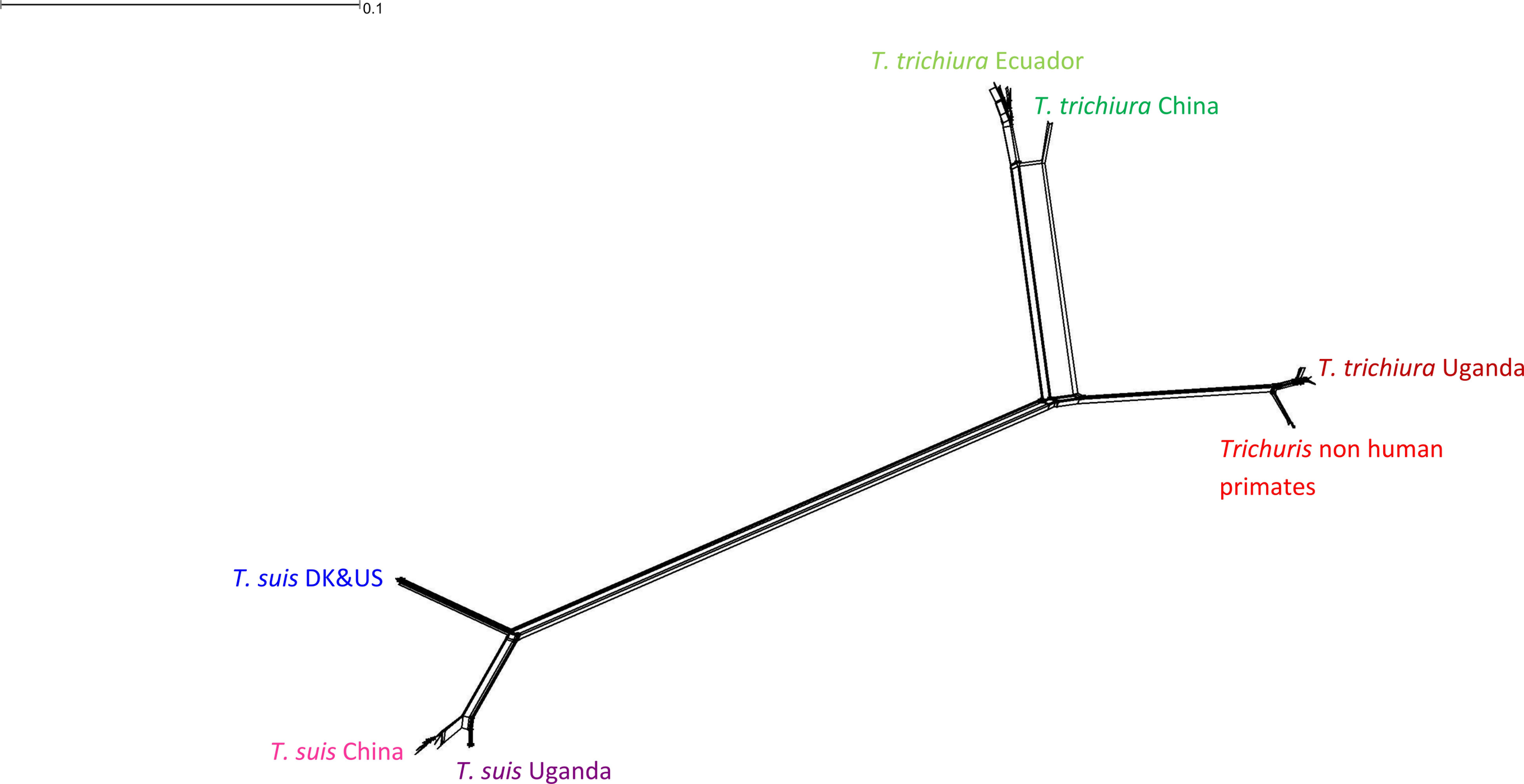
Figure 4. Summary of the evolutionary history showing possible dispersal routes of the human whipworms (dashed line) and pig whipworms (solid line) with the estimated time of divergence given as number of generations as estimated by Genetree. The native habitats in Africa of the different non-human primates (olive baboon (green), hamadrya baboons (purple), Yellow baboon (red) and African green monkey (yellow)) are indicated in the map. The origin of human *Trichuris* is believed to be in Africa where the parasite was transmitted to humans through early ancestors of primates while pigs evolved in China where it presumably acquired whipworms.

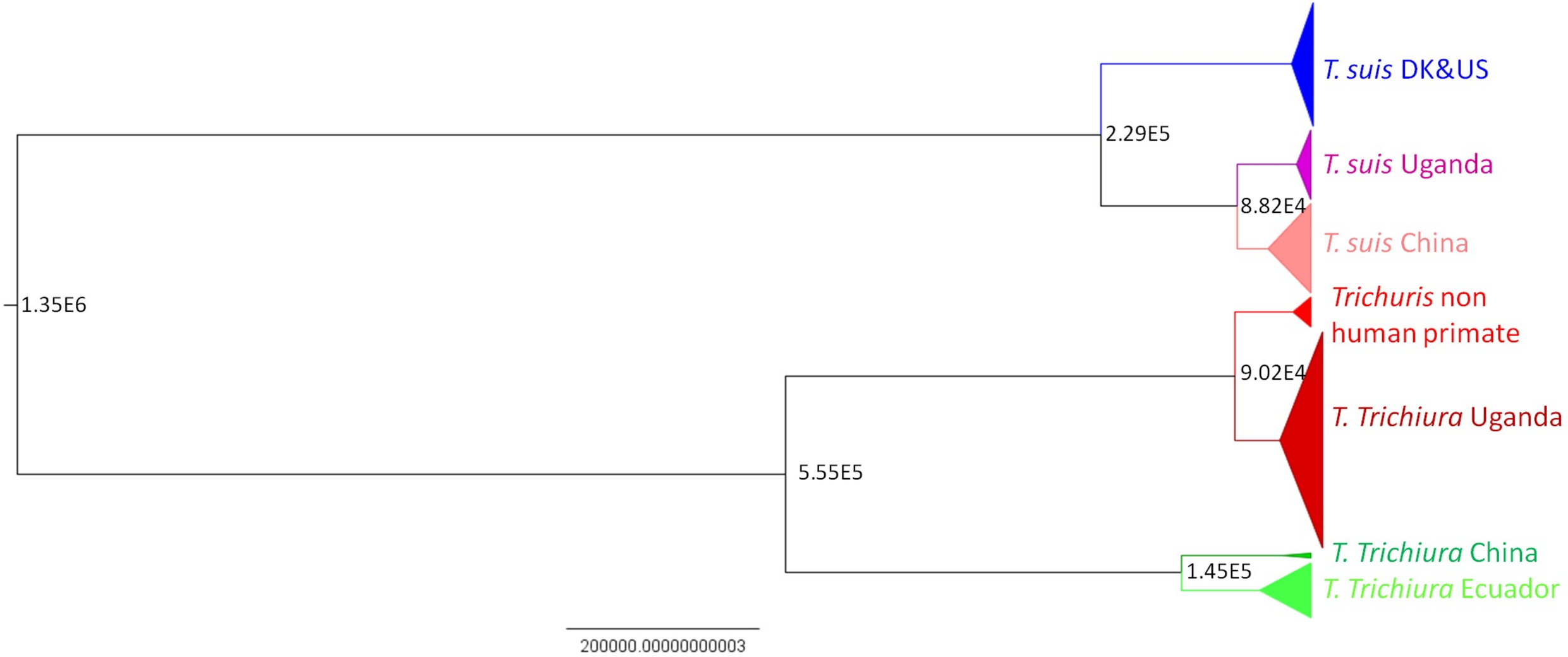
Supporting Information

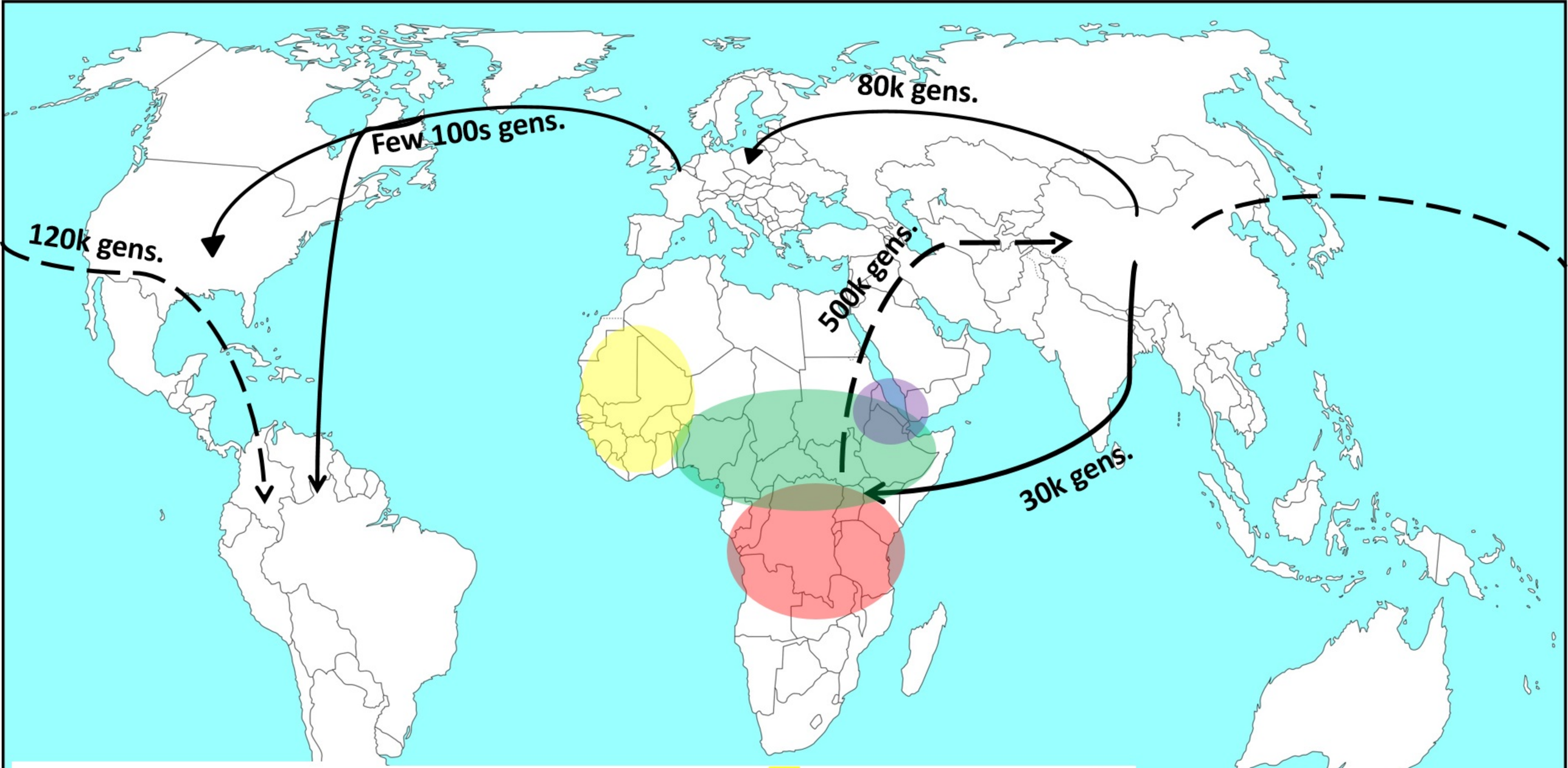
Supplementary Figure S1. The gene genealogy inferred by Genetree of (A) *T. suis* populations and (B) *T. trichiura* populations. Solid circles indicate mutations in the genealogy.

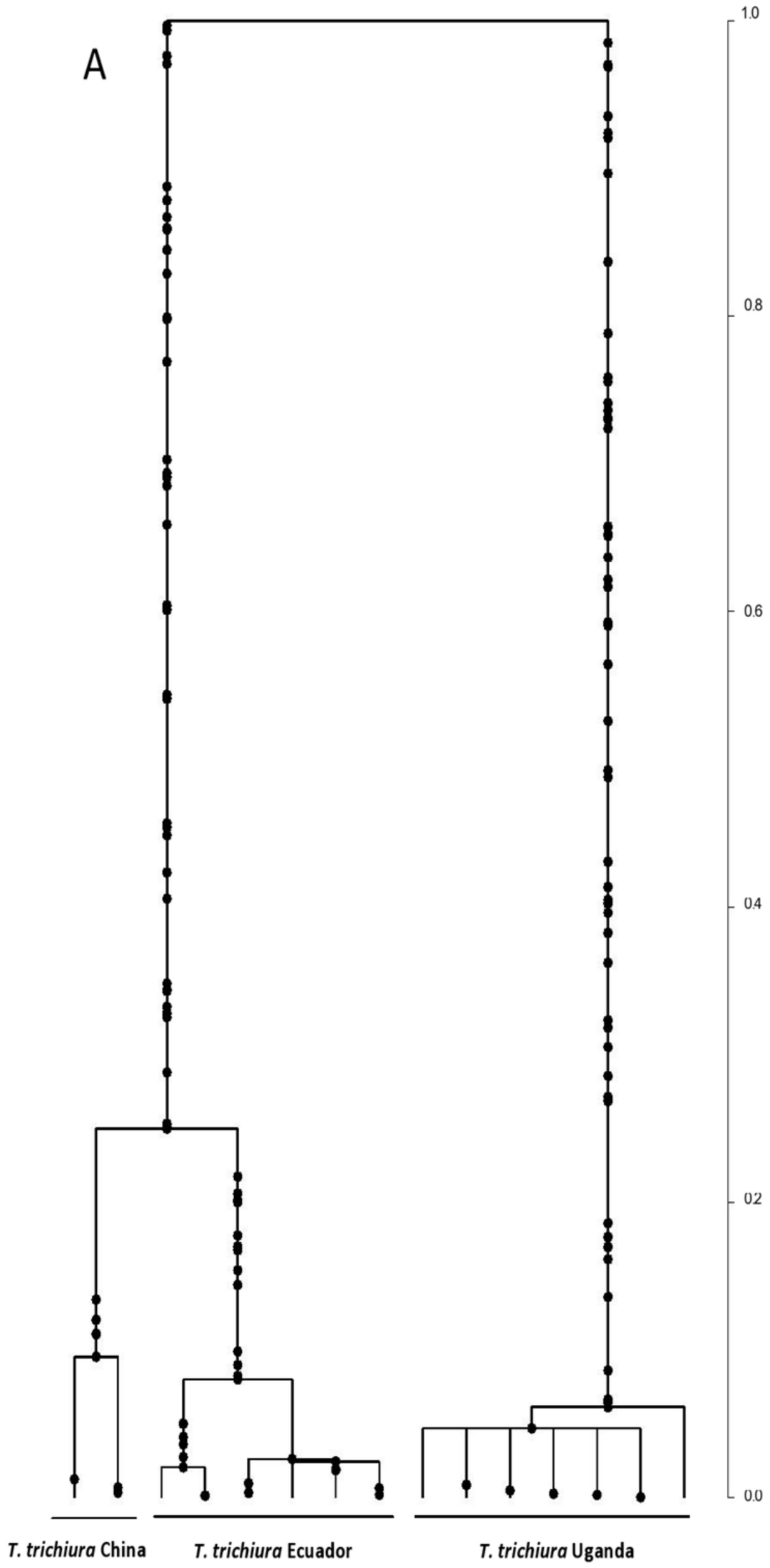
Supplementary Figure S2. Splitting time based on the isolation and migration model between (A) *T. suis* populations and (B) *T. trichiura* populations. The horizontal axis represents the number of generations since splitting which were estimated by dividing the splitting times between populations (t_0 and t_1) and time to most recent common ancestor (t_{mrca}) by the mutation rate per gene per generation (μ) while the vertical axis is the posterior probability density.

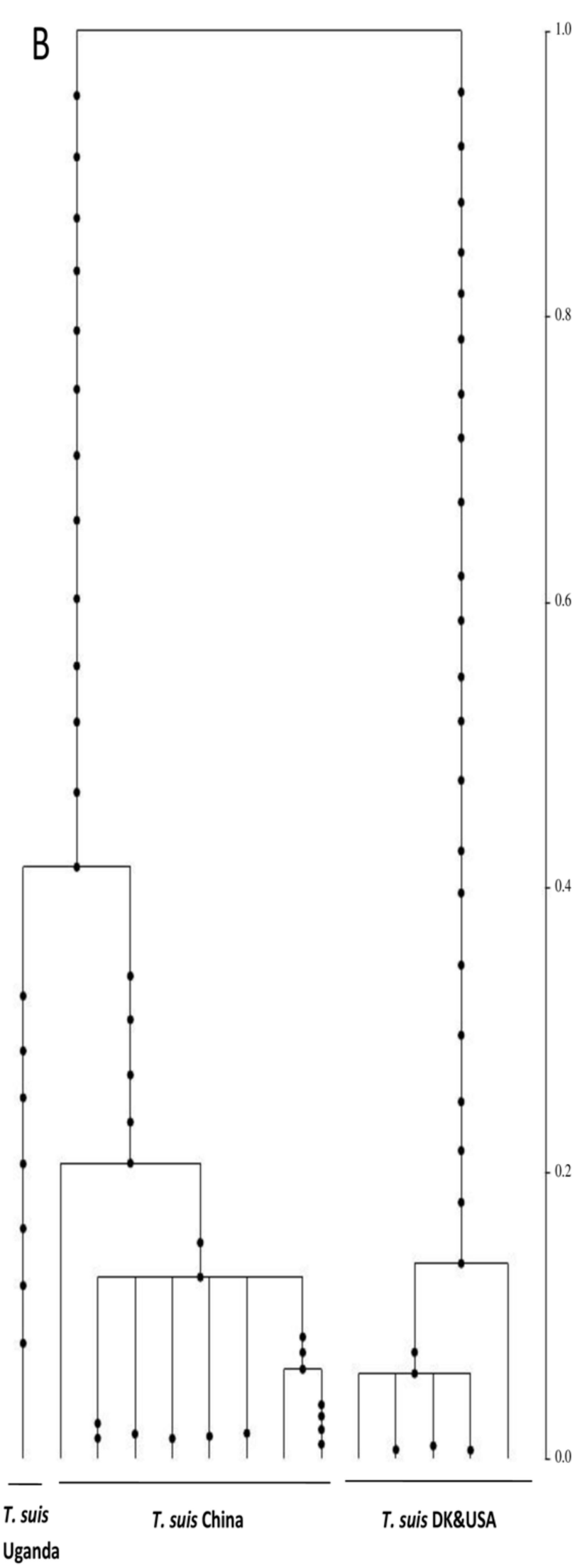

AF293969_Trichinella spiralis
T.suis_GU070737
P193_CH
P195_CH
P197_CH
P188_CH
P196_CH
P194_CH
P187_CH
P184_CH
P185_CH
P199_CH
P207_CH
P212_CH
P181_CH
P180_CH
P190_CH
P125_UG
P116_UG
P124_UG
P107_UG
P118_UG
P114_UG
P104_UG
P154_UG
P153_UG
P143_UG
P140_UG
P142_UG
P121_UG
P119_UG
P139_UG
P122_UG
P206_EC
P213_EC
P201_EC
P101_DK
P134_DK
P133_DK
P120_US
P112_US
P110_DK
P130_DK
P138_US
P137_US
P128_US
P111_DK
P131_DK
P129_US
P135_DK
P156_US
P132_DK
P113_US
P102_US
P149_UG
P148_US
P155_DK
P136_US
H86_EC
H83_EC
H81_EC
H80_EC
H78_EC
H79_EC
H76_EC
H77_EC
H82_EC
H89_EC
H90_EC
H75_EC
H37_CH
T.trichiura_GU385218
AM993017_Ttr2
AM993018_Ttr3
AM993020_Ttr6
AM993019_Ttr5
AM993021_Ttr7
B45_DK_K
B73_US
B67_US
B74_US
B51_DK_K
B52_DK_K
Gm6
B42_DK_K
H2_UG
H19_UG
B38_DK_C
B35_DK_C
H14_UG
Gm8
Gm7
B46_DK_K
B21_DK_C
B69_US
B6_DK_C
B64_DK_K
B22_DK_C
Gm13
B28_DK_C
H24_UG
B70_US
H15_UG
H36_UG
B53_DK_K
B18_DK_C
Gm14
Gm5
Gm4
H13_UG
B62_DK_K
H11_UG
H34_UG
B63_DK_K
H12_UG
B76_US
B37_DK_C
Gm10
Gm2
B68_US
H18_UG
H31_UG
B23_DK_C
B66_US
B75_US
B61_DK_K
B78_US
Gm11
Gm3
Gm9
H25_UG
B79_US
B5_DK_C
H30_UG
H16_UG
B77_US
H35_UG
B19_DK_C
B29_DK_C
B47_DK_K
B41_DK_K
B49_DK_K
0.05
81
74
100
99
100
100
89
80
73
91
100

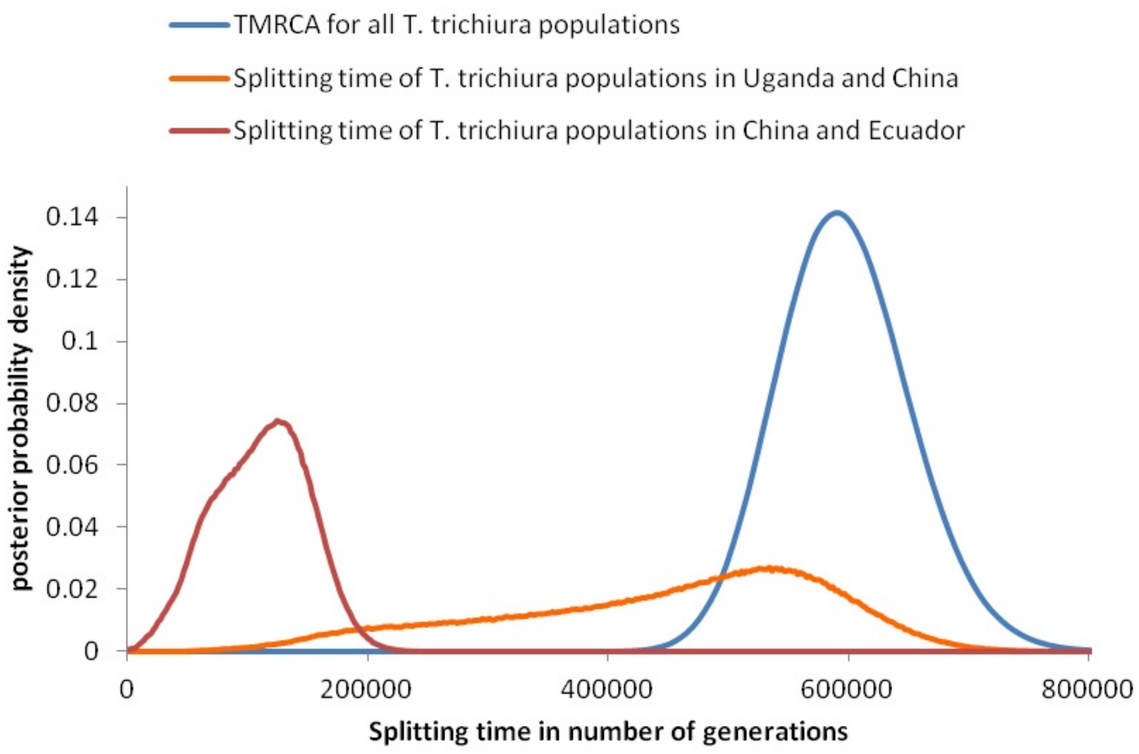

0.1
T. trichiura Ecuador
T. trichiura China
T. trichiura Uganda
Trichuris non human primates
T. suis DK&US
T. suis China
T. suis Uganda

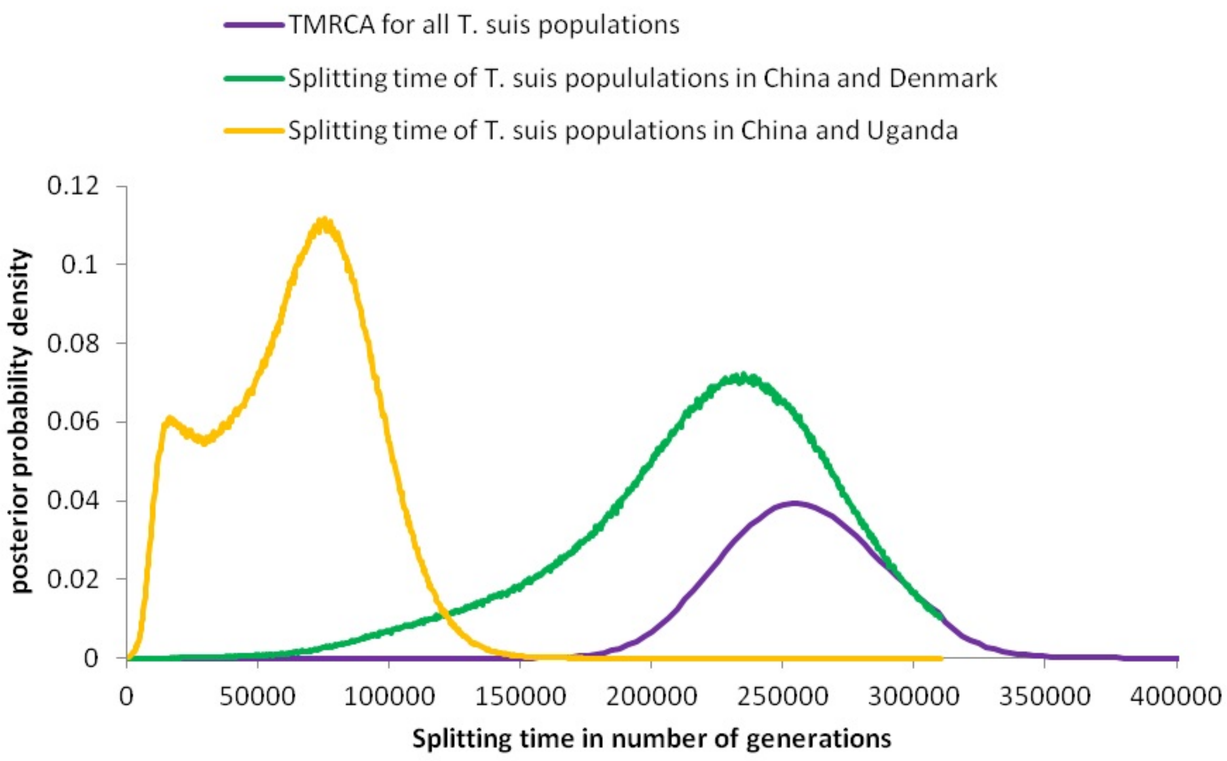

T. suis DK&US
2.29E5
T. suis Uganda
8.82E4
T. suis China
1.35E6
Trichuris non human primate
9.02E4
T. Trichiura Uganda
5.55E5
T. Trichiura China
1.45E5
T. Trichiura Ecuador
200000.00000000003


80k gens.
Few 100s gens.
120k gens.
500k gens.
30k gens.
Human whipworm
Pig whipworm
Green monkey native habitat
Olive baboon native habitat
Yellow baboon native habitat
Hamadryas baboons


A
1.0
0.8
0.6
0.4
0.2
0.0
T. trichiura China
T. trichiura Ecuador
T. trichiura Uganda


B
1.0
0.8
0.6
0.4
0.2
0.0
T. suis Uganda
T. suis China
T. suis DK&USA


TMRCA for all T. trichiura populations
Splitting time of T. trichiura populations in Uganda and China
Splitting time of T. trichiura populations in China and Ecuador
posterior probability density
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
200000
400000
600000
800000
Splitting time in number of generations


TMRCA for all T. suis populations
Splitting time of T. suis populululations in China and Denmark
Splitting time of T. suis populations in China and Uganda
posterior probability density
0.12
0.1
0.08
0.06
0.04
0.02
0
0
50000
100000
150000
200000
250000
300000
350000
400000
Splitting time in number of generations