

# Genetic Analysis of Lice Supports Direct Contact between Modern and Archaic Humans

David L. Reed<sup>1,3\*</sup>, Vincent S. Smith<sup>2</sup>, Shaless L. Hammond<sup>3</sup>, Alan R. Rogers<sup>4</sup>, Dale H. Clayton<sup>3</sup>

**1** Florida Museum of Natural History, Dickinson Hall, University of Florida, Gainesville, Florida, United States of America, **2** Graham Kerr Building, DEEB, IBLS, University of Glasgow, Glasgow, Scotland, **3** Department of Biology, University of Utah, Salt Lake City, Utah, United States of America, **4** Department of Anthropology, University of Utah, Salt Lake City, Utah, United States of America

**Parasites can be used as unique markers to investigate host evolutionary history, independent of host data. Here we show that modern human head lice, *Pediculus humanus*, are composed of two ancient lineages, whose origin predates modern *Homo sapiens* by an order of magnitude (ca. 1.18 million years). One of the two louse lineages has a worldwide distribution and appears to have undergone a population bottleneck ca. 100,000 years ago along with its modern *H. sapiens* host. Phylogenetic and population genetic data suggest that the other lineage, found only in the New World, has remained isolated from the worldwide lineage for the last 1.18 million years. The ancient divergence between these two lice is contemporaneous with splits among early species of *Homo*, and cospeciation analyses suggest that the two louse lineages codiverged with a now extinct species of *Homo* and the lineage leading to modern *H. sapiens*. If these lice indeed codiverged with their hosts ca. 1.18 million years ago, then a recent host switch from an archaic species of *Homo* to modern *H. sapiens* is required to explain the occurrence of both lineages on modern *H. sapiens*. Such a host switch would require direct physical contact between modern and archaic forms of *Homo*.**

Citation: Reed DL, Smith VS, Hammond SL, Rogers AR, Clayton DH (2004) Genetic analysis of lice supports direct contact between modern and archaic humans. PLoS Biol 2(11): e340.

## Introduction

One of the most intensely debated topics in evolutionary biology pertains to the origin of modern *Homo sapiens*. The debate concerns the precise manner in which anatomically modern humans arose from archaic ancestors. Empirical studies tend to support one of two prominent models of human origins, the Recent African Replacement model (Stringer and Andrews 1988) or the Multiregional Evolution model (Wolpoff et al. 1994). The Recent African Replacement model, as originally proposed, suggests that modern humans arose from an archaic ancestor in Africa ca. 130,000 years ago, and then replaced archaic humans in Asia, Africa, and Europe without introgression between archaic and modern humans. The Multiregional Evolution model (as proposed by Wolpoff et al. [1994] and revisited by Wolpoff et al. [2000]) suggests that gene flow existed not only among populations of modern *Homo sapiens*, but also between modern *H. sapiens* and archaic forms of *Homo* (e.g., *Homo neanderthalensis* and *Homo erectus*), which led to some degree of regional continuity. Both models can be subdivided into many variants. There are two common variants of the Multiregional Evolution model. In one variant, the transition from archaic to modern humans occurs incrementally across a large geographic region (i.e., both within and outside Africa); in the other variant, the transition from archaic to modern humans arises first in Africa then spreads through gene flow outside of Africa. This latter variant is very similar to a Diffusion Wave model recently put forth by Eswaran (2002). Both types of models of human origins (the Recent African Replacement and Multiregional Evolution models) have been examined with both human fossil and genetic data, but no single model or variant has been supported by all the data.

Fossils provide the only source of data available for most species of archaic humans and are therefore crucial to understanding the origin of modern humans. Unfortunately, missing taxa and fragmentary fossils limit our ability to reconstruct human evolutionary history based solely on fossil data. Molecular (DNA sequence) data have provided additional insight into the recent evolutionary history of humans, but these data are limited mainly to extant human populations. Ancient DNA was recently sequenced from *H. neanderthalensis* (Krings et al. 1997, 1999, 2000) and a 24,000-year-old specimen of modern *H. sapiens* (Caramelli et al. 2003), but even these ancient DNA studies do not agree on hypotheses of modern human origins (Templeton 2002; Serre et al. 2004). Only a few ancient specimens have been examined molecularly, and additional sequences are slow to emerge. Furthermore, DNA may never be retrieved from some specimens because it is difficult, if not impossible, to liberate sequenceable DNA from poorly preserved (Krings et al. 1997) or very old (Paabo and Wilson 1991) fossil material.

Received May 11, 2004; Accepted August 3, 2004; Published October 5, 2004  
DOI: 10.1371/journal.pbio.0020340

Copyright: © 2004 Reed et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abbreviations: COL, cytochrome oxidase subunit I; Cytb, cytochrome b; HTLV, human T-cell leukaemia/lymphoma virus; LRT, likelihood ratio test; ML, maximum likelihood; MRCA, most recent common ancestor; mtDNA, mitochondrial DNA; MY, million years; MYA, million years ago;  $N_{ef}$ , effective female population size; NW, New World; WW, worldwide

Academic Editor: Nick Barton, University of Edinburgh

\*To whom correspondence should be addressed. E-mail: dreed@flmnh.ufl.edu

Therefore, the degree to which we can reconstruct human evolutionary history depends, in part, upon additional types of data.

Several recent studies have inferred portions of human evolutionary history from the evolutionary history of their parasites (Chan et al. 1992; Ho et al. 1993; Ong et al. 1993; Escalante et al. 1998; Ashford 2000; Leal and Zlotoff 2000; Hoberg et al. 2001). Parasites can be a powerful tool for reconstructing host evolutionary history because they provide data that are independent of host data. For example, human papillomaviruses (Chan et al. 1992; Ho et al. 1993; Ong et al. 1993), tapeworms (Hoberg et al. 2001), and malarial parasites (Escalante et al. 1998) each have evolutionary origins in Africa, consistent with most human fossil and molecular data. Human T-cell leukaemia/lymphoma virus (HTLV) sequences show that most human viral strains are closely related to those of Old World apes and monkeys (Leal and Zlotoff 2000). In contrast, some Native American strains of HTLV have closer affinities to viral strains from Asian primates, suggesting a dual origin for this virus in humans (Leal and Zlotoff 2000 and references therein). Ashford (2000) recently reviewed the use of parasites as markers of human evolutionary history, pointing out that five parasites of humans (lice, tapeworms, follicle mites, a protozoan, and bedbugs) have closely related taxonomic pairs that suggest periods of host geographic isolation. Unfortunately, none of these five pairs has been studied rigorously with the primary goal of inferring host evolutionary history. Of these parasites, the ones most likely to provide the greatest insight into human evolutionary history are those that are known to have had a long-term coevolutionary association with their hosts, such as lice (*Insecta: Phthiraptera*) (Page 2003).

Lice are obligate parasites of mammals or birds that complete their entire life cycle on the body of the host; they cannot survive more than a few hours or days off the host (Buxton 1946). Mammal lice are closely tied to their hosts in both ecological (Reed and Hafner 1997) and evolutionary (Hafner et al. 1994) time. The lice found on primates are quite host specific, with most species occurring only on a single species of host (Durden and Musser 1994). Host specificity is reinforced by the fact that primate lice require direct physical contact between hosts for transmission (Buxton 1946; Durden 2001; Canyon et al. 2002; Burgess 2004). Host specificity often goes hand in hand with long-term coevolutionary patterns between hosts and parasites (Page 2003), making primate lice excellent candidates for inferring host evolutionary history. Humans are parasitized by two species of lice: head/body lice (*Pediculus humanus*), the focus of this paper, and pubic lice (*Phthirus pubis*), which serve as a phylogenetic outgroup in this study. *P. humanus* is found in two forms (head and body lice) that are morphologically similar, but ecologically distinct. Body lice live primarily in clothing and move onto the skin to feed once or twice a day. Head lice are confined to the scalp and feed more frequently. Body lice vector the bacteria responsible for epidemic typhus, trench fever, and relapsing fever; head lice are not known to vector any agent of human disease under natural conditions (Buxton 1946).

Recent molecular work by Leo et al. (2002) showed that, despite the ecological differences between head and body lice, the two forms are not genetically distinct. Kittler et al. (2003) confirmed this finding but also discovered two deeply

divergent clades within *P. humanus* that are uncorrelated with the head and body louse forms. The divergent clades of lice stand in contrast to mitochondrial sequence data from extant human populations, which coalesce to a single lineage very rapidly. The shallow coalescence in human mitochondrial sequence data is likely the result of a recent population bottleneck and subsequent population expansion (Rogers and Harpending 1992), which obscures much of the evolutionary history of humans prior to the bottleneck. The deep divergences within *P. humanus* have the potential to reveal aspects of human evolutionary history that cannot be recovered from human DNA markers.

We reconstructed the evolutionary history of *P. humanus* and several outgroup taxa using both morphological and molecular data. First, we used louse morphological data to test for patterns of cospeciation between primate lice and their hosts. Then we collected molecular data from a subset of the same taxa and calculated divergence dates for nodes in the louse phylogeny. This broad phylogenetic approach allowed us to date the origin of the human louse, *P. humanus*, and to date the two divergent lineages within the species. Finally, we collected population genetic data for *P. humanus* to compare with population-level characteristics of extant humans. Taken together, our phylogenetic and population-level data provide a well-resolved picture of the evolutionary history of *P. humanus*, which can be used to indirectly infer human evolutionary history. Specifically, we compared three distinct models of modern human origins (Recent African Replacement without Introgression, Multiregional Evolution, and Diffusion Wave) to see which model best fits the data from human lice.

## Results

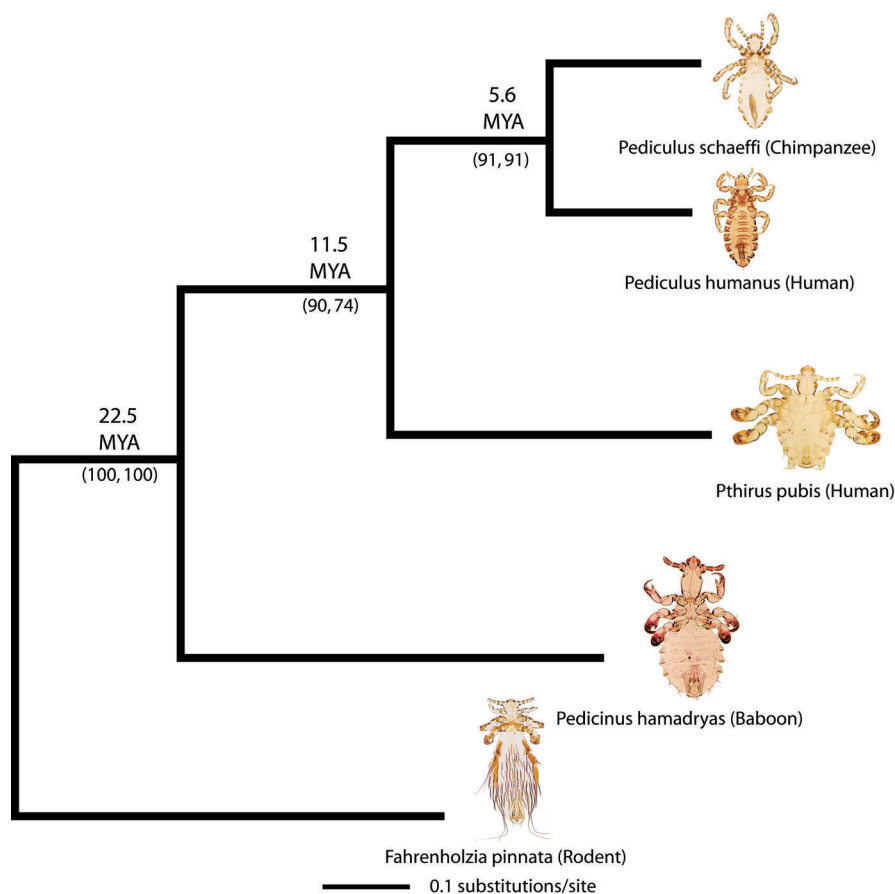
### Phylogenetic Analyses and Divergence Estimates

Both the morphological and molecular data sets produced a single phylogenetic relationship for the louse species in Figure 1. The phylogeny shows that *Pediculus* species on chimpanzees and humans are sister taxa, which together with *Phthirus* form a clade that is sister to *Pedicinus*, the most basal member of the ingroup (Figure 1). Bootstrap support for these relationships is high. Reconciliation analysis using Treemap v. 2.0 (M. A. Charleston and R. D. M. Page, software distributed by authors) revealed significant congruence ( $p < 0.01$ ) between the louse and primate phylogenies, thus validating the assumption of cospeciation (Kittler et al. 2003). Reconciliation analysis using Treemap showed four cospeciation events and one host switch. One particular node of cospeciation determined that as cercopithecoid and hominoid primates diverged 20–25 million years ago (MYA) (Benefit 1993; Leakey et al. 1995), *Pedicinus* diverged from the lineage leading to *Pediculus* and *Phthirus*. Since the nodes of cospeciation in congruent host and parasite trees are contemporaneous, the louse tree can be calibrated using the host tree.

We used the date of  $22.5 \pm 2.5$  MYA to calibrate the split between *Pedicinus* and *Phthirus* + *Pediculus* in the louse tree. This, in turn, yielded a divergence time of 11.5 MYA for the *Phthirus*/*Pediculus* split and 5.6 MYA for the split between *Pediculus schaeffi* and *P. humanus* (Table 1). Our estimated divergence between chimp and human lice (5.6 MYA) is strikingly similar to the 5.5 MYA estimates for the chimp/

**Figure 1.** Phylogeny of Primate Lice from Morphological and Molecular Data

The phylogeny is a strict consensus of morphology and a 1,525-bp fragment of COI and Cytb. Branch lengths were determined from the molecular data. Numbers in parentheses are bootstrap values from molecular and morphological data, respectively. Divergence dates are direct estimates from mtDNA data (see text). Louse images from light microscopy were taken by VSS.  
DOI: 10.1371/journal.pbio.0020340.g001



human divergence based on both mitochondrial and nuclear sequence data (Stauffer et al. 2001). To test the original calibration date of 22.5 MYA, we used the molecular estimate of the chimp/human split (5.5 MYA; Stauffer et al. 2001) to reverse calibrate the louse tree. This younger calibration point resulted in divergence estimates that were nearly identical to those from the previous calibration. For example, the 5.5 MYA calibration resulted in an estimated divergence of 22.65 MYA for the split between *Pedicinus* and *Pthirus* + *Pediculus*. Estimates of divergence time error were calculated from bootstrapped data sets (Table 1). Other studies have

shown that louse mitochondrial DNA (mtDNA) sequences evolve at a rate two to three times faster than that of host sequence rates (Page 1996; Page et al. 1998). The lice in this study are evolving at ca. 2.3 times the rate of their primate hosts, when nucleotide substitutions are estimated under a best-fit model of sequence evolution.

Phylogenetic analysis revealed two divergent clades within *P. humanus* (6% uncorrected sequence divergence for cytochrome oxidase subunit I [COI] and cytochrome *b* [Cytb]). One of the two lineages in our data set is worldwide (WW) in distribution (Figure 2, Worldwide clade), contains both head

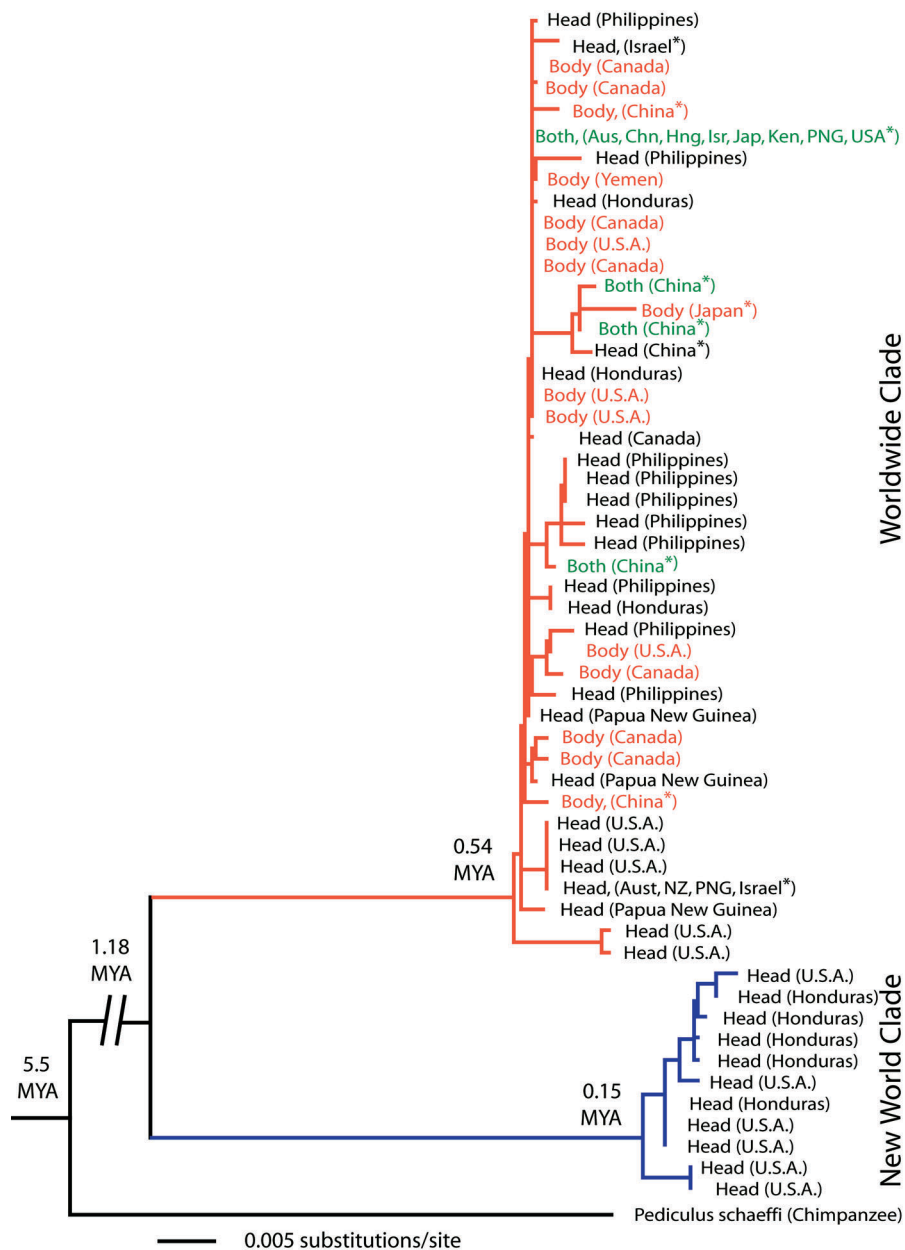
**Table 1.** Mean ( $\pm$  Standard Deviation), Minimum, and Maximum Estimates of Divergence Times (in Millions of Years) of Louse Lineages from 100 Bootstrapped Data Matrices

Node	Mean $\pm$ SD (MYA)	Minimum–Maximum (MYA)	Direct Estimates (MYA)
<i>Pedicinus</i>	18.91* $\pm$ 5.14	9.55*–35.03	22.65*
<i>Pthirus</i>	13.54 $\pm$ 2.21	9.43–19.33	11.51
<i>Pediculus</i>	7.08 $\pm$ 2.13	3.53–12.94	5.60
<i>P. humanus</i>	1.70 $\pm$ 1.10	0.59–4.75	1.18
WW clade	1.26 $\pm$ 1.18	0.04–4.25	0.54
NW clade	0.97 $\pm$ 1.24	0.04–4.74	0.15

Divergence dates were calculated without rate-smoothing methods because our data do not depart from the assumptions of a molecular clock (see text). Direct estimates of divergence times were calculated using the original data set (1,525-bp mtDNA) rather than the bootstrapped matrices. Trees were calibrated with a cercopithecoid/hominoid primate divergence of 22.5 MYA (from fossil evidence) except for the *Pedicinus* dates (asterisks), which were calibrated from the 5.5 MYA estimate of the human (*P. humanus*)-chimp (*P. schaeffi*) divergence date from molecular data (see text).  
DOI: 10.1371/journal.pbio.0020340.t001

**Figure 2.** Molecular Phylogeny of *P. humanus* from Geographically Diverse Human Populations

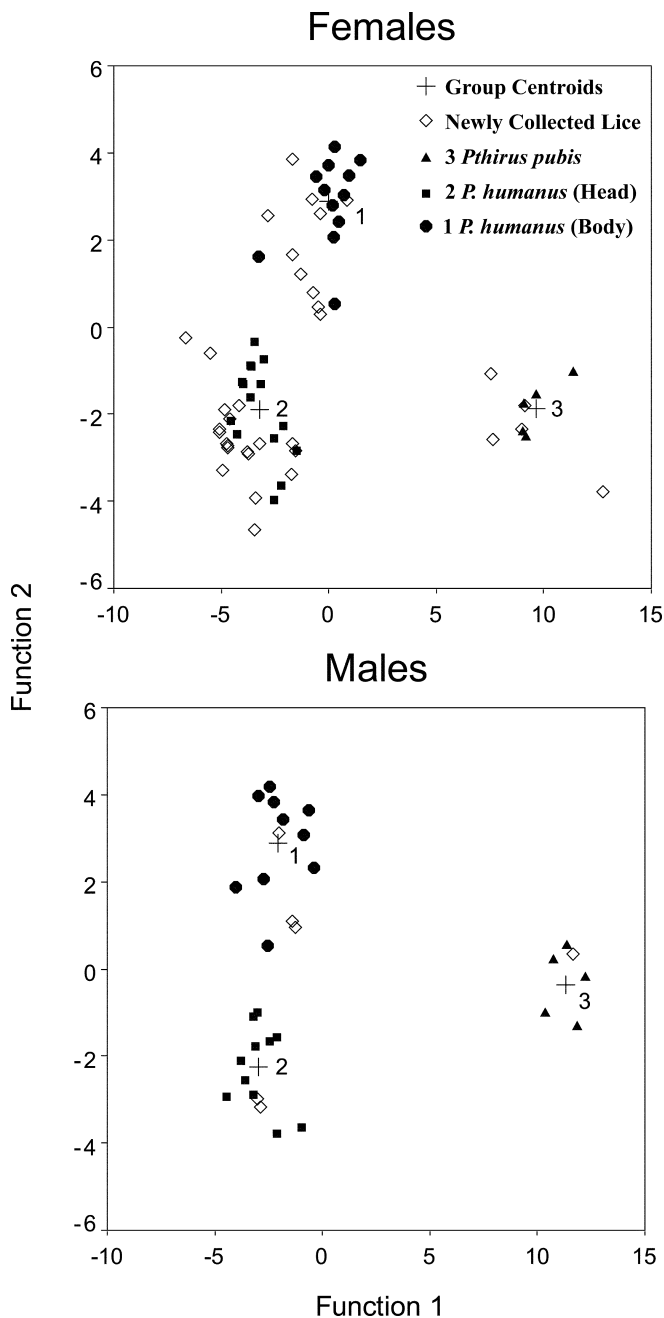
This species exhibits distinct “head” and “body” forms, which differ in ecology, and slightly in size. Head lice (black lettering) are smaller than body lice (red lettering) and are confined to the scalp, whereas body lice live primarily in clothing. Haplotypes shown in green were found in both head and body lice. There are no fixed genetic differences between the head and body forms, suggesting a lack of reproductive isolation, despite the fact that the two forms can be distinguished using discriminant function analysis of morphological data. These results are consistent with experimental data showing that head lice can transform morphologically into body lice within a few generations (Levene and Dobzhansky 1959). The Worldwide clade (red branches) shares a MRCA ca. 0.54 MYA and the geographically restricted New World clade (blue branches) has a much younger MRCA, ca. 0.15 MYA. Asterisks denote samples from Leo et al. (2002)  
DOI: 10.1371/journal.pbio.0020340.g002



and body louse forms, as determined by discriminant function analysis (Figure 3), and has a most recent common ancestor (MRCA) 0.54 MYA (Table 1). Even within this WW clade head and body lice are not reciprocally monophyletic, and a constraint to enforce such monophyly can be rejected using a Shimodaira-Hasegawa test ( $p < 0.01$ ) (Shimodaira and Hasegawa 1999). The other lineage (Figure 2, New World clade) is restricted to the New World (NW), contains only the head louse form, and has a MRCA only 0.15 MYA. The MRCA of all *P. humanus* was 1.18 MYA, which predates by a considerable margin the origin of modern *H. sapiens* based on mtDNA ( $\leq 0.20$  MYA; Cann et al. 1987; Vigilant et al. 1991; Ingman et al. 2000) as well as fossil evidence (0.15–0.16 MYA; White et al. 2003).

Our estimate of the age of the MRCA for *P. humanus* (1.18 MYA) is much older than that reported by Kittler et al. (2003), which was only 0.53 MYA based on mtDNA. Their estimate of

0.53 MYA was determined using a mtDNA sequence from a specimen of the chimp louse, *P. schaeffi*, that is quite aberrant when compared to other primate lice. Phylogenetic analysis of the Kittler et al. Cytb data (downloaded from GenBank), combined with our own data, shows that the Kittler et al. sequence for *P. schaeffi* is 40% divergent from *P. humanus* and 40% divergent from our own sequence of *P. schaeffi*. Phylogenetic analysis places their specimen of *P. schaeffi* outside all other primate lice and even outside the rodent louse (Figure 4), whereas our specimen of *P. schaeffi* is sister to *P. humanus*, based on both morphology and molecular data. We think that the Kittler et al. specimen has been attributed to the species *P. schaeffi* in error. In contrast to the mitochondrial data reported by Kittler et al. (2003), our analysis of their nuclear elongation factor (EF1-alpha) sequences produces a MRCA for *P. humanus* that is ca. 2 MYA. Similarly, 18S rRNA sequences for *P. humanus* from

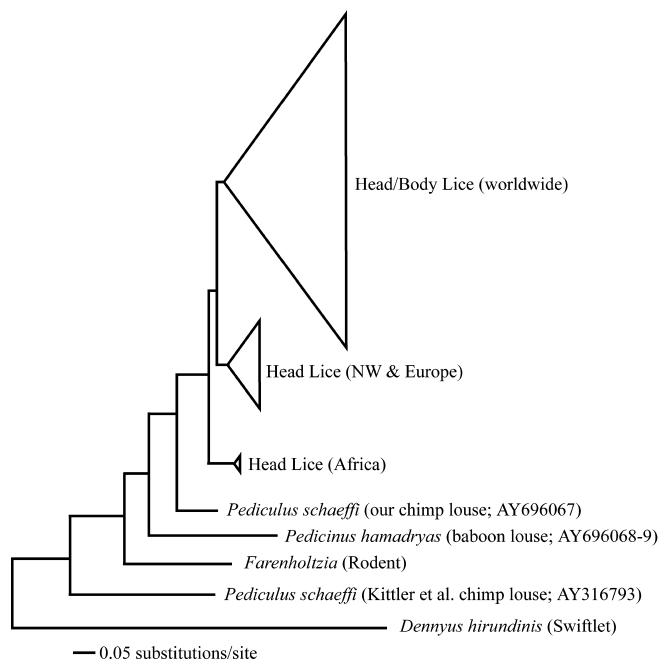


**Figure 3.** Plot of the First and Second Canonical Discriminant Functions for Specimens of Adult Head/Body Lice (*P. humanus*) and Pubic Lice (*Pthirus pubis*)

Solid points denote reference specimens of known identity acquired from museum collections. Unfilled points denote newly collected specimens used in the molecular analyses. In all cases the discriminant function analysis successfully classified each unknown case with a probability of >0.95.

DOI: 10.1371/journal.pbio.0020340.g003

Yong et al. (2003), combined with an 18S rRNA sequence from *P. schaeffi*, provide a MRCA for *P. humanus* that is ca. 2 MYA (for GenBank accession numbers, see Supporting Information). Together, these mitochondrial and nuclear markers support a MRCA for *P. humanus* greater than 1.18 MYA, which is an order of magnitude older than the MRCA for its human host.



**Figure 4.** Neighbor-Joining Tree Using a Best-Fit Model of Nucleotide Substitution (Tamura-Nei +  $\Gamma$ ) for a Combined Data Set of Cytb Sequences from Our Study and from Kittler et al. (2003)

The clades of *P. humanus* identified by Kittler et al. (2003) are nearly identical to those from our data, with the exception of their basal African clade, which was not represented in our data set. One clade contains both head lice and body lice and is WW in distribution. Another clade is comprised solely of head lice from the NW (our data) and Europe (samples from Kittler et al. 2003), and the most basal clade contains isolates 4, 18, and 33 from Kittler et al. (2003), which are head lice from Africa. The size of the triangles representing the three clades are proportional in size to the number of taxa within the clade. This phylogeny is rooted with a divergent louse, *Dennyus hirundinis*, which is a bird louse in the suborder *Amblycera*. Note the placement of the Kittler et al. (2003) specimen of *P. schaeffi*, which falls outside all other primate lice and the rodent louse *Fahrenholtzia*.

DOI: 10.1371/journal.pbio.0020340.g004

### Population Genetic Analysis of *P. humanus*

We can calculate an expected date of mitochondrial coalescence for *P. humanus* if we assume for the moment that the entire population of lice mated at random (i.e., panmixia). The estimate of expected coalescence is based on the effective female population size ( $N_{ef}$ ), which was estimated from the sample of all *P. humanus* specimens to be 1.1 million female lice from the equation  $\Theta = 2N_{ef}\mu$ . The estimate of  $N_{ef}$  provides an expected coalescence time for the two divergent mitochondrial lineages of *P. humanus* of 1.10 million generations or ca. 0.11 MYA, which is an order of magnitude younger than the observed divergence time of 1.18 MYA. In a large randomly mating population consisting of 1.1 million female lice, one would expect to maintain two distinct haplotypes for only ca. 0.11 million years (MY). This suggests that we can reject panmixia if we assume that  $N_{ef}$  prior to the bottleneck was roughly similar to what we see today. If estimates of  $N_{ef}$  were drastically higher (ca. 60 million female lice) prior to the bottleneck, then expected time to coalescence could be much longer. The  $F_{st}$  value, a measure of genetic population differentiation, calculated for the WW and NW clades was 0.96, indicating substantial population structure, which also supports the rejection of panmixia.

The WW clade of *P. humanus* shows evidence of a recent population expansion (Fu and Li's  $D^* = -2.80$  [Fu and Li 1993];  $p < 0.02$ ). We estimated the date of this population expansion from the mismatch distribution. The estimate was calculated by comparing the average pairwise difference within the WW clade of *P. humanus* (4.21 mutations) to the pairwise difference between *P. humanus* and *P. schaeffi* (220 mutations), which diverged 5.6 MYA. The population expansion of the WW clade is estimated to be 0.11 MYA, similar to the estimated date of population expansion of modern humans out of Africa ca. 0.10 MYA (di Rienzo and Wilson 1991; Rogers and Harpending 1992; Harpending et al. 1993). In contrast, the NW clade of *P. humanus* does not exhibit the signature of a recent population expansion (Fu and Li's  $D^* = 0.17$ ), but instead shows a more stable population size.

### Contemporaneous Divergences in *Pediculus* and Archaic *Homo* spp.

The age of the MRCA of *P. humanus* dates to 1.18 MYA (for mtDNA), which is roughly midway between the estimated ages of *H. neanderthalensis* (0.60 MY) and *H. erectus* (1.8 MY). We used a maximum likelihood (ML) analysis to test whether our two divergent lineages of lice could have diverged in tandem with *H. sapiens* and *H. neanderthalensis* (Neandertals). *H. neanderthalensis* is the only other species of *Homo* for which DNA sequence data are available (Krings et al. 1999). The test evaluated whether relative branch lengths (scaled according to mutation rate) in the host tree, specifically for the branch between *H. sapiens* and *H. neanderthalensis*, are consistent with the parasite DNA sequence data (Huelsenbeck et al. 1997). In cospeciating assemblages, host and parasite branch lengths are highly correlated due to a shared evolutionary history (Page 1996). A likelihood ratio test (LRT) rejected ( $p < 0.0001$ ) the *H. sapiens*/*H. neanderthalensis* split as a node of cospeciation with the two clades of *P. humanus* because the branch length between *H. sapiens* and *H. neanderthalensis* is far too short to explain the louse DNA sequence data. In other words, the split between *H. sapiens* and *H. neanderthalensis* is too recent to have been contemporaneous with the divergence of the two lineages of lice. If one artificially lengthens the branch between *H. sapiens* and *H. neanderthalensis* to approximate the split between *H. sapiens* and *H. erectus* (anywhere from 1.2 to 1.8 MYA), the LRT fails to reject this hypothesis of cospeciation.

### Discussion

Morphological and molecular data agree that primates and their lice have been cospeciating for over 20 MY. Indeed, it is this cospeciation that permits us to use host fossil evidence to calibrate portions of the louse phylogenetic tree. This has resulted in the discovery of two extant lineages of human lice that diverged 1.18 MYA. This ancient divergence is surprising because humans, and presumably their lice, are thought to have passed through a population bottleneck ca. 0.05–0.10 MYA (Rogers and Harpending 1992). Such bottlenecks reduce genetic diversity by eliminating uncommon haplotypes, thereby making it less likely that multiple haplotypes survive bottleneck events. For example, mtDNA sequences from human populations coalesce to a single lineage very quickly ( $\leq 0.20$  MYA), presumably the result of the population bottleneck. The deep divergences found in *P. humanus* could

conceivably be the result of sequencing a nuclear copy of a mitochondrial gene. However, several lines of evidence strongly suggest otherwise. Because we amplified two different mitochondrial genes (COI and Cytb) that show the same divergent lineages and similar percent sequence divergences, copies of both mitochondrial genes would have had to enter the nucleus simultaneously, which is unlikely. In addition, we amplified each gene with a nested set of overlapping primers, and we never amplified more than one gene copy, even during bouts of cloning. Nucleotide base composition for our COI and Cytb data do not deviate from the mean values for all louse COI and Cytb sequences in GenBank (unpublished data), which would not be the case for a nuclear copy of a mitochondrial gene. Finally, the deep divergences seen in our mitochondrial genes are confirmed by preliminary analyses of nuclear data (EF1-alpha and 18S rRNA, unpublished data). Therefore, we are confident that the DNA sequences used in this study are mitochondrial in origin, and we must attempt to explain the occurrence of such ancient mitochondrial haplotypes in human lice.

### Gene Trees and Ancient Polymorphisms

Gene trees (e.g., mitochondrial lineages) can be considerably older than species trees, and therefore our louse mitochondrial lineages could predate the actual origin of the species *P. humanus* (i.e., its speciation time). It is useful to determine an expected time to coalescence from the estimated  $N_{ef}$  of 1.1 million female lice, even though this estimate seems high for a parasite of humans, who themselves have had very small effective population sizes (as few as 10,000 individuals) and recently went through a population bottleneck (Rogers and Harpending 1992). Although we do not necessarily expect human and louse effective population sizes to be directly correlated, it is difficult to imagine that humans could have maintained such a large effective population of lice during a bottleneck event. Regardless, the expected time to coalescence was estimated to be 0.10 MYA, an order of magnitude younger than the observed divergence time of 1.18 MYA. The deeper gene tree that our data provide also could have been produced either by balancing selection or by subdivision of the louse population into several distinct groups with very limited gene flow. A Fu and Li test does not detect balancing selection when both lineages of *P. humanus* are evaluated together ( $p = 0.11$ ); therefore, we must consider the alternative explanation of extensive population subdivision.

### Population Substructure and Host Geographic Isolation

Substantial isolation among populations of lice on modern *H. sapiens* could disrupt gene flow and allow the retention of very old lineages, making the age of *P. humanus* seem much older than it actually is. However, there is no evidence of such pervasive geographic isolation in the modern human hosts of these lice. Other species of lice have been shown to have substantial geographic substructure (i.e., isolation) even when hosts show no geographic isolation (Johnson et al. 2002). If populations of *P. humanus* are more highly subdivided than those of their hosts, then we might expect *P. humanus* to have retained ancient mitochondrial polymorphisms, even through host bottleneck events. One prediction of this hypothesis would be that both clades of *P. humanus* (the WW and NW clades) would show signs of the recent population

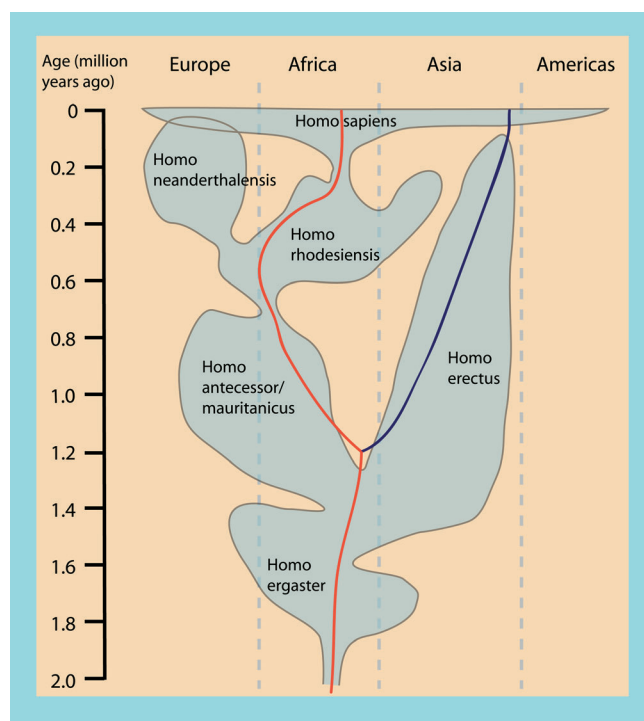
expansion of humans during the last ca. 0.10 MY. However, only the WW clade shows evidence of this event, which very closely matches the timing of human population expansion.

Because the WW clade is commonly found worldwide, and shows a population expansion concurrent with that of modern *H. sapiens*, we conclude that this lineage has a common evolutionary history with modern *H. sapiens*. In contrast, the NW clade appears to have diverged from the WW lineage 1.18 MYA, and has had a distinctly different evolutionary history. We are left unable to explain the retention of two ancient louse lineages, each with a different evolutionary history, within the confines of a single host, modern *H. sapiens*. Given the history of cospeciation between primate lice and their hosts, it is necessary to look beyond modern *H. sapiens* to determine whether the two divergent lineages of *P. humanus* are legacies of a more ancient divergence.

### Contemporaneous Divergences in *Pediculus* and Archaic *Homo* spp.

ML analyses rejected *H. neanderthalensis* as having diverged from *H. sapiens* contemporaneously with the two divergent lineages of lice. The mitochondrial MRCA of Neandertals and humans is 0.60 MYA (Kriings et al. 1997), which is only about half as old as the MRCA of the two ancient lineages of *P. humanus*, 1.18 MYA. The same ML test failed to reject the codivergence of these lice with *H. erectus* and *H. sapiens* when their divergence was set anywhere between 1.2 and 1.8 MYA. Therefore, the deep divergence within *P. humanus* is entirely consistent with a cospeciation event within the genus *Homo* ca. 1.2–1.8 MYA, but not 0.60 MYA. Unfortunately, no DNA sequence data exist for *H. erectus* or any other archaic species of *Homo* to enable a more direct test of cospeciation.

There is much debate regarding the past 2 MY of hominid evolution. However, one area of broad agreement is that, prior to 2 MYA, our ancestors were confined to Africa, then left the continent ca. 1.8 MYA. This first migration out of Africa resulted in archaic species of *Homo* that were widespread in distribution, and at times both contemporaneous with, and geographically isolated from, the lineage leading to modern *H. sapiens* (Figure 5). The 1.18 MY of isolation required to preserve the two ancient louse lineages must have occurred, in part, among these archaic species of *Homo*. It should be noted here that some interpretations of the Multiregional Evolution model do not necessarily consider modern *H. sapiens* to be a distinctly different species from archaic humans (e.g., *H. erectus* and *H. neanderthalensis*). We refer to them as “species” mostly for convenience of writing. Whereas the WW lineage has population genetic characteristics that are similar to those of modern *H. sapiens*, the geographically restricted NW lineage does not. It likely evolved on a now extinct species of *Homo* only to switch to modern *H. sapiens* very recently. For example, Figure 5 depicts one possible scenario where the NW lineage evolved on *H. erectus* and switched to modern *H. sapiens*. Interestingly, Hoberg et al. (2001) reported that two species of tapeworms of humans diverged ca. 0.78–1.71 MYA, and one of the two species, *Taenia asiatica*, is entirely restricted to Asia. This is consistent with the depiction in Figure 5, if one assumes that *T. asiatica* evolved on *H. erectus*. Although divergence dates are not available, it is intriguing that some Native American strains of HTLV have closer affinities to Asian primate strains



**Figure 5.** Temporal and Geographical Distribution of Hominid Populations Redrawn from Stringer (2003)

This figure depicts one view of human evolutionary history based on fossil data. Other interpretations differ primarily in the taxonomy and geographical distribution of hominid species. The temporal distribution of the two divergent lineages of *P. humanus* is superimposed on the hominid tree to show host evolutionary events that were contemporaneous with the origin of *P. humanus*. Whereas the NW lineage is depicted on *H. erectus* in this figure, several alternative hypotheses are consistent with our data when other evolutionary histories of hominids are considered (unpublished data). The WW clade is shown in red and the NW clade in blue (see text for descriptions of clades).

DOI: 10.1371/journal.pbio.0020340.g005

than to other human strains of HTLV, suggesting an independent Asian origin of this virus in humans. One must still explain how these parasites came to be on modern *H. sapiens*, but taken together, the parasitological evidence (especially the deep divergences in tapeworms and lice) suggests that they might have evolved on *H. erectus* and switched recently to *H. sapiens*. If true, this implies that *H. erectus* was contemporaneous with modern *H. sapiens* in eastern Asia, as suggested by Swisher et al. (1996), and it begs a discussion of recent human origins.

### Recent Human Origins

Explanations of modern human origins are dominated by two competing models, Recent African Replacement and Multiregional Evolution. The Recent African Replacement model assumes that anatomically modern *H. sapiens* arose as the result of a speciation event ca. 0.13 MYA, and then replaced without introgression (i.e., admixture) non-African archaic humans (e.g., *H. neanderthalensis* and *H. erectus*). In contrast, the sensu stricto Multiregional Evolution model assumes that modern *H. sapiens* evolved from early African descendants (up to 2 MYA). Characteristics of modernity were spread geographically through intercontinental gene flow,

but local regional characteristics were maintained through admixture between modern and archaic forms (Wolpoff et al. 1994). Most debates on modern human origins in the recent literature focus on one central question: “Was there admixture (i.e., introgression) between modern and archaic humans?” Both the Recent African Replacement and Multi-regional Evolution models have been proposed with numerous variations that include introgression, one of which was recently put forth by Eswaran (2002). Eswaran’s Diffusion Wave model proposes that a diffusion wave of modern *H. sapiens* left Africa (ca. 0.13 MYA) and replaced archaic humans through a process of introgression, natural selection, and gradual demic expansion.

The evolutionary history of *P. humanus* is somewhat consistent with all three models of modern human origins mentioned above; however, the number of ad hoc assumptions required to reconcile host and parasite evolutionary histories varies among the three views of human origins. Each model can account for the deep divergence between the two clades of *P. humanus* because each recognizes divergences between archaic species of *Homo* ca. 2 MYA. However, the model that best fits the louse data must account not only for the 1 MY of isolation between archaic and modern forms of lice, but also for a recent population expansion in just one louse lineage, the WW clade. The model must also explain how archaic louse DNA might have been incorporated into the lice of modern *H. sapiens*.

The sensu stricto model of Multiregional Evolution (Wolpoff et al. 1994) predicts continual gene flow between the geographically separated populations of humans following their early migration out of Africa (ca. 2 MYA), which is inconsistent with the louse data. This intercontinental gene flow among humans is required in the Multiregional Evolution model to maintain the continuity of the species *H. sapiens*. This scenario does not provide the louse populations with the degree of isolation necessary (ca. 1.18 MY) to maintain the two divergent louse lineages, unless we assume that gene flow between human populations was considerably greater than gene flow among their populations of lice. There is no reason to assume such a disparity in gene flow between hosts and parasites. The Multiregional Evolution model also predicts that we should detect the same genetic fingerprint of recent population expansion in both clades of *P. humanus*, which we do not.

The Recent African Replacement model provides the isolation necessary between archaic and modern forms, because it assumes that modern *H. sapiens* left Africa ca. 0.10 MYA, more than a million years after archaic species of *Homo* left Africa, and that the modern and archaic humans remained distinct (i.e., no introgression). Furthermore, it explicitly assumes a recent population expansion in modern *H. sapiens*, which would account for the population expansion seen in our WW clade. However, in the strict sense, this model also predicts that modern *H. sapiens* replaced archaic forms of humans without introgression (i.e., hybridization), which leaves no obvious mechanism for archaic louse DNA to reach the lice of modern *H. sapiens*. This lack of host introgression implies, but does not require, a lack of direct physical contact between modern and archaic humans. It is conceivable that direct contact between modern and archaic humans was sufficient to allow the lice to switch hosts without making the assertion that the hosts were interbreeding. Therefore, the

Recent African Replacement model is fairly consistent with the louse data, so long as one assumes some level of direct contact (e.g., fighting, sharing/stealing of clothing, etc.) between modern and archaic humans.

Eswaran’s Diffusion Wave model (2002) is similar to the Multiregional Evolution model in that it permits some level of introgression between modern and archaic humans. However, it is also similar to the Recent African Replacement model in that it assumes the same recent population expansion of modern humans out of Africa (ca. 0.10 MYA), thus providing both the isolation and population expansion necessary to accommodate our louse data. The additional assumption of introgression between modern and archaic forms of humans, which is proposed to have occurred only in the last 0.10 MY, provides a ready vehicle that would have transported archaic louse DNA into the modern louse population. Eswaran’s model applied to lice suggests that at the beginning of the diffusion wave of modern *H. sapiens* leaving Africa (ca. 0.13 MYA), modern and archaic humans had distinct types of lice owing to  $\geq 1$  MY of isolation. As modern humans began to replace archaic forms, direct contact between hosts during introgression allowed archaic lice to switch to modern *H. sapiens* hosts.

As previously stated, the recent literature addressing human origins boils down to models that do not permit introgression (strict-sense replacement models) and those of many types that do (admixture models, including variants allied with both the Multiregional Evolution and Recent African Replacement models). All things being equal, our parasite data are most consistent with a limited amount of admixture between modern and archaic humans, because this process presents the opportunity for host switching. However, introgression between modern and archaic humans over a protracted period of time would erode the isolation required to maintain the two louse lineages that we have observed. For example, some variants of the Multiregional Evolution model reject a single origin of modernity in Africa ca. 0.13 MYA in favor of a piecemeal acquisition of modern traits over a long period of time. This long-term admixture is precisely what would disrupt the isolation required to maintain the two louse lineages. Eswaran’s Diffusion Wave model, on the other hand, confines admixture to the last ca. 0.10 MY.

Our data cannot directly address whether host introgression occurred, because nonsexual, direct contact between hosts is sufficient for parasite transmission. We are confident that “direct contact” would be required for a host switch because these obligate parasitic lice cannot move between individuals without direct physical contact (Buxton 1946; Durden 2001; Canyon et al. 2002; Burgess 2004) and furthermore, they die within 24 h of being removed from their host. However, an examination of *Pthirus pubis*, the human pubic louse, might shed light on the subject of human admixture because unlike head and body lice, pubic lice are primarily transmitted during intercourse.

If our scenario involving lice switching from *H. erectus* to *H. sapiens* were true, then the host switch would have brought together two long-separated taxa of lice. It is impossible to know whether this long separation affected the reproductive compatibility of the two louse taxa once reunited. Discriminant function analysis shows no morphological differences between members of the two divergent molecular haplotypes



of head lice (see Figure 3). There are other well-defined species of lice (e.g., see Johnson et al. 2002) whose populations show even greater sequence divergence (19% uncorrected sequence divergence) and yet have no discernible morphological differences between populations. It is likely that the two long-separated types of lice have experienced some level of introgression since their secondary contact on modern *H. sapiens*. The recency of this introgression of archaic louse DNA into modern lice also accounts for the younger coalescence time for the NW clade (0.15 MYA) compared to the WW clade (0.54 MYA). Presumably, the archaic form (i.e., morphotype) of louse either was extirpated along with its host or was assimilated into modern *P. humanus*. Regardless of the mechanism, ancient louse lineages can be found among the lice of modern *H. sapiens*.

A recent review by Ashford (2000) reported five parasites that occur on humans as closely related pairs of taxa (lice, tapeworms, follicle mites, a protozoan, and bedbugs). The fact that there are *five* such pairs caused Ashford to ask, were humans once two distinct populations that rejoined after a long separation? The ancient divergences seen in mitochondrial data from *P. humanus* are clearly consistent with some level of long-term host isolation, and preliminary evidence from nuclear markers (EF1-alpha and 18S rRNA) reveals similarly ancient divergences (unpublished data). Furthermore, the two tapeworm species from humans showed amazingly concordant divergences (Hoberg et al. 2001) and distributional patterns. Our data suggest that the isolation Ashford refers to may be between species of *Homo* rather than within modern *H. sapiens* itself. We conclude that the parasites may be very useful in the study of human evolutionary history, because they represent an independent marker of human evolution that has yet to be studied in detail.

## Materials and Methods

**Specimen collection and preparation.** We collected human head and body lice (*P. humanus*) from many localities, ranging from remote areas such as the Papua New Guinea highlands to metropolitan areas like Boston (Table 2). We also obtained *P. schaeffi* (from chimpanzees), *Pthirus pubis* (from humans), *Pedicinus hamadryas* (from baboons), and *Fahrenholzia pinnata* (from a rodent) to use as outgroup taxa in phylogenetic analyses. All lice were preserved in 95% EtOH and stored at  $-80^{\circ}\text{C}$ . DNA was extracted from lice by separating the thorax and abdomen and placing both in digestion buffer (Qiagen DNeasy tissue kit; Qiagen, Valencia, California, United States). Digestion proceeded for 48 h at  $55^{\circ}\text{C}$ , then followed the manufacturer's protocol. After digestion, each louse was reassembled on a microscope slide as a voucher specimen corresponding to each DNA sequence. Voucher specimens were deposited in the Price Institute of Phthirapteran Research (PIPeR) collection at the University of Utah.

**Phylogenetic analyses: morphological data.** Considerable morphological variation exists among different species of primate lice. We examined 155 unordered morphological characters for 113 specimens of *P. humanus* (from humans), *P. schaeffi* (from chimpanzees), *Pthirus pubis* (from humans), *Pthirus gorillae* (from gorillas), *Pedicinus hamadryas* (from baboons), and *F. pinnata* (from a rodent). Morphological data were scored in the software package MacClade v. 4.05 (W. P. Maddison and D. R. Maddison; Sinauer, Sunderland, Massachusetts, United States), and heuristic searches consisting of random stepwise addition (1,000 replicates) and tree bisection/reconnection branch swapping were performed in PAUP\* v. 4.0b10 (D. L. Swofford; Sinauer). Branch support was estimated with bootstrapping (tree bisection/reconnection swapping, 1,000 replicates). The complete data matrix is available from TreeBASE (<http://www.treebase.org/>) as study accession number SN1969.

**Primate-lice cospeciation.** The morphological data set (six species, 155 characters) was compared to the host phylogeny

((((human, chimp), gorilla), baboon), rodent) using reconciliation analysis in Treemap v. 2.0 with default parameters. Treemap determines whether the two phylogenies are more congruent than expected by chance based on randomizations of both the host and parasite phylogeny. Significant congruence between host and parasite phylogenies is interpreted as being the result of a shared evolutionary history (i.e., repeated bouts of cospeciation).

**Discriminant function analysis.** We examined the morphology of *P. humanus* lice in detail to test for morphological correlates of the differences detected at the molecular level. Busvine (1978) examined a large series of head and body louse specimens and found no discrete morphological differences between the two forms. However, he noted that several morphological characters related to size and shape might be useful in this regard. To test this hypothesis we measured head width, thoracic width, total body length, and second-leg tibia length from a series of 50 slide-mounted adult museum specimens collected by earlier workers prior to our study. The microhabitat (head or body, i.e., clothing) from which these museum specimens were collected was well documented. Canonical discriminant analysis was used to build a predictive model to attempt to distinguish between the head and body forms of *P. humanus*. The predictor variables were used to build a set of discriminant functions that maximized variation among groups while minimizing within-group variation. The first two canonical discriminant functions explained 100% of the variation within the data.

These discriminant functions, which were built using existing museum specimens, were then applied to our newly collected specimens in a blind test to determine whether the specimens could be identified as head or body lice from morphology alone. We were able to classify our samples as head or body lice with a probability of  $\geq 0.95$ . Indeed, the assignment of adult specimens proved to be 100% accurate when checked against microhabitat data for the new specimens.

**Phylogenetic analyses: molecular data.** Fresh specimens suitable for the collection of molecular data were obtained for five of the six species of lice. We sequenced 1,525 combined base pairs (bp) of the mitochondrial (mtDNA) genes COI (854 bp) and Cytb (671 bp) from 69 individuals of *P. humanus*, *P. schaeffi*, *Pthirus pubis*, *Pedicinus hamadryas*, and *F. pinnata*. PCR primers were as follows: (5'-3') COI, C1-J-1718 GGAGGTTTGTCTAATTGATTAG and H7005 CCGGATC-CACNACRTARTANGTRTCRTG; Cytb, L11122 GAAATTTGGGT-CWTRCTNGG and H11823 GGCATATGCCAATARGAARTATCA. PCR parameters included  $94^{\circ}\text{C}$  for 30 s,  $48^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 1.5 min (five cycles), then 30 cycles of  $94^{\circ}\text{C}$  for 30 s,  $52^{\circ}\text{C}$  for 30 s, and  $72^{\circ}\text{C}$  for 1.5 min. Amplified fragments were sequenced in both directions, assembled using Sequencher v. 4.1 (GeneCodes, Ann Arbor, Michigan, United States), and deposited in the NCBI database (see Supporting Information). To ensure that we were not amplifying nuclear copies of mitochondrial genes, we performed additional PCR amplifications using nested sets of overlapping primers.

The computer program modelTest (Posada and Crandall 1998) was used as a guide to determine a best-fit (Cunningham et al. 1998) ML model for the molecular data. This model (GTR+I+G) was incorporated into ML branch and bound and heuristic searches in PAUP\* with 100 bootstrap replicates. An LRT was used to compare ML estimates from a clock-enforced and an unconstrained analysis. Our data did not depart significantly from the assumption of a molecular clock.

**Dating nodes in the louse phylogeny.** We used the significant cospeciation shown between the primate and louse phylogenies (see Results) as a basis for dating nodes in the louse phylogeny. We used a calibration point of  $22.5 \pm 2.5$  MYA for the split between *Pedicinus* and *Pediculus* + *Pthirus*. The date is based on fossil evidence (Benefit 1993; Leakey et al. 1995) of the split between cercopithecoid primates that host only lice in the genus *Pedicinus* and hominoid primates that host only lice in the genera *Pthirus* and *Pediculus* (Durden and Musser 1994). All reconciliations of the host and parasite trees in our cospeciation analysis determined that this particular host/parasite node represents a cospeciation event (see Results). Using the computer software r8s (M. J. Sanderson, software distributed by the author), we constrained the divergence of *Pedicinus* + *Pthirus*/*Pediculus* to  $22.5 \pm 2.5$  MYA and allowed all other nodes in the louse tree to be determined from our DNA sequence data. Error estimates on divergence dates were calculated by generating 100 bootstrapped data matrices in Phylip (J. Felsenstein, software distributed by the author). Each of these bootstrapped datasets was calibrated with the same  $22.5 \pm 2.5$  MYA divergence.

**Cospeciation within *Homo*.** An ML-based analysis was used to determine whether deep divergences in the louse tree were contemporaneous with divergences of now extinct species of *Homo*.

**Table 2.** Specimens of *P. humanus* and Outgroup Taxa Examined in This Study, Their Collection Locality, and Number of Specimens Examined

Scientific Name	Location	Number of Specimens
<i>P. humanus</i> (head)	Honduras <sup>a</sup>	12
<i>P. humanus</i> (head)	Papua New Guinea <sup>a</sup>	4
<i>P. humanus</i> (head)	Philippines <sup>a</sup>	16
<i>P. humanus</i> (head)	United Kingdom <sup>a</sup>	1
<i>P. humanus</i> (head)	United States <sup>a</sup>	4
<i>P. humanus</i> (head)	Australia <sup>b</sup>	2
<i>P. humanus</i> (head)	China <sup>b</sup>	12
<i>P. humanus</i> (head)	Hungary <sup>b</sup>	1
<i>P. humanus</i> (head)	Israel <sup>b</sup>	3
<i>P. humanus</i> (head)	Kenya <sup>b</sup>	3
<i>P. humanus</i> (head)	New Zealand <sup>b</sup>	3
<i>P. humanus</i> (head)	Papua New Guinea <sup>b</sup>	4
<i>P. humanus</i> (head)	Burundi <sup>c</sup>	1
<i>P. humanus</i> (head)	Russia <sup>c</sup>	1
<i>P. humanus</i> (head)	France <sup>c</sup>	1
<i>P. humanus</i> (body)	Boston, Massachusetts, United States <sup>a</sup>	2
<i>P. humanus</i> (body)	United States <sup>a</sup>	7
<i>P. humanus</i> (body)	Socotra Island, Yemen <sup>a</sup>	2
<i>P. humanus</i> (body)	China <sup>b</sup>	17
<i>P. humanus</i> (body)	Japan <sup>b</sup>	5
<i>P. humanus</i> (body)	Kenya <sup>b</sup>	3
<i>P. humanus</i> (body)	United States <sup>b</sup>	3
<i>P. humanus</i> (body)	Tunisia <sup>c</sup>	1
<i>P. humanus</i> (body)	Burundi <sup>c</sup>	1
<i>P. humanus</i> (body)	Zimbabwe <sup>c</sup>	1
<i>P. humanus</i> (body)	Peru <sup>c</sup>	1
<i>P. humanus</i> (body)	France <sup>c</sup>	2
<i>P. humanus</i> (body)	Russia <sup>c</sup>	1
<i>P. schaeffi</i>	Uganda <sup>a</sup>	4
<i>Pthirus pubis</i>	United States <sup>a</sup>	6
<i>Pthirus pubis</i>	United Kingdom <sup>a</sup>	1
<i>Pedicinus hamadryas</i>	United States (captive) <sup>a</sup>	5

<sup>a</sup> From this study.

<sup>b</sup> From Leo et al. (2002).

<sup>c</sup> Downloaded from GenBank (see Supporting Information for accession numbers).

DOI: 10.1371/journal.pbio.0020340.t002

We calculated the branch length between *H. sapiens* and *H. neanderthalensis* (the only species of *Homo* for which DNA sequence data are available) based on the human mitochondrial hypervariable region II of the D loop (see Supporting Information). This value was scaled according to the average distance between these two taxa and their sister taxon (chimpanzee), providing a relative branch length within the primate tree (e.g., the branch between human and Neandertal is one-fifth the length of the branch that unites them with the chimpanzee). This relative branch length was incorporated into a louse constraint tree, in effect forcing the two clades of *P. humanus* to conform to a prescribed relative branch length. The resulting likelihood score was compared with the unconstrained tree score using an LRT (d.f. = taxa - 2). If the constrained and unconstrained tree scores are not significantly different, then the host tree topology and branch lengths describe the parasite DNA sequence data as well as the parasite tree itself (Huelsenbeck and Crandall 1997). However, if a significant difference is detected, then the host tree does not fit the parasite data well enough to be explained by cospeciation, and the hypothesis of cospeciation is rejected.

**Population genetic analyses.** Population genetic analyses were performed on a pruned dataset, which contained only specimens of *P.*

*humanus* (i.e., no outgroup taxa). The computer software package DnaSP (Rozas and Rozas 1999) was used to generate mismatch distributions, to calculate Fu and Li's D\* statistic for the *P. humanus* clades, and to calculate additional population parameters (e.g.,  $F_{st}$  and  $\Theta$ ). Ten additional haplotypes from Leo et al. (2002) were used in these population-level analyses (see Supporting Information). To estimate an expected time to coalescence for *P. humanus*, we used estimates of theta ( $\Theta$ , from the software package DnaSP [Rozas and Rozas 1999]) and mutation rate ( $\mu$ ) for *P. humanus* (see below) to calculate louse  $N_{ef}$  from the equation  $\Theta = 2N_{ef}\mu$ . The mutation rate ( $\mu = 9.0 \times 10^{-9}$  substitutions per site per generation) was calculated by determining the expected number of substitutions per site between *P. humanus* and *P. schaeffi* under the Tamura-Nei +  $\Gamma$  model of nucleotide substitution. This mutation rate for *P. humanus* is roughly five to six times faster than that of human mtDNA, excluding the D-loop (Ingman et al. 2000), when both mutation rates are scaled to absolute time (i.e., number of substitutions per site per year).  $N_{ef}$  was then used to determine the expected time to coalescence (in generations) given the formula  $2N_{ef}(1 - 1/n)$ , where  $n$  is the number of haplotypes detected in the population. One can also ask the similar question, what is the probability that two lineages, which are expected to differ by  $k_e$  substitutions, actually differ by  $k^*$  or more

substitutions? The expression used,  $(k_e / (1 + k_e))^k$ , is derived from the geometric distribution (see, for example, Golding and Strobeck 1982) and for our data suggests that the large number of substitutions found between the WW and NW clades is far greater than that which is expected ( $p < 0.0006$ ).

## Supporting Information

### Accession Numbers

GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/>) accession numbers for items discussed in the text are as follows: the EF1-alpha sequences from Kittler et al. (2003), AY316794–AY316834; the 18S rRNA sequences for *P. humanus* from Yong et al. (2003), AY236410–AY236418, AF139478–AF139482, AF139484, AF139486, and AF139488; the 18S rRNA sequence from *P. schaeffi*, AY695939; the Cytb sequence from *P. schaeffi*, AY316793; the human mitochondrial hypervariable region II of the D loop, M76311, AY195756, AY217615, AF282972, AF142095, X97709, X98472, X93336, X93337, X93347, and X93348; the ten haplotypes from Leo et al. (2002) used in population-level analyses, AF320286; the divergent louse *Dennyus hirundinus* shown in Figure 4, AF545694 and U96434.

The amplified PCR fragments discussed in Materials and Methods have been deposited in the NCBI database under accession numbers AY695939–AY696069.

### References

- Ashford RW (2000) Parasites as indicators of human biology and evolution. *J Med Microbiol* 49: 771–772.
- Benefit BR (1993) The permanent dentition and phylogenetic position of *Victoriapithecus* from Maboko Island, Kenya. *J Hum Evol* 25: 83–172.
- Burgess IF (2004) Human lice and their control. *Annu Rev Entomol* 49: 457–481.
- Busvine JR (1978) Evidence from double infestations for the specific status of human head lice and body lice (Anoplura). *Syst Entomol* 3: 1–8.
- Buxton PA (1946) The louse. Baltimore: Williams and Wilkins Co, 164 p.
- Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* 325: 31–36.
- Canyon DV, Speare R, Muller R (2002) Spatial and kinetic factors for the transfer of head lice (*Pediculus capitis*) between hairs. *J Invest Dermatol* 119: 629–631.
- Caramelli D, Laluzze-Fox C, Vernesi C, Lari M, Casoli A, et al. (2003) Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. *Proc Natl Acad Sci U S A* 100: 6593–6597.
- Chan SY, Bernard HU, Ong CK, Chan SP, Hofmann B, et al. (1992) Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and variants: A showcase for the molecular evolution of DNA viruses. *J Virol* 66: 5714–5725.
- Cunningham CW, Zhu H, Hillis DM (1998) Best-fit maximum likelihood models for phylogenetic inference: Empirical tests with known phylogenies. *Evolution* 52: 978–987.
- di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc Natl Acad Sci U S A* 88: 1597–1601.
- Durden LA (2001) Lice (Phthiraptera). In: Samuel WM, Pybus MJ, Kocan AA, editors. Parasitic diseases of wild mammals. Ames: Iowa State University Press. pp. 3–17.
- Durden LA, Musser GG (1994) The sucking lice (Insecta: Anoplura) of the world: A taxonomic checklist with records of mammalian hosts and geographical distributions. *Bull Am Mus Nat Hist* 218: 1–90.
- Escalante AA, Freeland DE, Collins WE, Lal AA (1998) The evolution of primate malaria parasites based on the gene encoding cytochrome *b* from the linear mitochondrial genome. *Proc Natl Acad Sci U S A* 95: 8124–8129.
- Eswaran V (2002) A diffusion wave out of Africa. *Curr Anthropol* 43: 749–774.
- Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics* 133: 693–709.
- Golding GB, Strobeck C (1982) The distribution of nucleotide site differences between two finite sequences. *Theor Popul Biol* 22: 96–107.
- Hafner MS, Sudman PD, Villablanca FX, Spradling TA, Demastes JW, et al. (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* 265: 1087–1090.
- Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Curr Anthropol* 34: 483–496.
- Ho L, Chan S, Burk R, Das B, Fujinaga K, et al. (1993) The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 67: 6413–6423.
- Hoberg EP, Alkire NL, de Queiroz A, Jones A (2001) Out of Africa: Origins of the *Taenia* tapeworms in humans. *Proc R Soc Lond B Biol Sci* 268: 781–787.
- Huelsenbeck JP, Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu Rev Ecol Systematics* 28: 437–466.

## Acknowledgments

We thank F. Adler, C. Bell, S. Bush, J. Coyne, M. Gitzendanner, M. Hafner, H. Harpending, K. Johnson, W. Potts, J. Seger, and three reviewers for helpful comments. For providing specimens we thank D. Colwell (Agriculture and Agri-Food Canada), K. Rice (Southwest Foundation for Biomedical Research), R. Todd (Insect Control and Research, Inc.), D. Altschuler (National Pediculosis Association), J. O'Connell (Boston Health Care for the Homeless Program), D. Dean (Salt Lake City Public Health Center), Debby Cox (Chimpanzee Sanctuary, Wildlife Conservation Trust, and The Jane Goodall Institute), and Nigel Hill (London School of Hygiene and Tropical Medicine). We also thank C. Lyal and P. Brown at the Natural History Museum, London, and D. Arnold, K.C. Emerson Entomology Museum, Stillwater, Oklahoma, for access to museum collections of *Phthiraptera*. We thank J. Light for sequences of the rodent outgroup taxon, *Fahrenholzia*. This research was supported by a University of Utah Funding Incentive Seed Grant and National Science Foundation (NSF) grants to DHC (DEB-0107947 and DEB-0118794), an NSF Postdoctoral Fellowship to DLR (DBI-0102112), and a Wellcome Trust Junior Biodiversity Fellowship to VSS.

**Conflicts of interests.** The authors have declared that no conflicts of interest exist.

**Author contributions.** DLR, VSS, and DHC designed the study. DLR, VSS, and SLH collected and analyzed the data. DLR, VSS, ARR, and DHC wrote the paper. ■

- Huelsenbeck JP, Ronquist F, Yang Z (1997) Statistical tests of host-parasite cospeciation. *Evolution* 51: 410–419.
- Ingman M, Kaessmann H, Paabo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408: 708–713.
- Johnson KP, Williams BL, Drown DM, Adams RJ, Clayton DH (2002) The population genetics of host specificity: Genetic differentiation in dove lice (Insecta: Phthiraptera). *Mol Ecol* 11: 25–38.
- Kittler R, Kayser M, Stoneking M (2003) Molecular evolution of *Pediculus humanus* and the origin of clothing. *Curr Biol* 13: 1414–1417.
- Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, et al. (1997) Neandertal DNA sequences and the origin of modern humans. *Cell* 90: 19–30.
- Krings M, Geisert H, Schmitz RW, Krainitzki H, Paabo S (1999) DNA sequence of the mitochondrial hypervariable region II from the Neandertal type specimen. *Proc Natl Acad Sci U S A* 96: 5581–5585.
- Krings M, Capelli C, Tschentscher F, Geisert H, Meyer S, et al. (2000) A view of Neandertal genetic diversity. *Nat Genet* 26: 144–146.
- Leakey MG, Ungar PS, Walker A (1995) A new genus of large primate from the late Oligocene of Lothidok, Turkana District, Kenya. *J Hum Evol* 28: 519–531.
- Leal ES, Zanotto PMA (2000) Viral diseases and human evolution. *Mem Inst Oswaldo Cruz* 95: 193–200.
- Leo NP, Campbell NJH, Yang X, Mumcuoglu K, Barker SC (2002) Evidence from mitochondrial DNA that head and body lice of humans (Phthiraptera: Pediculidae) are conspecific. *J Med Entomol* 39: 662–666.
- Levene H, Dobzhansky T (1959) Possible genetic difference between the head louse and the body louse. *Am Nat* 87: 347–353.
- Ong C, Chan S, Campo M, Fujinaga K, Mavromara-Nazos P, et al. (1993) Evolution of human papillomavirus type 18: An ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol* 67: 6424–6431.
- Paabo S, Wilson AC (1991) Miocene DNA sequences—a dream come true? *Curr Biol* 1: 45–46.
- Page RD, Lee PL, Becher SA, Griffiths R, Clayton DH (1998) A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. *Mol Phylogenet Evol* 9: 276–293.
- Page RDM (1996) Temporal congruence revisited: Comparison of mitochondrial DNA sequence divergence in cospeciating pocket gophers and their chewing lice. *Syst Biol* 45: 151–167.
- Page RDM (2003) Tangled trees: Phylogeny, cospeciation, and coevolution. Chicago: University of Chicago Press. 378 p.
- Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
- Reed DL, Hafner MS (1997) Host specificity of chewing lice on pocket gophers: A potential mechanism for cospeciation. *J Mammal* 78: 655–660.
- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9: 552–569.
- Rozas J, Rozas R (1999) DnaSP version 3: An integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174–175.
- Serre D, Langaney A, Chech M, Teschler-Nicola M, Paunovic M, et al. (2004) No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biol* 2: e57 DOI: 10.1371/journal.pbio.0020057.
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16: 1114–1116.

- Stauffer RL, Walker A, Ryder OA, Lyons-Weiler M, Hedges SB (2001) Human and ape molecular clocks and constraints on paleontological hypotheses. *J Hered* 92: 469–474.
- Stringer CB (2003) Out of Ethiopia. *Nature* 423: 692–695.
- Stringer CB, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. *Science* 239: 1263–1268.
- Swisher CC III, Rink WJ, Anton SC, Schwarcz HP, Curtis GH, et al. (1996) Latest *Homo erectus* of Java: Potential contemporaneity with *Homo sapiens* in Southeast Asia. *Science* 274: 1870–1874.
- Templeton AR (2002) Out of Africa again and again. *Nature* 416: 45–51.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253: 1503–1507.
- White TD, Asfaw B, DeGusta D, Gilbert H, Richards GD, et al. (2003) Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423: 742–747.
- Wolpoff MH, Thorne AG, Smith FH, Frayer DW, Pope GG (1994) Multiregional evolution: A world-wide source for modern human populations. In: Nitecki M, Nitecki D, editors. *Origins of anatomically modern humans*. New York: Plenum Press. pp. 174–199.
- Wolpoff MH, Hawkes J, Caspari R (2000) Multiregional, not multiple origins. *Am J Phys Anthropol* 112: 129–136.
- Yong Z, Fournier PE, Rydkina E, Raoult D (2003) The geographical segregation of human lice preceded that of *Pediculus humanus capitis* and *Pediculus humanus humanus*. *C R Biol* 326: 565–574.