Catalog of supplementary information.

Table S1. Characteristics of some major human infectious diseases.

Note S1. Choice of diseases.

Note S2. "Temperate" vs. "tropical"

Note S3. Robustness tests.

Note S4. Paucity of major diseases before Stage 4.

Note S5. Animal origins of temperate diseases.

Note S6. Animal origins of tropical diseases.

Note S7. Tropical domestic animals.

Note S8. Our speculative guesses about the next big killers.

Note S9. Details on each of our 25 major diseases.

Supplementary references.

Table S1. Characteristics of some major human infectious diseases, divided into two sets: ones predominantly tropical or of tropical origin, and ones regularly occurring at temperate as well as tropical latitudes. Column 1: * = a crowd disease of large human populations (see text). Column 2: the pathogen causing the disease. Column 3: whether transmitted to humans by an animal vector, or else directly from another human and if so by what means (breathed-out air = aerosol, sexual, skin contact, etc.). Column 4: approximate duration, from infection until death or recovery. Column 5: case-fatality ratio in traditional human populations without modern medical treatment or nutrition (from ref. S1 and other sources). Column 6: whether patients who recover thereby develop long-lasting or lifetime immunity against re-infection. Columns 7 and 8: whether the pathogen has a reservoir apart from infected humans, either in the environment or in animals (from ref. S2 and other sources). The term "reservoir" means a non-human source of pathogens immediately capable of infecting humans without evolutionary change. Reservoirs may enable a pathogen to persist even in the absence of a susceptible human population large enough by itself to sustain the pathogen. Column 9: our tentative conclusion about the most likely source from which the pathogen reached humans. Column 10: to which of the five evolutionary stages explained in Box 1 we assign the disease. See Notes S1, S2, and S9 for further explanation, and for details and references on each disease.

Table S1. "TEMPERATE" DISEASES

1. Disease (* = crowd disease)	2. Agent	3. Route of transmission	4. Duration	5. Untreated case-fatality ratio	6. Long-lasting immunity?	7.Environmental reservoir?	8. Animal reservoir?	9. Likely origins	10. Stage
Diphtheria *	*Corynebacterium diphtheriae*	Human: aerosol, contact	≤ 1 week	35-90%	Usually	Clothes, dust (several weeks)	No	Domestic herbivores	5
Hepatitis B	Hepadnavirus: Hepatitis B virus	Human: perinatal, horizontal, sexual, parenteral	Months	<5-10%	Yes, if cleared	On surfaces 7+ days	No	Apes	5
Influenza A *	Myxovirus: Influenzavirus A	Human: aerosol	1 week	varies by strain	Strain-dependent	No	Wild birds	Ducks & pigs, ultimately wild birds	4c
Measles *	Paramyxovirus: Morbillivirus	Human: aerosol	Weeks	10-25%	Yes	No	No	Cattle	5
Mumps *	Paramyxovirus: Rubulavirus	Human: saliva, aerosol	Weeks	1-2%	Yes	No	No	Mammal, possibly pigs	5
Pertussis *	*Bordetella pertussis*	Human: aerosol	Weeks	High for children	Usually, 10-20 years	No	No	Mammal (wide host range)	5
Plague *	*Yersinia pestis*	Vector: fleas Human: aerosol	1-2 weeks	25-90%	Long lasting	Soil	Rodents	Rodents	3
Rotavirus A	Reoviridae: Rotavirus	Human: fecal-oral	1 week	1 - 10%	Partial	Possibly water	No	Domestic herbivores, other mammals	5?
Rubella *	Togavirus: Rubivirus	Human: aerosol, transplacental	Weeks	Adults – low, fetuses - high	Usually	No	No	Unknown	5
Smallpox *	Poxvirus: variola virus	Human: aerosol, skin	Weeks	20-50%	Yes	Textiles: days to weeks	No	Camels?	5
Syphilis	*Treponema pallidum*	Human: sexual, transplacental	Formerly months, now years	6-50% congenital; ~50% adult?	No	No	No	Unknown	5
Tetanus	*Clostridium tetani*	Wounds	2 weeks	50%	No	Soil	(Intestines of animals)	Unknown	2
Tuberculosis	*Mycobacterium tuberculosis*	Human: aerosol	Years	50%	No	No	No?	Ruminants?	5
Typhoid *	*Salmonella enterica var. typhi*	Human: fecal-oral	Weeks	10-20%	Usually	No	No	Unknown	5
Typhus *	*Rickettsia prowazekii*	Vector: louse	1-2 weeks	10-40%	Long lasting	No	No, but occasionally flying squirrels	Rodents?	4c

Table S1, cont. "TROPICAL" DISEASES

1. Disease	2. Agent	3. Route of transmission	4. Duration	5. Untreated case-fatality ratio	6. Long-lasting immunity?	7. Environmental reservoir?	8. Animal reservoir?	9. Likely origins	10. Stage
AIDS	Retrovirus: HIV-1	Human: sexual	Years	Nearly 100%	Rarely	No	No	Chimpanzee	5
Chagas' disease	Trypanosome: *Trypanosoma cruzi*	Vector: kissing bugs	Months to decades	≥ 30%	No	No	Many mammals	Many wild and domestic mammals	3
Cholera	*Vibrio cholerae*	Human: fecal-oral	1 week	≤ 60%	Partial?	Aquatic environments	Aquatic copepods, algae, crustacea	Aquatic organisms?	4c
Dengue fever (hemorrhagic)	Flavivirus: dengue virus serotypes DEN-1,2,3,4	Vector: mosquitoes	1 – 2 weeks	Variable, up to 15% in children	Yes, strain-dependent	No	Old World primates	Old World primates	4c
East African sleeping sickness	Trypanosome: *Trypanosoma brucei rhodesiense*	Vector: tsetse flies	Weeks to 9 months	Nearly 100%	No	No	Wild and domestic ruminants	Wild and domestic ruminants	2
Falciparum malaria	Plasmodium: *Plasmodium falciparum*	Vector: mosquitoes	9 days to years	5-25%	Partial, strain-dependent	No	No	Ultimately birds, but perhaps ancient	5
Visceral leishmaniasis	*Leishmania donovani*	Vector: sand flies	Years	80-95%	No	No	Dogs, rodents	Dogs, rodents	2, but 5 in India
Vivax malaria	Plasmodium: *Plasmodium vivax*	Vector: mosquitoes	2 weeks to years	Relatively low	Limited	No	Occasionally New World monkeys	Asian macaques	5, but 4c in Americas
West African sleeping sickness	Trypanosome: *Trypanosoma brucei gambiense*	Vector: tsetse flies	Months to 6 years	Nearly 100%	No	No	Minor: wild and domestic ruminants	Wild and domestic ruminants	4c
Yellow fever	Flavivirus: Yellow fever virus	Vector: mosquitoes	1-2 weeks	variable, can be >50%	Yes	No	Nonhuman primates	African primates	4a

Note S1. Choice of diseases.

In the text and in Table S1 we attempted to select for analysis 25 infectious diseases of the greatest evolutionary and historical significance. While everyone would agree that such a list should include smallpox, measles, tuberculosis, and falciparum malaria, some other choices pose a gray area, and other scientists might arrive at a list somewhat different from ours. The considerations behind our choices were as follows:

One consideration is current mortality, which immediately leads to the inclusion of falciparum malaria and tuberculosis. In addition, not just mortality, but also non-fatal morbidity, can be historically and evolutionarily significant. For instance, syphilis has caused high frequency of sterility in some populations, such as Andaman Islanders. While rubella causes only modest postnatal mortality, it causes up to 90% incidence of fetal malformation or death in fetuses of mothers infected during the first trimester of pregnancy. A widely used measure of a disease's mortality and morbidity is its so-called DALY score (disability-adjusted life years): "a health gap measure that extends the concept of potential years of life lost due to premature death to include equivalent years of healthy life lost by virtue of individuals being in states of poor health or disability"[S3]. "One DALY can be thought of as one lost year of healthy life, and the burden of disease as a measure of the gap between current health status and an ideal situation where everyone lives into old age free from disease and disability"[S4].

Lopez et al.[S4] have tabulated DALY values for the whole world for all causes, including not just infectious and non-infectious diseases but also other causes, such as traffic accidents, fires, drownings, and war. From their table, we have taken the 17 specific, infectious, microbial,

non-commensal disease entities with the highest DALY scores as of 2001, ranging from AIDS (#1) and malaria (#2) down to Chagas' disease (#16) and dengue fever (#17). More specifically, we took the eight temperate disease entries and the seven tropical disease entries with the highest DALY scores in the table of Lopez et al.[S4], but two of the tropical entries actually refer to pairs of diseases as explained below. Five sets of specific considerations arose in extracting the table of Lopez et al.[S4]:

1. We included only microbes. We excluded helminths and other macroparasites, such as filiariasis, schistosomiasis, and hookworm. We also excluded commensals, such as Staphylococcus aureus, Streptococcus pneumoniae and other streptococci, Neisseria meningitidis, and Haemophilus influenzae B -- i.e., microbes that are sometimes part of our normal flora and do not harm us but that can also cause serious illness under certain conditions, e.g., as secondary infections in patients already weakened by measles, and in the very young and old.

2. Three of the high DALY scores of Lopez et al.[S4] were not assigned by them to specific microbes but to "wastebasket categories" of shared symptoms caused by many different microbes, whose relative contributions are often poorly understood and not separately diagnosed in many or most individual cases. From the category "lower respiratory disease," we adopted influenza virus as the single microbe making the largest contribution; the other major causes are commensals that are often not specifically diagnosed. From the category "diarrheal diseases," we adopted rotavirus A as the single microbe making the largest contribution, followed at a distance by cholera; other notable diarrheal agents include Escherichia coli strains, Campylobacter, Shigella, Salmonella, Giardia, and Entamoeba histolytica, but they are usually lumped under the diagnosis of "diarrhea" and their individual contributions are poorly known.

From the category "meningitis," we adopted no microbes; its leading causes are respiratory commensals.

3. Lopez et al.[S4] attributed a DALY score to "malaria," without distinguishing among the four human malarias. We adopted falciparum malaria because it is by far the most important human malaria, and also vivax malaria because its estimated contribution to malarial morbidity would give it a DALY score within the top seven tropical diseases, although towards the lower end.

4. Lopez et al.[S4] attribute a DALY score to "trypanosomiasis" exclusive of Chagas' disease (New World trypanosomiasis), i.e., African sleeping sickness. Our Table S1 lists West and East African sleeping sickness separately, because they differ markedly in duration, role of animal reservoir, and disease stage, and are considered different subspecies of one microbe species.

5. Lopez et al.[S4] attribute a DALY score to "leishmaniases." Our Table S1 specifically lists only visceral leishmaniasis, the form imposing the heaviest burden; New World and especially Old World cutaneous leishmaniases impose much lower burdens.

The DALY scores of Lopez et al.[S4] thus led to our including 17 microbial diseases. But considerations of historical mortality support inclusion of eight other diseases that formerly were major killers and ceased to be so only recently: smallpox now eradicated, and plague, typhoid, typhus, yellow fever, mumps, diphtheria, and rubella now greatly decreased, due to modern vaccination, sanitation, and/or vector control. European-introduced diseases that proved major killers of previously unexposed Native American, Pacific Islander, Khoisan, and Aboriginal Australian populations included measles, mumps, pertussis, smallpox, syphilis, tuberculosis, and typhus[S5-S7]. These diseases caused higher historical mortalities in these non-European

populations than in contemporary European populations for several reasons: the former populations' lack of acquired immunity arising from previous individual exposure, lack of genetic resistance arising from historical exposure of the population, and first exposure in adulthood whereas most Europeans became exposed in childhood (some diseases such as measles, rubella, and hepatitis B cause higher mortality in adults than in children). Conversely, we also include several diseases that cause major mortality today but whose global toll in the past was either zero (AIDS) or considerably lower (cholera, dengue fever).

Likely origins. Conclusions about the animal sources of a particular human microbe are obviously limited by testing bias -- i.e., which animal species were selected for testing, and which pathogens were sought in them. Given that fact, a certain animal microbe currently considered to be the closest match to a corresponding human microbe might not be its closest relative at all; the true animal ancestor could eventually prove to be a microbe found in some animal species not yet studied. This is clearly true for rubella virus; its confinement to large human populations shows that it must have evolved within the last 11,000 years, and is thus likely still to bear a close relationship to some ancestral animal microbe, yet rubella virus has no remotely close relatives among known microbes, implying that we have not yet found its close animal relative. In the case of *Plasmodium reichenowi* of chimpanzees, we know of its existence just from one isolate; if that isolate hadn't been found, we would not know the closest relative of human P. falciparum. At present, it appears that smallpox is most closely related to camelpox, but that conclusion is tentative because there might be some more closely-related poxviruses still awaiting discovery. Camelpox itself might prove to be a misnomer; the virus has been isolated from camels but has rarely been sought in other species. It is of note that the identity of the

closest animal relative is still unresolved or debated for two of our most important human diseases, smallpox and falciparum malaria.

The column "likely origins" of Table S1 lists the animal species that, based on present information, seems to us to be the most likely proximate source of the zoonotic infection leading to that human disease. For pathogens infecting humans but dependent for their persistence on an animal reservoir (e.g., East African sleeping sickness), the immediate origins of the human disease are probably in that animal reservoir. For pathogens now confined to humans, the origins in evolutionary time may be sought in the most closely related animal pathogen, if that has been correctly identified. But one can also ask about ultimate origins. For instance, influenza A's main proximate sources are domesticated pigs and ducks, into whose populations the virus arrives from its ultimate source in wild aquatic birds. Human measles may have evolved proximately from rinderpest of domestic cattle, which may in turn have received rinderpest from an ultimate source of wild ruminants. Similarly, it is sometimes suggested that the proximate source of human smallpox virus was a poxvirus transmitted from a domestic animal, but that that ancestral poxvirus ultimately derived from an African rodent. In Note S9's detailed discussions of individual diseases, we attempt to distinguish between proximate and ultimate hosts for those three diseases (influenza A, measles, and smallpox). But even when the closest animal relatives of a human pathogen have been correctly identified, uncertainty may persist (e.g., in the case of measles, smallpox, and tuberculosis) about the historical direction of inferred spreads between humans, domestic and wild ruminants, and rodents.

Note S2. "Temperate" vs. "tropical".

We separate diseases into these two categories, based on their distributions and likely origins. There is little uncertainty about the 10 diseases that we denote as "tropical": nine are still largely confined to the tropics (although vivax malaria and yellow fever have also caused some seasonal outbreaks in areas within temperate latitudes), cholera arose in tropical Southeast Asia although in recent centuries it has spread to cause epidemics in temperate latitudes, and AIDS now occurs independently of latitude but is known to have arisen recently in tropical Africa. While the 15 diseases that we denote as "temperate" appear or appeared regularly at temperate latitudes, most also appear or appeared in the tropics as well. Most are evidently of Old World origin, but their latitude of origin is generally unknown. This separation of 10 "tropical" from 15 "temperate" (or to be more precise, temperate plus tropical) diseases still permits us to discern numerous differences between these two groups of diseases, discussed in the text.

Note S3. Robustness tests

Would a different selection of "the major human infectious diseases" lead to different conclusions? To test whether our conclusions are robust to the selection criteria, we drew up three alternative lists of "the major human infectious diseases" and performed all 10 analyses that our text describes for our chosen list of 25 diseases.

About the diseases that impose the heaviest burdens today or did so in the past, probably any other author drawing up a list would agree on including AIDS, influenza, malaria, plague, smallpox, tuberculosis, and up to a dozen other choices. However, other authors may differ on where to draw the line among the next most important diseases. Specifically, if one takes DALY values for the whole world as an objective (albeit not perfect) assessment of modern burdens, and if one ranks diseases according to DALY values, how far down the list should one go and still consider a disease to be "major"?

The highest DALY values[S4] are for influenza A and AIDS. Eleven diseases have DALY values equal to 5% or more of those influenza and AIDS values. There is then a nearly two-fold drop to the six next highest DALY values equal to between only 3% and 0.7% of the influenza/AIDS values (for hepatitis B, visceral leishmaniasis, West and East African trypanosomiasis, Chagas' disease, and dengue fever). One might thus reason that a minimum modern list would consist only of the 11 most burdensome diseases; a maximum list (the one that we actually used for our analyses reported in the text), those 11 plus those six next highest diseases; and an intermediate list, like the maximum list but deleting the diseases with the two lowest DALY values (Chagas' disease and dengue fever). It would be difficult to defend a DALY-value-based list of diseases more inclusive than that maximum list of 17, because any

additional diseases would then have DALY values less than 0.7% of the values of the leading diseases, i.e. they would have to be considered as minor rather than major and would mean "scraping the bottom of the barrel."

In addition, our maximum list of 25 diseases includes eight that we selected on the basis of their high historical burdens despite modest or negligible modern DALY values. We do not have an objective method, comparable to the method based on DALY values, for assessing historical impacts. However, there can be little doubt about the historical impacts of five diseases (plague, smallpox, typhoid, typhus, and yellow fever), which we accordingly added to all three of our lists (maximum, intermediate, and minimum). Hence our minimum list of 16 diseases consists of the 11 most burdensome modern diseases (AIDS, cholera, falciparum malaria, influenza A, measles, pertussis, rotavirus A, syphilis, tetanus, tuberculosis, and vivax malaria) plus the five additional leading historical diseases (plague, smallpox, typhoid, typhus, and yellow fever). Our maximum list of 25 diseases adds six modern diseases (Chagas' disease, dengue fever, East African sleeping sickness, hepatitis B, visceral leishmaniasis, and West African sleeping sickness), plus three more historical diseases less significant than plague, smallpox, typhoid, typhus, and yellow fever (diphtheria, mumps, and rubella). Our intermediate list of 21 diseases differs from our maximum list of 25 by deleting its two modern diseases with the lowest DALY scores (Chagas' disease and dengue fever) and its two historical diseases considered least significant (mumps, rubella).

We then performed, on all three lists, the 10 analyses reported in the text for our maximum lists: Old World vs. New World origins; and tropical-vs.-temperate differences in insect vector transmission, Stage 5, long-lasting immunity, animal reservoirs, environmental reservoirs, acute course, crowd epidemic disease, domestic animal origins, and wild non-human-

primate origins. With one minor exception, all 10 analyses yielded qualitatively the same directions of differences for all three tests. (That single exception is that Stage-5 diseases barely outnumbered non-Stage-5 diseases among both tropical and temperate diseases on the minimum list, but that Stage-5 diseases predominated among temperate diseases while non-Stage-5 diseases predominated among tropical diseases on the maximum and intermediate lists). As expected merely from considerations of sample size, significance levels of the statistical comparisons were highest for the maximum lists (n = 25), lower for the intermediate list (n = 21), and lowest for the minimum list (n = 16).

Thus, our conclusions are robust to variation in the selection criteria, provided that one agrees with our preference for DALY values as an objective assessment of modern disease burdens, and with our non-objective judgments about historic burdens. We encourage other authors who can think of other selection criteria to test them for comparison.

Note S4. Paucity of major diseases before Stage 4

Combining temperate with tropical diseases, 14 of our 25 major diseases are Stage-5 (confined to humans), 6 are Stage-4 (transmitted to humans both from animals and from other humans), and only 5 belong to Stage 2 (East African sleeping sickness, tetanus) or Stage 3 (Chagas' disease, plague, visceral leishmaniasis). The paucity of Stage-2 and Stage-3 diseases on our list of major human diseases inflicting high mortality is striking, because some Stage-2 and Stage-3 pathogens (such as anthrax, rabies, Ebola, and Marburg) are nevertheless virulent, feared, and objects of much attention. À priori, one could have imagined that many of these virulent pathogens would have become major killers on the scale of smallpox and cholera -- but few (notably, plague) actually have.

Was this conclusion, that most major human diseases belong to Stage 5 or Stage 4, just a fore-ordained outcome and a consequence of circular logic? For instance, one might reason à priori that that outcome was inevitable, e.g. because human-to-human contact will always be more frequent than animal-to-human contact, and because confinement of a pathogen to humans implies long evolutionary time for a pathogen to evolve special virulence for humans. More often, though, authors have used à priori reasoning to predict exactly the opposite outcome: "The received wisdom, set forth in most medical texts and elsewhere, is that 'successful' or 'well-adapted' parasites are relatively harmless to their hosts…that it is clearly in the interest of the parasite population not to harm its host population too much" (ref. S8, p. 648). In reality, Anderson and May[S8] argue, both of those opposite outcomes are theoretically possible and are realized in practice by different host/pathogen systems, because the mechanism by which the

pathogen damages its host is often but not always linked to the mechanism by which the pathogen becomes transmitted.

Our data base permits a test of these predictions. Nine of our 14 Stage-5 diseases, but also two of our non-Stage-5 diseases, sicken or kill their victims by causing a host reaction that serves to transmit the pathogen (diarrhea in the cases of typhoid and rotavirus, respiratory inflammation with coughing and sneezing in the cases of measles and pertussis and five other diseases). Conversely, five of our Stage-5 diseases, and nine of our non-Stage-5 diseases, sicken or kill by mechanisms unrelated to the pathogen's mode of transmission (e.g., sexual in the cases of AIDS and syphilis and hepatitis B, via an insect vector in the cases of malarias and yellow fever). Nor are Stage-5 diseases necessarily major killers: some, such as chickenpox and Shigella sonnei and human rhinovirus infections, are normally mild. Thus, many alternative outcomes are indeed possible, and only empirical data rather than à priori reasoning can tell us the predominant outcomes.

Why does it turn out that, among Stage-2 and Stage-3 diseases, only five have become major killers? At least some are clearly limited in their world toll by limited geographic range and/or low infection rates of animal hosts, inefficient transmission from infected animals to humans, or all three factors. For instance, although untreated rabies is always fatal, it is transmitted to humans only by bites of infected animals (usually dogs or bats), and relatively few dogs and bats become infected and then bite humans, so only a modest number of humans die of rabies. Similarly, while the case-fatality ratio of anthrax can be 20 - 100%, human infections mainly involve people who handle infected livestock, but relatively few livestock are infected and few people shear sheep or prepare cowhides, hence anthrax's human death toll remains small. Among trypanosomiases, East African sleeping sickness currently infects (and would kill

without treatment) several hundred thousand people, but only in tropical Africa within the ranges of its vector (tsetse flies) and of its infected animal hosts; and Chagas' disease in the New World tropics is also limited geographically by its vector (reduviid bugs) and animal hosts (and is now even more limited by modern housing). Wider vector and host distributions could turn these two trypanosomiases into bigger killers. Finally, current evidence suggests that Ebola and SARS reach humans mainly via chains of mammalian hosts (bats to apes and duikers, and bats to civets and other mammals, respectively). Both of these diseases are virulent and feared, but limited human contact and infection rates of host mammals may contribute to limiting their human toll as well.

Note S5. Animal origins of temperate diseases

Among the eight candidates for domestic animal origins (six of which are crowd diseases), the evidence is compelling for influenza A, because new human strains are continuing to arise today from domestic ducks and pigs, which in turn become infected from wild birds. Measles virus probably evolved from rinderpest virus of cattle, diphtheria bacillus and pertussis and rotavirus from pathogens of one or more of several domestic herbivores (cow, sheep, goat, pig, horse). In the latter three cases each ancestral pathogen has also been reported in wild mammals, but the proximate source of the human pathogen is more likely to have been a domestic than a wild mammal, because human contacts with domestic mammals are much closer and more frequent. Mumps may have arrived from pigs, smallpox from camels, and tuberculosis from cattle, but in each of these cases there remains some uncertainty about the ancestral pathogen's identity or host range or (in the case of tuberculosis) direction of spread.

Note S6. Animal origins of tropical diseases

Domestic animal origins can be immediately excluded for six of the 10 tropical diseases of Table S1: AIDS, dengue fever, vivax malaria, and yellow fever are derived instead from wild primates, cholera from aquatic algae and invertebrates, and falciparum malaria ultimately from birds. The pathogens ancestral to the agents causing Chagas' disease, West and East African sleeping sickness, and visceral leishmaniasis infect both domestic and wild mammals. However, at least in the case of Chagas' disease, wild mammals are much more likely to have furnished the earliest ancestor, because the agent of Chagas' disease is still transmitted to humans today from over 150 species of wild animals, and Native Americans in tropical South and Central America (where Chagas' disease occurs) traditionally kept only few species and individuals of domestic animals (just dogs, llamas, Muscovy Ducks, and guinea pigs)[S9].

Note S7. Tropical domestic animals

The sole abundant domestic animal to have originated in the tropics is the chicken (tropical Southeast Asia). Smaller and more localized populations of water buffalo, Muscovy Duck, and llama and guinea pig originated in tropical Southeast Asia, tropical South America, and the Andes, respectively, while even smaller and more localized populations of banteng and gaur originated in Bali and tropical Southeast Asia, respectively[S9]. In addition, before Columbus but only within the last few thousand years, some domestic animals of Old World temperate origin (cow, sheep, goat, horse) spread into tropical Africa, where their numbers remain low due to disease.

Note S8. Our speculative guesses about the next big killers

Hoping that we shall not look foolish in 10 years, we offer some speculations about new major diseases of the future. While the principles of pathogen transmission have not changed during the last 11,000 years, changing modern conditions are exposing us to new pathogen reservoirs and new modes of transmission. The world is now different from ancient Eurasia, in which most people were farmers and herders in close contact with domestic animals, and in which many of our 25 major diseases of Table S1 evolved. Here are some of our guesses.

First, here is our opinion about which microbes will not become the next big killers. We believe that some pathogens arouse fear and attract resources out of proportion to their real threat to humanity. In particular, case-fatality ratio is often mistakenly equated with danger, but some pathogens greatly feared for their proven ability to kill a few hundred people are too inefficiently transmitted to kill millions. For example, while anthrax, BSE, Ebola, and Marburg do kill a high fraction of infected victims (100% in the case of BSE), their inefficient transmission assures few infected victims; 100% of a small number is still a small number. Anthrax and BSE are not transmitted from human to human (unless by cannibalism and organ transplants in the case of BSE-type prions); anthrax is treatable by antibiotics; and human-to-human transmission of Ebola and Marburg fades out after a few transfers and is unlikely to occur in the first place outside of specific settings, such as rural hospitals lacking practices and supplies for controlling infections. For anthrax, Ebola, hantavirus, and Marburg the rapid onset and severity of symptoms have made identification and containment feasible. Similarly, some emergent Stage-2 pathogens, such as West Nile Virus, hantavirus, and Lyme Disease's agent, have aroused more fear than their lack of human-to-human transmission and modest burden of morbidity and mortality seem to us

to warrant. But we acknowledge that virulent pathogens currently exacting a low death toll due to inefficient transmission could become dangerous if they evolved new modes of more efficient transmission (e.g., by aerosolized respiratory droplets) -- as may have almost occurred in 1989 with the Reston subtype of Ebola virus[S10].

If not anthrax and Marburg, which pathogens do we instead fear? A straightforward prediction, often neglected in our focus on the exotic, is that some new pandemics will emerge from pathogen taxa of which strains have already caused historical pandemics. Prime candidates are two microbes persisting in animal and/or environmental reservoirs capable of generating new strains: influenza virus and, under local conditions of failing hygiene, possibly cholera. We envision these two important diseases of the past again becoming important diseases of the future. Another prime candidate is tuberculosis, of which new strains have arisen through the development of drug resistance, and which is already causing pandemics among human subpopulations with weakened resistance, such as HIV patients.

A further prediction involves emerging pathogens transmitted by routes that render their spread difficult to control. Sexually transmitted diseases (STD's) fall in this category, because it is difficult to persuade us to abstain from or change our sexual behavior. HIV offers a grim warning: despite its swift and huge global impact, the AIDS epidemic would have been much worse if the sexual transmissibility of HIV (actually rather modest[S11]) had equaled that of other sexually transmitted agents such as HPV. Similarly, it would be difficult to control emerging pathogens transmitted by our pets (which increasingly include many exotic species as well as traditional domesticated breeds), in chains connecting wild or feral animals to outdoor pets to humans[S12]. While we have in some cases reluctantly accepted the culling of millions of farmyard chickens and cows as the price of stemming the spread of avian influenza and BSE, it

is hard to imagine our killing millions of our children's beloved puppies, bunnies, and kittens and of our village hunting dogs, even if those animals did offer a likely entry portal for a dangerous pathogen. Hence monitoring for the emergence of new STD's and pet-associated diseases would both be good investments.

Still other opportunities exist for novel pathogens to emerge from the big and often-discussed changes associated with modern human societies (see Table 1 in Box 3), such as urbanization, global travel and trade, evolution of drug-resistant microbes, climate change, and increasing numbers of elderly, antibiotic-treated, immuno-suppressed people. For instance, the high rate of urbanization in Africa could transform yellow fever, Chikungunya virus disease, and other African arbovirus diseases into urban diseases, as has already happened with dengue hemorrhagic fever. Globalization, by connecting distant places, permits long-distance transfers of microbes and their vectors, as witnessed by recent North American cases of cholera and SARS brought by infected passengers on jet flights from South America and Asia respectively. Although we do have a vaccine against yellow fever, most of the world's human population remains unvaccinated (e.g., because of the disease's still-unexplained absence from Asia), so that a yellow fever epidemic especially in China or India could be devastating. Global warming is causing tropical birds and arthropods to expand their ranges into temperate zones[S13]. As tropical arboviruses and arthropod vectors join the ranks of those expanding species, we anticipate temperate flare-ups of many diseases now dismissed as "just" tropical.

All these examples illustrate that future disease control efforts must not only continue the traditional approaches to combating on-going pandemics but must also place more emphasis on "disease forecasting," the early detection of novel pandemics. As summarized in this review,

most of our major human pathogens have emerged from animal reservoirs, and some types of animals and modes of transmission have been especially productive of human killers.

Note S9. Details on each of our 25 major diseases

AIDS (HIV-1)

HIV-1 is derived from Simian Immunodeficiency Virus (SIVcpz), found in chimpanzees (subspecies *Pan troglodytes troglodytes*)[S14]. The cross-species transmission of SIVcpz and its subsequent adaptation to humans most likely occurred by the early 20th century[S15]. This cross-species transmission occurred in western equatorial Africa, inferred from viral diversity and the presence there of SIVcpz strains most similar to HIV-1 strains[S15]. SIVcpz has crossed into humans on at least three or four occasions, distinguished by phylogenetic analysis of HIV-1.

VIRUS FAMILY AND GENUS (Note: Virus taxonomy follows ICTV[S16])**:** Retroviridae: *Lentivirus*

VIRUS TYPE: enveloped, positive single-stranded RNA

NEAREST MODERN RELATED VIRUS: SIVcpz

PROXIMATE HOST: Chimpanzee (*Pan troglodytes troglodytes*)

GEOGRAPHIC ORIGIN: Western Equatorial Africa

CHAGAS' DISEASE

Chagas' disease is caused by *Trypanosoma cruzi*. It descends from a line of trypanosomes that includes a number of marsupial host species and only a few Old World bat host species. The clade that gave rise to *T. cruzi* is thought to have diverged from other trypanosomes on the landmass that became Australia and South America, "when Africa became isolated from the other continents" during the mid-Cretaceous[S17]. *T. cruzi* may have adapted to humans as they populated regions where sylvatic cycles were already occurring in a variety of mammalian hosts. There is paleopathological evidence that some of the early settlers of South America were infected as of approximately 9000 years ago[S18]. The closest known relatives of *T. cruzi* today, based on 18S rRNA, include *T. dionisii*, *T. vespertilionis*, *T. conorhini*, *T. minasense*, *T. rangeli*, and *T. leeuwenhoeki* [S19].

There remains disagreement about the wider phylogenetics of trypanosomes; although Lukes *et al.*[S20], Stevens and Gibson[S21], and Hamilton *et al.*[S22] have argued for monophyly, Hughes and Piontkivska[S19, S23] have used their 18S rRNA work to maintain the idea that trypanosomes are paraphyletic.

PROTOZOAN FAMILY AND GENUS: Trypanosomatidae: *Trypanosoma*

ARTHROPOD VECTOR: Triatomine (reduviid) bugs

NEAREST MODERN RELATED PROTOZOAN: *T. cruzi* is the outgroup of a group of trypanosomes which include *T. dionisii*, *T. conorhini*, and *T. vespertilionis*.

PROXIMATE HOST: over 150 different mammal species[S24]

GEOGRAPHIC ORIGIN: New World tropics

CHOLERA

Vibrio cholerae is mostly closely related to *Vibrio mimicus*[S25]. *V. cholerae* has been found on aquatic algae, crustaceans, copepods, and amoebae[S26], although birds[S27, S28] and terrestrial herbivores[S29] can also harbor the pathogen. The GbpA binding protein, which helps *V. cholerae* bind to the chitinous exoskeletons of zooplankton, also appears to help the bacteria colonize the intestines of larger fauna[S30]. The first cholera pandemic began in Asia in 1817; the seventh, caused by the El Tor biotype, began in 1961. The new serotype 0139 appeared in 1992.

BACTERIAL FAMILY AND GENUS: Vibrionaceae: *Vibrio*

BACTERIA TYPE: Gram-negative bacillus

NEAREST MODERN RELATED BACTERIUM: *V. mimicus*

PROXIMATE HOST: marine organisms

GEOGRAPHIC ORIGIN: Tropical Southeast Asia

DENGUE FEVER

Dengue fever is caused by dengue fever virus serotypes 1-4, which are within the Flaviviridae family. Dengue virus occupies its own clade when nucleotide sequences are used for analysis, but clusters with Kedougou virus when amino acid sequences are analyzed[S31]. Wang *et al*, who used envelope protein sequences in their phylogenetic analysis, speculate that the ancestral sylvatic form of the virus, affecting nonhuman primates, probably arose in Asia/Oceania before diverging into human forms, possibly in different geographic regions[S32]. Others posit an African origin for dengue, supported by divergence of viral strains and studies of vector evolution[S33]. Several of the most closely related known extant viruses appear to have African nonhuman primate reservoirs[S34]. Because of limited samples from both regions, the geographic origins of the four dengue virus serotypes remain unresolved[S32, S34].

VIRUS FAMILY AND GENUS: Flaviviridae: *Flavivirus*

VIRUS TYPE: positive single-stranded RNA

ARTHROPOD VECTOR: *Aedes* mosquitoes

NEAREST MODERN RELATED VIRUS: Kedougou virus?

PROXIMATE HOST: Old World nonhuman primates

GEOGRAPHIC ORIGIN: Old World tropics

DIPHTHERIA

Martin *et al*.[S35] determined via ribosomal protein and RNA that *Corynebacterium diphtheriae*, the causative agent of diphtheria, is most closely related to *C. pseudotuberculosis* and *C. ulcerans*, and less closely to *C. vitaruminis* and *C. kutscheri*. *C. pseudotuberculosis*[S36] and *C. ulcerans*[S35] are known to infect domestic animals and livestock, including cows, horses, goats, and sheep. In addition, pathogenic *C. ulcerans* has been isolated from ground squirrels[S37] and otters[S38]. Voles are susceptible to *C. kutscheri*[S39].

BACTERIAL FAMILY AND GENUS: Corynebacteriaceae: *Corynebacterium*

BACTERIA TYPE: gram-positive

NEAREST MODERN RELATED BACTERIUM: *C. pseudotuberculosis* and *C. ulcerans*

PROXIMATE HOST: probably domestic livestock

GEOGRAPHIC ORIGIN: Old World; did not arrive in New World until after A.D. 1700[S5].

FALCIPARUM MALARIA

P. falciparum is most closely related to *P. reichenowi*, a parasite of chimpanzees, based on an analysis of nucleotide polymorphisms[S40]. Several studies have suggested that extant populations of *Plasmodium falciparum* diverged relatively recently[S40, S41], but the timing of this divergence remains vigorously debated. Some authors argue that the association of *P. falciparum* with hominids is ancient and predates the divergence between humans and

chimpanzees, while others argue for an "intermediate origin"[S42] (i.e., only thousands of years ago). Martin *et al*.[S43] suggest, based on a functional analysis of red blood cell receptors, that the fact that *P. falciparum* can infect *Aotus* monkeys in the New World is due to convergent red blood cell evolution in the host species. The plasmodia most closely related to, and presumed ultimately ancestral to, *P. falciparum* and *P. reichenowi* are ones infecting birds; the debate concerns whether that jump from birds to humans and chimpanzees occurred over 5 million years ago or within the past 10,000 years.

PROTOZOAN FAMILY AND GENUS: Plasmodiidae: *Plasmodium*

ARTHROPOD VECTOR: *Anopheles* mosquitoes

NEAREST MODERN RELATED PROTOZOAN: *P. reichenowi* (chimpanzees)

ULTIMATE HOST: birds

GEOGRAPHIC ORIGIN: Tropical Africa

HEPATITIS B

The origin and evolution of hepatitis B virus are still contested. The phylogenetic tree prepared by Simmonds[S44] suggests that human hepatitis B viruses are closely related to ape HBVs, with strains found in chimpanzees, orangutans, gibbons, and gorillas. Analysis of a region of the S gene shows human strains clustering with a chimpanzee strain and all other

nonhuman primate HBV strains forming a separate clade, although there is evidence for recombination among several strains[S45]. Robertson's[S46] unrooted tree places the Old World nonhuman primate HBVs in a separate clade from human HBVs. Like the tree from Takahashi *et al*.[S47], Robertson's tree shows a woolly monkey strain of HBV closely related to human HBV type F, found in the New World; perhaps the former represents a recent transfer from humans.

VIRUS FAMILY AND GENUS: Hepadnaviridae: *Orthohepadnavirus*

VIRUS TYPE: circular, double-stranded DNA virus (reverse transcriptase)

NEAREST MODERN RELATED VIRUS: chimpanzee HBV

PROXIMATE HOST: apes

GEOGRAPHIC ORIGIN: Old World (except type F)

INFLUENZA A

Waterfowl and shorebirds, including ducks, are the primary nonhuman reservoirs for influenza A strains. Influenza viruses may be transmitted directly from them to humans[S48], via poultry that have had contact with other birds[S48, S49], or by reassortment with avian, porcine, and/or human strains[S50].

VIRUS FAMILY AND GENUS: Orthomyxoviridae: *Influenzavirus A*

VIRUS TYPE: negative single-stranded RNA

NEAREST MODERN RELATED VIRUS: avian influenza strains

PROXIMATE HOSTS: poultry, occasionally pigs

ULTIMATE HOSTS: waterfowl

GEOGRAPHIC ORIGIN: Old World, probably Asia

LEISHMANIASIS (Visceral)

Leishmaniasis is caused by protozoa of the genus *Leishmania*. The visceral leishmaniases (kala-azar) are transmitted to humans by phlebotomine sandflies, usually from infected dogs and occasionally from other humans. There have been disagreements concerning the taxonomic nomenclature of the *Leishmania* causing visceral disease[S51] and the cladistic robustness of the phyla[S52], but the organisms most commonly implicated are Leishmania of the *L. donovani* complex, including *L. donovani donovani* and *L. donovani infantum*[S53]. *L. donovani chagasi*, which is probably the same subspecies as *L.d. infantum*[S51], has been isolated from visceral cases in the Americas. *L. amazonensis*, *L. tropica*, and *L. columbiensis* and *L. major* IV usually cause cutaneous leishmaniasis but have occasionally been reported from visceral cases[S54-S57] as well. *L. d. donovani* itself appears to remain an Old World pathogen, possibly originating in East Africa[S58, S59]. The geographic origins and spread of *L. d. infantum* (/*L. d. chagasi*) remain unresolved; although present throughout the Mediterranean, Asia, the Middle East, and Africa, *L. d. infantum* infection has also been detected in a variety of native fauna of

the Brazilian Amazon[S60] and in canids in French Guyana [S61] as well. Old World and New World cutaneous leishmaniases are distinct from, and less severe than, visceral leishmaniasis.

PROTOZOAN FAMILY AND GENUS: Trypanosomatidae: *Leishmania*

ARTHROPOD VECTOR: Phlebotomine flies

NEAREST MODERN RELATED PROTOZOA: Various *Leishmania* species

PROXIMATE HOST: Canids, most commonly dogs, but human-to-human transmission also occurs

GEOGRAPHIC ORIGINS: Old World, possibly East Africa (*L. donovani donovani*).

MEASLES

Measles virus is in the *Morbillivirus* genus of the Paramyxoviridae family. It is most closely related to rinderpest virus, which infects cattle and other wild and domestic ruminants[S62, S63]. This conclusion has been reached via phylogenetic analysis of fusion and matrix proteins[S64], attachment proteins[S65, S66],and amino acid sequences[S67]. Measles virus was first described in the 10th century CE[S68], but not earlier by Hippocrates.

VIRUS FAMILY AND GENUS: Paramyxoviridae: *Morbillivirus*

VIRUS TYPE: negative single-stranded RNA

NEAREST MODERN RELATED VIRUS: Rinderpest virus (cattle and other ruminants)

PROXIMATE HOST: Probably cattle

ULTIMATE HOST: Probably wild ruminants

GEOGRAPHIC ORIGIN: Old World

MUMPS

Chang *et al*., using NP, P, F, HN, and L proteins[S69], Westover and Hughes, using fusion proteins[S64], and Seal *et al*.[S65, S66] place mumps virus on its own in a clade that also contains the simian parainfluenza viruses (which affect many mammals) and human parainfluenza virus type 2. Kawano *et al*.[S70] placed mumps in a clade with PIV-2 and SV-5 (aka PIV5), the latter of which affects a variety of mammals. Two recent analyses[S71, S72] that have included the porcine rubulavirus LPMV suggest that it may also be closely related to mumps virus. The human parainfluenza virus types 4a and 4b appear more distantly related in the rubulavirus group except when fusion proteins alone[S64] or matrix proteins were used[S69]. Mumps itself has been infecting humans since before the time of Hippocrates (5th century BC)[S73].

VIRUS FAMILY AND GENUS: Paramyxoviridae: *Rubulavirus*

VIRUS TYPE: negative single-stranded RNA

NEAREST MODERN RELATED VIRUS: porcine rubulavirus, simian and human parainfluenza viruses

PROXIMATE HOST: unresolved mammal species; possibly pigs

GEOGRAPHIC ORIGIN: Old World

PERTUSSIS

Pertussis is caused mainly by *Bordetella pertussis,* and less often and more mildly by *Bordetella parapertussis*. These two species are thought to have evolved separately from an ancestral bacterium, *Bordetella bronchiseptica*, which "infects and can cause diseases in a wide range of mammals from marsupials through to ungulates, rodents and carnivores"[S74]. *B. bronchiseptica* is also known to occasionally infect humans. The evolution of human-adapted strains of *Bordetella* is thought to have occurred separately for *B. pertussis and B. parapertussis*; Bjornstad and Harvill[S74] found each of these two species to be more closely related to *B. bronchiseptica* than to each other. Van der Zee *et al*.[S75] found *B. bronchiseptica* to cluster more tightly with *B. parapertussis*, while *B. pertussis* was largely a separate clade. The similarity of all *B. parapertussis* strains analyzed led Van der Zee *et al*.[S75] to conclude that *B. parapertussis* evolved and spread fairly recently, while *B. pertussis* has been in humans longer.

BACTERIAL FAMILY AND GENUS: Alcaligenaceae: *Bordetella*

BACTERIA TYPE: aerobic, gram-negative, rod-shaped bacterium

NEAREST MODERN RELATED BACTERIUM: *B. bronchiseptica*

PROXIMATE HOST: Mammals of unknown species, because *B. bronchiseptica* has a wide host range among mammals and can also be found in birds.

GEOGRAPHIC ORIGIN: Old World; may have arrived in New World in the 16th century[S5].

PLAGUE

Yersinia pestis is commonly found in rodents and their parasites, with the ability to infect and sicken other mammals as well. Using 16S rDNA, Kim *et al*.[S76] found *Yersinia pestis*, the causative agent of plague, to be extremely close to *Yersinia pseudotuberculosis*, which is also pathogenic, but less frequently fatal, in humans. Indeed, Achtman *et al*.[S77] classify *Y. pestis* as "a recently emerged clone of *Yersinia pseudotuberculosis*". They[S77] place bounds of 1500 to 20,000 years ago on the evolution of *Y. pestis*, assuming a molecular clock similar to that of *E. coli*. Dykhuizen[S78] suggests that *Y. pestis* may have diverged from other strains of *Y. pseudotuberculosis* around the time of the first recorded plague pandemics. Three biovars of *Y. pestis* have been identified from analysis of phenotypes, and it has been suggested that biovar Orientalis is associated with the plague currently in circulation[S79]. Achtman *et al*.[S80], however, disagree and urge caution in making connections between genotypes and specific pandemics.

BACTERIAL FAMILY AND GENUS: Enterobacteriaceae: *Yersinia*[S81]

BACTERIA TYPE: "gram-negative, non-motile, nonsporulating coccobacillus"[S81]

ARTHROPOD VECTOR: fleas

NEAREST MODERN RELATED BACTERIUM: *Y. pseudotuberculosis*

PROXIMATE HOST: rodents

GEOGRAPHIC ORIGIN: Old World, traditionally thought to be Eurasia[S82], but there is also some evidence for a long history in Africa[S83]

ROTAVIRUS A

Six groups of rotaviruses (groups A through F) are known in humans and animals; group A causes most human disease. Ito *et al*.[S84] used amino acid sequences to determine that some strains of rotavirus affecting humans cluster with bovine, porcine, and equine rotaviruses, but they are more distant from avian rotaviruses. Rabbit-derived strains have also been isolated from ill children in Europe[S85, S86]. Rahman *et al*[S87], analyzing a human infection in Bangladesh that was genotypically similar to a porcine infection, and Van der Heide *et al*.[S88], comparing bovine rotaviruses found in the Netherlands to a human case that had been reported in Hungary, both noted the potential importance of reassortment in the development of human rotavirus strains.

VIRUS FAMILY AND GENUS: Reoviridae: *Rotavirus*

VIRUS TYPE: double-stranded RNA

NEAREST MODERN RELATED VIRUS: Many closely-related viruses have been found in domestic animal hosts as well as in rodents and simians.

PROXIMATE HOST: livestock animals (cattle, pigs, horses)

GEOGRAPHIC ORIGIN: unknown

RUBELLA

Rubella virus, known only to affect humans, has been classified taxonomically within the Togaviridae as the lone member of its genus[S89]. Recent "phylogenetic comparison of the RDRP or helicase domains places rubella equidistant between alphaviruses and hepatitis E viruses" (Frey, pers. comm.), and it shows some similarity in non-structural protein open reading frames to beet yellow necrotic virus and hepatitis E[S90].

VIRUS FAMILY AND GENUS: Togaviridae: *Rubivirus*

VIRUS TYPE: positive single-stranded RNA

NEAREST MODERN RELATED VIRUS: no known close relatives

PROXIMATE HOST: unknown

GEOGRAPHIC ORIGIN: unknown

SLEEPING SICKNESS (East and West African)

The trypanosomes responsible for sleeping sickness are *T. brucei rhodesiense* (East Africa) and *T. brucei gambiense* (West and Central Africa). These two forms differ so markedly in duration of disease (much longer for West than East African sleeping sickness) and in the importance of animal-to-human transmission (far more frequent for East African sleeping sickness than for West African sleeping sickness, which tends to be maintained in human populations) that our paper analyzes them separately. *T. brucei brucei* is not thought to be pathogenic in humans[S91]. MacLeod *et al*., using minisatellite variant mapping, found that *T. brucei* subtypes capable of infecting humans adapted from other primates in several separate events[S91]. A recent evaluation of 18S rRNA sequences show *T. brucei* to be closely related to *T. evansi* and *T. equiperidum* (both equine)[S19], and to other primate-adapted trypanosomes. This result accords with previous work by Hughes and Piontkivska[S23], and Stevens and Gibson[S21].

PROTOZOAN FAMILY AND GENUS: Trypanosomatidae: *Trypanosoma*

ARTHROPOD VECTOR: Tsetse flies

NEAREST MODERN RELATED PROTOZOAN: *T. evansi* and *T. equiperdum*

PROXIMATE HOST: *T. brucei* infects many mammal species, but wild[S92-S94] and domestic[S95-S99] ruminants are the principal reservoirs.

GEOGRAPHIC ORIGIN: Tropical Africa

SMALLPOX

Smallpox is caused by variola virus, a member of the *Orthopoxvirus* genus in the family Poxviridae[S16]. Several analyses of gene families[S100], protein sequences[S100], and genomes[S101-S103] have shown that, of the known orthopoxviruses, variola virus is most closely related to camelpox. In one study, however, a variola-specific PCR picked up only cowpox isolates, and not camelpox or other isolates[S104]. One personal communication suggests that the closest extant relative of variola is tatera poxvirus, which affects African gerbils[S105], although Fenner notes that this and other poxviruses have not been well-studied[S106]. In their recent genomic analysis of variola and other orthopoxvirus strains, Esposito *et al*.[S107] list camelpox virus and a West African strain of tatera poxvirus as the "[orthopoxviruses] most closely related to [variola virus]". The *Orthopoxvirus* genus has a broad host range, but it has been suggested that the genus's origins are in African rodents, and it seems possible that camelpox, smallpox, and cowpox represent the adaptation of rodent poxviruses to domestic animals and humans. Monkeypox and cowpox, despite their names, are mainly diseases of rodents that occasionally infect monkeys and cows. It should be noted, however, that unlike some other orthopoxviruses, smallpox and camelpox have only been isolated from their currently known hosts (i.e. humans and camels).

VIRUS FAMILY AND GENUS: Poxviridae: *Orthopoxvirus*

VIRUS TYPE: enveloped, double stranded DNA virus

NEAREST MODERN RELATED VIRUS: Camelpox virus, found in camels, and tatera poxvirus, from African gerbils

PROXIMATE HOST: Possibly camels

ULTIMATE HOST: Most likely African rodents

GEOGRAPHIC ORIGIN: Old World, probably Africa

SYPHILIS (Venereal)

Treponema pallidum pallidum, the causative agent of venereal syphilis, is one of three *T. pallidum* subspecies, the other two being *T. pallidum pertenue* (yaws) and *T. pallidum endemicum* (bejel/endemic syphilis). A related treponemal disease, pinta, is caused by *T. carateum*. Of these, only *T. pallidum pallidum* is sexually-transmitted; the others are commonly transmitted among children and tend to affect primarily the skin, bones, and cartilage. Syphilis is found worldwide today, while yaws is predominantly a disease of rural tropical areas, bejel/endemic syphilis is found in dry rural regions, and pinta occurs in the Americas.

The origin of venereal syphilis and its relationship to other treponemes has been obscured by several factors: treponemal isolates are difficult to propagate in the lab, clear paleopathological evidence of disease is scant and has been subject to debate, and the pathogenic treponemes are serologically and morphologically indistinct. As a result, no consensus has emerged on the geography or the timing of the appearance of the different diseases.

Much of the work done to establish the origin of venereal syphilis has relied on paleopathological evidence – or the absence thereof – to diagnose the condition in antiquity. Rothschild[S108] concluded from analyses of skeletal remains that the venereal form of *T. pallidum* was present in the pre-Columbian New World and could be reliably distinguished from conditions like yaws and bejel, but other researchers disagree. There is evidence of syphilitic skeletons in the Old World, including remains in England[S109] that predate Columbus' voyage.

Treponemes, including some *T. pallidum* subspecies, are known to infect nonhuman primates as well as humans. In the 1960s and 1970s, serosurveys of baboons and other primates in yaws-endemic regions of West and Central Africa revealed numerous foci of treponematosis with standard human treponemal assays[S110-S112], but few isolates were preserved. Genes from one simian isolate were recently sequenced, and appeared most closely related to *T. pallidum pertenue* ; a pathogenic treponeme from rabbits, *T. paraluiscuniculi*, was also sequenced and found to be more distantly related[S113, S114]. Additional molecular studies of human and animal treponemes may help inform the debate over the origins and evolution of pathogenic strains.

BACTERIAL FAMILY AND GENUS: Spirochaetaceae or Treponemataceae: *Treponema*

BACTERIA TYPE: spirochaete

NEAREST MODERN RELATED BACTERIUM: *Treponema pallidum pertenue, T. pallidum endemicum*

PROXIMATE HOST: Humans (venereal syphilis); nonvenereal treponematoses appear to infect multiple primates

GEOGRAPHIC ORIGIN: Ongoing controversy as to whether New World or Old World

---

TETANUS

Tetanus (lockjaw) is caused by the bacterium *Clostridium tetani*, which is commonly found in soil and in intestines of animals and humans, and may be present in animal feces. It is usually transmitted to adults through contamination of wounds. Effective vaccines are routinely used in much of the world. In the case of neonates, birth on a soil surface[S115], nonsterile umbilical cord treatment, including the application of animal dung[S116] or dung-contaminated ghee[S117], and circumcision[S118] are risk factors.

BACTERIAL FAMILY AND GENUS: Clostridiaceae: *Clostridium*

BACTERIA TYPE: Gram-positive, anaerobic, spore-forming, rod-shaped bacterium

NEAREST MODERN RELATED BACTERIUM: *C. cochlearium*[S119]

PROXIMATE HOST: too widespread to specify.

GEOGRAPHIC ORIGIN: unknown; present worldwide

TUBERCULOSIS

Mycobacterium tuberculosis is closely related to a number of other mycobacterial species. Using their analyses of 16S rRNA and hsp65 genes, both Devulder *et al*.[S120] and Kim *et al*.[S121] grouped it with *M. caprae* (mainly cattle, plus red deer and wild boar), *M. microti* (voles and other mammals), *M. bovis* (cows and other mammals[S122]), and *M. africanum* (human, W./C. Africa)[S123]. Devulder *et al.* also arrived at the same grouping when rpoB and sod genes were used. Krauss *et al*.[S123] state, "*M. africanum* is thought to be intermediate between *M. tuberculosis* and *M. bovis*; its retention as a distinct species is probably not justified". Based on an analysis of genomic deletions, Mostowy *et al*.[S122] suggest that M. tuberculosis diverged from the group prior to *M. africanum*, *M. microti*, *M. pinnipedi*, *M. caprae*, or *M. bovis.*

BACTERIAL FAMILY AND GENUS: Mycobacteriaceae: *Mycobacterium*

BACTERIA TYPE: rod-shaped, non-motile

NEAREST MODERN RELATED BACTERIUM: probably *Mycobacterium bovis*

PROXIMATE HOST: possibly cattle, but the direction of evolution is debated, and the presence of *M. bovis* in cattle may be relatively recent

GEOGRAPHIC ORIGIN: Probably Old World[S124], but independent, New World origins are also debated.

TYPHOID

Using tuf and atpD genes, as well as 16S and 23S Rrna[S125], and 24 chromosomal loci[S126], *Salmonella enterica typhi* (aka *S. typhi*) was shown to be closely related to *S. paratyphi A.* While the latter, as well as more distantly related strains of *S. paratyphi*, can infect mammals and birds and occasionally humans, *S. typhi* is particularly specialized to humans[S126]. Selander *et al*. discuss several temporal and geographic evolutionary scenarios[S126]; Frankel *et al*.[S127] posited an Indonesian origin of *S. typhi*, but an African or other geographic origin is also possible[S126,128]. About 3% of infected people become chronic carriers.

BACTERIAL FAMILY AND GENUS: Enterobacteriaceae: *Salmonella*

BACTERIA TYPE: Gram-negative, rod-shaped

NEAREST MODERN RELATED BACTERIUM: *S. paratyphi*

PROXIMATE HOST: unclear and perhaps non-applicable, because related *Salmonella* species have a wide host range (e.g. pigs, birds, lizards, etc.)

GEOGRAPHIC ORIGIN: Old World

TYPHUS

Rickettsia prowazekii is the pathogen that causes epidemic human typhus; related rickettsiae of animals also infect humans. Andersson *et al*.[S129], using spacer regions and 23S rRNA, and Lee *et al*.[S130], using gro-EL gene analysis, found *R. prowazekii* to be closely related to *R. typhi*, which causes murine typhus and can also infect humans. *R. prowazekii* also infects flying squirrels[S123, S131], and *R. typhi* is common among rodents and may also be found in game birds[S123]. Another member of the typhus group rickettsiae, *R. felis*, has been isolated from North American opossums[S132]. The presence of *R. prowazekii* and related species in North American squirrels has been taken to suggest New World origins, but the global distribution of epidemic and murine typhus, and the ability of *R. prowazekii* to be spread among humans without the need for another host species, make geographic origins difficult to determine.

BACTERIAL FAMILY AND GENUS: Rickettsiaceae: *Rickettsia*

BACTERIA TYPE: rickettsial

ARTHROPOD VECTOR: fleas, body lice, and possibly others

NEAREST MODERN RELATED BACTERIUM: *Rickettsia typhi* (rodents, game birds)

PROXIMATE HOST: perhaps rodents

GEOGRAPHIC ORIGIN: unresolved

VIVAX MALARIA

Reconstruction of the evolutionary history of *Plasmodium vivax* was initially obscured by several apparently conflicting pieces of evidence from genetic diversity of African and Asian types, malaria-resistant phenotypes common among human populations in Africa, and close genetic similarity to a New World species (*P. simium*). Several recent studies converge on an Asian origin of *Plasmodium vivax*, probably from macaques[S133-S135], although the timing of the adaptation of *P. vivax* to humans remains unresolved. Escalante *et al.*[S134] constructed a phylogeny from one plastid gene and two nuclear genes, while Jongwutiwes *et al.*[S133] and Mu *et al.*[S135] used mitochondrial genome sequences in their analyses.

PROTOZOAN FAMILY AND GENUS: Plasmodiidae: *Plasmodium*

ARTHROPOD VECTOR: *Anopheles* mosquitoes

NEAREST MODERN RELATED PROTOZOAN: *P. simium* (New World monkeys); probably by very recent (post-Columbian) adaptation of *P. vivax* to New World monkey hosts[S33]

PROXIMATE HOST: Asian macaques

GEOGRAPHIC ORIGIN: Tropical Southeast Asia

YELLOW FEVER

Yellow fever is caused by yellow fever virus, a flavivirus. Forest primates in Africa and South America appear to serve as the primary hosts for sylvatic cycles of yellow fever[S136], although other forest mammals may also carry the virus[S137]. The viruses ancestral to yellow fever appear to have originated in the Old World, probably Africa[S138]. Yellow fever virus has been shown to be most closely related to the group of viruses that includes Wesselsbron virus (sheep, rodents; Africa and SE Asia[S139]), Edge Hill virus (marsupials; Australia[139]), Bouboui virus (monkeys; Africa[S139]), Jugra virus, Banzi virus (rodents; Africa[139]), Potiskum and Saboya viruses (rodent; Africa[S139]), Uganda S virus (birds; Africa[S139]), and Zika virus (monkeys; Africa & Asia)[S31, S139, S140]. Mutebi *et al*. compared the 3' noncoding region of yellow fever, Sepik, Uganda S, and Zika viruses, and placed yellow fever virus with Sepik virus as a separate clade from Uganda S virus and Zika virus[S141]. Other analyses have placed Sepik virus as an outgroup [S31, S33].

VIRUS FAMILY AND GENUS: Flaviviridae: *Flavivirus*

VIRUS TYPE: positive single-stranded RNA

ARTHROPOD VECTOR: *Aedes* and *Haemogogus* mosquitoes

NEAREST MODERN RELATED VIRUS: sylvatic yellow fever virus (nonhuman primates and other mammals)

PROXIMATE HOST: African nonhuman primates

GEOGRAPHIC ORIGIN: Old World tropics, probably Africa

Supplementary References

S1. Heymann, D. L. ed. *Control of Communicable Diseases Manual* (American Public Health Association, Washington D.C., 2004)

S2. Krauss, H. *et al.* eds. *Zoonoses: Infectious Diseases Transmissible from Animals to Humans* (ASM Press, Washington D.C., 2003)

S3. Murray, C. J. & Lopez, A. D. The incremental effect of age-weighting on YLLs, YLDs, and DALYs: a response. *Bull. World Health Organ.* **74**, 445-446 (1996).

S4. Lopez, A. D., Mathers, C. D., Ezzati, M., Jamison, D. T. & Murray, C. J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747-1757 (2006).

S5. Ramenofsky, A. F. *Vectors of Death: The Archaeology of European Contact*, 300pp. (University of New Mexico Press, Albuquerque, 1988).

S6. Verano, J. W. & Ubelaker, D. H. *Disease and Demography in the Americas*, 294pp. (Smithsonian Institution Press, Washington, 1992).

S7. Dobyns, H. F. in *Their Number Become Thinned: Native American Population Dynamics in Eastern North America*, 396pp. (University of Tennessee Press, Knoxville, 1983).

S8. Anderson, R.M. & May, R.M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, 1991).

S9. Diamond, J. *Guns, Germs, and Steel: the Fates of Human Societies* (Norton, New York, 1997).

S10. Jahrling, P. B. *et al*. Experimental infection of cynomolgus macaques with Ebola-Reston filoviruses from the 1989-1990 U.S. epizootic. *Arch. Virol. Suppl.* **11**, 115-134 (1996).

S11. Wawer, M. J. *et al*. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J. Infect. Dis.* **191**, 1403-1409 (2005)

S12. Daszak, P., Cunningham, A. A. & Hyatt, A. D. Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science* **287**, 443-449 (2000)

S13. Lovejoy, T.E. & Hannah, O., eds. *Climate Change and Biodiversity*, 418 pp. (Yale University Press, New Haven, 2005).

S14. Gao, F. *et al*. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436-441 (1999).

S15. Hahn, B. H., Shaw, G. M., De Cock, K. M. & Sharp, P. M. AIDS as a zoonosis: scientific and public health implications. *Science* **287**, 607-614 (2000).

S16. Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. *Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses*, 1259pp. (2005).

S17. Stevens, J. R., Noyes, H. A., Dover, G. A. & Gibson, W. C. The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *T. cruzi*. *Parasitology* **118 (Pt 1)**, 107-116 (1999).

S18. Aufderheide, A. C. *et al*. A 9,000-year record of Chagas' disease. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2034-2039 (2004).

S19. Piontkivska, H. & Hughes, A. L. Environmental kinetoplastid-like 18S rRNA sequences and phylogenetic relationships among Trypanosomatidae: paraphyly of the genus *Trypanosoma*. *Mol. Biochem. Parasitol.* **144**, 94-99 (2005).

S20. Lukes, J. *et al*. Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. *J. Mol. Evol.* **44**, 521-527 (1997).

S21. Stevens, J. R. & Gibson, W. C. The evolution of pathogenic trypanosomes. *Cad. Saude Publica* **15**, 673-684 (1999).

S22. Hamilton, P. B., Stevens, J. R., Gaunt, M. W., Gidley, J. & Gibson, W. C. Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *Int. J. Parasitol.* **34**, 1393-1404 (2004).

S23. Hughes, A. L. & Piontkivska, H. Phylogeny of Trypanosomatidae and Bodonidae (Kinetoplastida) based on 18S rRNA: evidence for paraphyly of *Trypanosoma* and six other genera. *Mol. Biol. Evol.* **20**, 644-652 (2003).

S24. Teixeira, A. R. *et al*. Emerging Chagas disease: trophic network and cycle of transmission of *Trypanosoma cruzi* from palm trees in the Amazon. *Emerg. Infect. Dis.* **7**, 100-112 (2001).

S25. Thompson, C. C. *et al*. Use of recA as an alternative phylogenetic marker in the family Vibrionaceae. *Int. J. Syst. Evol. Microbiol.* **54**, 919-924 (2004).

S26. Abd, H., Weintraub, A. & Sandstrom, G. Intracellular survival and replication of *Vibrio cholerae* O139 in aquatic free-living amoebae. *Environ. Microbiol.* **7**, 1003-1008 (2005).

S27. Schlater, L. K. *et al*. A non-O1 *Vibrio cholerae* isolated from a goose. *Avian Dis.* **25**, 199-201 (1981).

S28. Ogg, J. E., Ryder, R. A. & Smith, H. L.,Jr. Isolation of *Vibrio cholerae* from aquatic birds in Colorado and Utah. *Appl. Environ. Microbiol.* **55**, 95-99 (1989).

S29. Rhodes, J. B., Schweitzer, D. & Ogg, J. E. Isolation of non-O1 *Vibrio cholerae* associated with enteric disease of herbivores in western Colorado. *J. Clin. Microbiol.* **22**, 572-575 (1985).

S30. Kirn, T. J., Jude, B. A. & Taylor, R. K. A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature* **438**, 863-866 (2005).

S31. Kuno, G., Chang, G. J., Tsuchiya, K. R., Karabatsos, N. & Cropp, C. B. Phylogeny of the genus *Flavivirus*. *J. Virol.* **72**, 73-83 (1998).

S32. Wang, E. *et al*. Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. *J. Virol.* **74**, 3227-3234 (2000).

S33. Gaunt, M. W. *et al*. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J. Gen. Virol.* **82**, 1867-1876 (2001).

S34. Holmes, E. C. & Twiddy, S. S. The origin, emergence and evolutionary genetics of dengue virus. *Infect. Genet. Evol.* **3**, 19-28 (2003).

S35. Martin, J. F., Barreiro, C., Gonzalez-Lavado, E. & Barriuso, M. Ribosomal RNA and ribosomal proteins in corynebacteria. *J. Biotechnol.* **104**, 41-53 (2003).

S36. Biberstein, E. L., Knight, H. D. & Jang, S. Two biotypes of *Corynebacterium pseudotuberculosis*. *Vet. Rec.* **89**, 691-692 (1971).

S37. Olson, M. E., Goemans, I., Bolingbroke, D. & Lundberg, S. Gangrenous dermatitis caused by *Corynebacterium ulcerans* in Richardson ground squirrels. *J. Am. Vet. Med. Assoc.* **193**, 367-368 (1988).

S38. Foster, G. *et al*. Corynebacterium ulcerans in free-ranging otters. *Vet. Rec.* **150**, 524 (2002).

S39. Barrow, P. A. Corynebacterium kutscheri infection in wild voles (*Microtus agrestis*). *Br. Vet. J.* **137**, 67-70 (1981).

S40. Rich, S. M. & Ayala, F. J. Population structure and recent evolution of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6994-7001 (2000).

S41. Carter, R. & Mendis, K. N. Evolutionary and historical aspects of the burden of malaria. *Clin. Microbiol. Rev.* **15**, 564-594 (2002).

S42. Conway, D. J. Tracing the dawn of *Plasmodium falciparum* with mitochondrial genome sequences. *Trends Genet.* **19**, 671-674 (2003).

S43. Martin, M. J., Rayner, J. C., Gagneux, P., Barnwell, J. W. & Varki, A. Evolution of human-chimpanzee differences in malaria susceptibility: relationship to human genetic loss of N-glycolylneuraminic acid. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12819-12824 (2005).

S44. Simmonds, P. The origin and evolution of hepatitis viruses in humans. *J. Gen. Virol.* **82**, 693-712 (2001).

S45. Simmonds, P. & Midgley, S. Recombination in the genesis and evolution of hepatitis B virus genotypes. *J. Virol.* **79**, 15467-15476 (2005).

S46. Robertson, B. H. Viral hepatitis and primates: historical and molecular analysis of human and nonhuman primate hepatitis A, B, and the GB-related viruses. *J. Viral Hepat.* **8**, 233-242 (2001).

S47. Takahashi, K., Brotman, B., Usuda, S., Mishiro, S. & Prince, A. M. Full-genome sequence analyses of hepatitis B virus (HBV) strains recovered from chimpanzees infected in the wild: implications for an origin of HBV. *Virology* **267**, 58-64 (2000).

S48. Subbarao, K. & Shaw, M. W. Molecular aspects of avian influenza (H5N1) viruses isolated from humans. *Rev. Med. Virol.* **10**, 337-348 (2000).

S49. Lin, Y. P. *et al*. Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9654-9658 (2000).

S50. Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. The evolution of human influenza viruses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**, 1861-1870 (2001).

S51. Dantas-Torres, F. *Leishmania infantum* versus *Leishmania chagasi*: do not forget the law of priority. *Mem. Inst. Oswaldo Cruz* **101**, 117-8; discussion 118 (2006).

S52. Zemanova, E., Jirku, M., Mauricio, I. L., Miles, M. A. & Lukes, J. Genetic polymorphism within the *Leishmania donovani* complex: correlation with geographic origin. *Am. J. Trop. Med. Hyg.* **70**, 613-617 (2004).

S53. Mauricio, I. L., Gaunt, M. W., Stothard, J. R. & Miles, M. A. Genetic typing and phylogeny of the *Leishmania donovani* complex by restriction analysis of PCR amplified gp63 intergenic regions. *Parasitology* **122**, 393-403 (2001).

S54. Wilson, M. E., Jeronimo, S. M. & Pearson, R. D. Immunopathogenesis of infection with the visceralizing *Leishmania* species. *Microb. Pathog.* **38**, 147-160 (2005).

S55. Alborzi, A., Rasouli, M. & Shamsizadeh, A. *Leishmania tropica*-isolated patient with visceral leishmaniasis in southern Iran. *Am. J. Trop. Med. Hyg.* **74**, 306-307 (2006).

S56. Rodriguez-Bonfante, C., Bonfante-Garrido, R., Grimaldi, G.,Jr, Momen, H. & Cupolillo, E. Genotypically distinct *Leishmania colombiensis* isolates from Venezuela cause both cutaneous and visceral leishmaniasis in humans. *Infect. Genet. Evol.* **3**, 119-124 (2003).

S57. Akman, L. *et al*. Multi-site DNA polymorphism analyses of *Leishmania* isolates define their genotypes predicting clinical epidemiology of leishmaniasis in a specific region. *J. Eukaryot. Microbiol.* **47**, 545-554 (2000).

S58. Ibrahim, M. E. & Barker, D. C. The origin and evolution of the *Leishmania donovani* complex as inferred from a mitochondrial cytochrome oxidase II gene sequence. *Infect. Genet. Evol.* **1**, 61-68 (2001).

S59. Pratlong, F. *et al*. Sudan: the possible original focus of visceral leishmaniasis. *Parasitology* **122**, 599-605 (2001).

S60. Lainson, R. & Rangel, E. F. *Lutzomyia longipalpis* and the eco-epidemiology of American visceral leishmaniasis, with particular reference to Brazil: a review. *Mem. Inst. Oswaldo Cruz* **100**, 811-827 (2005).

S61. Rotureau, B. Ecology of the *Leishmania* species in the Guianan ecoregion complex. *Am. J. Trop. Med. Hyg.* **74**, 81-96 (2006).

S62. Anderson, E. C. Morbillivirus infections in wildlife (in relation to their population biology and disease control in domestic animals). *Vet. Microbiol.* **44**, 319-332 (1995).

S63. Kock, R. A. *et al*. Rinderpest epidemic in wild ruminants in Kenya 1993-97. *Vet. Rec.* **145**, 275-283 (1999).

S64. Westover, K. M. & Hughes, A. L. Molecular evolution of viral fusion and matrix protein genes and phylogenetic relationships among the Paramyxoviridae. *Mol. Phylogenet. Evol.* **21**, 128-134 (2001).

S65. Seal, B. S., Crawford, J. M., Sellers, H. S., Locke, D. P. & King, D. J. Nucleotide sequence analysis of the Newcastle disease virus nucleocapsid protein gene and phylogenetic relationships among the Paramyxoviridae. *Virus Res.* **83**, 119-129 (2002).

S66. Seal, B. S. Nucleotide and predicted amino acid sequence analysis of the fusion protein and hemagglutinin-neuraminidase protein genes among Newcastle disease virus isolates. Phylogenetic relationships among the Paramyxovirinae based on attachment glycoprotein sequences. *Funct. Integr. Genomics* **4**, 246-257 (2004).

S67. Shihmanter, E., Panshin, A. & Lipkind, M. Nucleotide sequence of the matrix protein gene of avian paramyxovirus, serotype 3b: evidence on another member of the suggested new genus of the subfamily Paramyxovirinae. *Comp. Immunol. Microbiol. Infect. Dis.* **28**, 37-51 (2005).

S68. Rima, B. K. & Duprex, W. P. Morbilliviruses and human disease. *J. Pathol.* **208**, 199-214 (2006).

S69. Chang, P. C. *et al*. Complete nucleotide sequence of avian paramyxovirus type 6 isolated from ducks. *J. Gen. Virol.* **82**, 2157-2168 (2001).

S70. Kawano, M. *et al*. Sequence determination of the hemagglutinin-neuraminidase (HN) gene of human parainfluenza type 2 virus and the construction of a phylogenetic tree for HN proteins of all the paramyxoviruses that are infectious to humans. *Virology* **174**, 308-313 (1990).

S71. Chua, K. B., Wang, L. F., Lam, S. K. & Eaton, B. T. Full length genome sequence of Tioman virus, a novel paramyxovirus in the genus *Rubulavirus* isolated from fruit bats in Malaysia. *Arch. Virol.* **147**, 1323-1348 (2002).

S72. Berg, M. *et al*. Analysis of the fusion protein gene of the porcine rubulavirus LPMV: comparative analysis of paramyxovirus F proteins. *Virus Genes* **14**, 55-61 (1997).

S73. Muhlemann, K. The molecular epidemiology of mumps virus. *Infect. Genet. Evol.* **4**, 215-219 (2004).

S74. Bjornstad, O. N. & Harvill, E. T. Evolution and emergence of *Bordetella* in humans. *Trends Microbiol.* **13**, 355-359 (2005).

S75. van der Zee, A., Mooi, F., Van Embden, J. & Musser, J. Molecular evolution and host adaptation of *Bordetella* spp.: phylogenetic analysis using multilocus enzyme electrophoresis and typing with three insertion sequences. *J. Bacteriol.* **179**, 6609-6617 (1997).

S76. Kim, W. *et al*. Comparison of 16S rDNA analysis and rep-PCR genomic fingerprinting for molecular identification of *Yersinia pseudotuberculosis*. *Antonie Van Leeuwenhoek* **83**, 125-133 (2003).

S77. Achtman, M. *et al*. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 14043-14048 (1999).

S78. Dykhuizen, D. E. Yersinia pestis: an instant species? *Trends Microbiol.* **8**, 296-298 (2000).

S79. Thomson, N. R. & Parkhill, J. in *Yersinia: molecular and cellular biology* (eds Carniel, E. & Hinnebusch, B. J.) 1-11 (Horizon Bioscience, Wymondham, 2004).

S80. Achtman, M. *et al*. Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17837-17842 (2004).

S81. Poland, J. D. & Dennis, D. T. in *Bacterial infections of humans :epidemiology and control* (eds Evans, A. S. & Brachman, P. S.) 545-558 (Plenum Medical Book Co., New York, 1998).

S82. Drancourt, M. *et al*. Genotyping, Orientalis-like *Yersinia pestis*, and plague pandemics. *Emerg. Infect. Dis.* **10**, 1585-1592 (2004).

S83. Picard, B. Molecular epidemiology of large bacterial endemics in Sub-Saharan Africa. *Bull. Soc. Pathol. Exot.* **93**, 219-223 (2000).

S84. Ito, H., Sugiyama, M., Masubuchi, K., Mori, Y. & Minamoto, N. Complete nucleotide sequence of a group A avian rotavirus genome and a comparison with its counterparts of mammalian rotaviruses. *Virus Res.* **75**, 123-138 (2001).

S85. De Leener, K. *et al*. Human infection with a P[14], G3 lapine rotavirus. *Virology* **325**, 11-17 (2004).

S86. Matthijnssens, J. *et al*. Full genomic analysis of human rotavirus strain B4106 and lapine rotavirus strain 30/96 provides evidence for interspecies transmission. *J. Virol.* **80**, 3801-3810 (2006).

S87. Rahman, M. *et al*. Detection and characterization of human group C rotaviruses in Bangladesh. *J. Clin. Microbiol.* **43**, 4460-4465 (2005).

S88. van der Heide, R. *et al*. Molecular characterizations of human and animal group a rotaviruses in the Netherlands. *J. Clin. Microbiol.* **43**, 669-675 (2005).

S89. Schlesinger, S. & Schlesinger, M. J. in *Fields' virology* (eds Fields, B. N., Knipe, D. M., Howley, P. M. & Griffin, D. E.) 895-916 (Lippincott Williams & Wilkins, Philadelphia, 2001).

S90. Frey, T. K. Molecular biology of rubella virus. *Adv. Virus Res.* **44**, 69-160 (1994).

S91. MacLeod, A., Welburn, S., Maudlin, I., Turner, C. M. & Tait, A. Evidence for multiple origins of human infectivity in *Trypanosoma brucei* revealed by minisatellite variant repeat mapping. *J. Mol. Evol.* **52**, 290-301 (2001).

S92. Herder, S., Simo, G., Nkinin, S. & Njiokou, F. Identification of trypanosomes in wild animals from southern Cameroon using the polymerase chain reaction (PCR). *Parasite* **9**, 345-349 (2002).

S93. Njiokou, F., Simo, G., Nkinin, S. W., Laveissiere, C. & Herder, S. Infection rate of *Trypanosoma brucei* s.l., *T. vivax*, *T. congolense* "forest type", and *T. simiae* in small wild vertebrates in south Cameroon. *Acta Trop.* **92**, 139-146 (2004).

S94. Njiokou, F. *et al*. Wild fauna as a probable animal reservoir for *Trypanosoma brucei gambiense* in Cameroon. *Infect. Genet. Evol.* **6**, 147-153 (2006).

S95. Fevre, E. M. *et al*. The origins of a new *Trypanosoma brucei rhodesiense* sleeping sickness outbreak in eastern Uganda. *Lancet* **358**, 625-628 (2001).

S96. Hutchinson, O. C., Fevre, E. M., Carrington, M. & Welburn, S. C. Lessons learned from the emergence of a new *Trypanosoma brucei rhodesiense* sleeping sickness focus in Uganda. *Lancet Infect. Dis.* **3**, 42-45 (2003).

S97. Fevre, E. M. *et al*. A burgeoning epidemic of sleeping sickness in Uganda. *Lancet* **366**, 745-747 (2005).

S98. Fevre, E. M. *et al*. Central point sampling from cattle in livestock markets in areas of human sleeping sickness. *Acta Trop.* **97**, 229-232 (2006).

S99. Fevre, E. M., Picozzi, K., Jannin, J., Welburn, S. C. & Maudlin, I. Human African trypanosomiasis: epidemiology and control. *Adv. Parasitol.* **61**, 167-221 (2006).

S100. Hughes, A. L. & Friedman, R. Poxvirus genome evolution by gene gain and loss. *Mol. Phylogenet. Evol.* **35**, 186-195 (2005).

S101. Gubser, C. & Smith, G. L. The sequence of camelpox virus shows it is most closely related to variola virus, the cause of smallpox. *J. Gen. Virol.* **83**, 855-872 (2002).

S102. Gubser, C., Hue, S., Kellam, P. & Smith, G. L. Poxvirus genomes: a phylogenetic analysis. *J. Gen. Virol.* **85**, 105-117 (2004).

S103. Afonso, C. L. *et al*. The genome of camelpox virus. *Virology* **295**, 1-9 (2002).

S104. Meyer, H., Neubauer, H. & Pfeffer, M. Amplification of 'variola virus-specific' sequences in German cowpox virus isolates. *J. Vet. Med. B Infect. Dis. Vet. Public Health* **49**, 17-19 (2002).

S105. Fenner, F. Smallpox: emergence, global spread, and eradication. *Hist. Philos. Life. Sci.* **15**, 397-420 (1993).

S106. Fenner, F. Adventures with poxviruses of vertebrates. *FEMS Microbiol. Rev.* **24**, 123-133 (2000).

S107. Esposito, J. J. *et al*. Genome sequence diversity and clues to the evolution of variola (smallpox) virus. *Science* **313**, 807-812 (2006).

S108. Rothschild, B. M. History of syphilis. *Clin. Infect. Dis.* **40**, 1454-1463 (2005).

S109. von Hunnius, T. E., Roberts, C. A., Boylston, A. & Saunders, S. R. Histological identification of syphilis in pre-Columbian England. *Am. J. Phys. Anthropol.* **129**, 559-566 (2006).

S110. Baylet, R., Thivolet, J., Sepetjian, M. & Bert, J. Seroepidemiological studies on primate treponematosis in Senegal. *Bull. Soc. Pathol. Exot. Filiales* **64**, 836-841 (1971).

S111. Baylet, R., Thivolet, J., Sepetjian, M., Nouhouay, Y. & Baylet, M. Natural open treponematosis in the Papio papio baboon in Casamance. *Bull. Soc. Pathol. Exot. Filiales* **64**, 842-846 (1971).

S112. Fribourg-Blanc, A. & Mollaret, H. H. Natural treponematosis of the African primate. *Primates Med.* **3**, 113-121 (1969).

S113. Centurion-Lara, A. *et al*. Conservation of the 15-kilodalton lipoprotein among *Treponema pallidum* subspecies and strains and other pathogenic treponemes: genetic and antigenic analyses. *Infect. Immun.* **65**, 1440-1444 (1997).

S114. Gray, R. R. *et al*. Molecular Evolution of the tprC, D, I, K, G, and J Genes in the Pathogenic Genus *Treponema*. *Mol. Biol. Evol.* (2006).

S115. Quddus, A., Luby, S., Rahbar, M. & Pervaiz, Y. Neonatal tetanus: mortality rate and risk factors in Loralai District, Pakistan. *Int. J. Epidemiol.* **31**, 648-653 (2002).

S116. Parashar, U. D., Bennett, J. V., Boring, J. R. & Hlady, W. G. Topical antimicrobials applied to the umbilical cord stump: a new intervention against neonatal tetanus. *Int. J. Epidemiol.* **27**, 904-908 (1998).

S117. Bennett, J., Ma, C., Traverso, H., Agha, S. B. & Boring, J. Neonatal tetanus associated with topical umbilical ghee: covert role of cow dung. *Int. J. Epidemiol.* **28**, 1172-1175 (1999).

S118. Bennett, J. *et al*. Circumcision and neonatal tetanus: disclosure of risk and its reduction by topical antibiotics. *Int. J. Epidemiol.* **28**, 263-266 (1999).

S119. Wilde, E., Collins, M. D. & Hippe, H. *Clostridium pascui* sp. nov., a new glutamate-fermenting sporeformer from a pasture in Pakistan. *Int. J. Syst. Bacteriol.* **47**, 164-170 (1997).

S120. Devulder, G., Perouse de Montclos, M. & Flandrois, J. P. A multigene approach to phylogenetic analysis using the genus *Mycobacterium* as a model. *Int. J. Syst. Evol. Microbiol.* **55**, 293-302 (2005).

S121. Kim, H. *et al*. Differentiation of *Mycobacterium* species by analysis of the heat-shock protein 65 gene (hsp65). *Int. J. Syst. Evol. Microbiol.* **55**, 1649-1656 (2005).

S122. Mostowy, S. *et al*. Revisiting the evolution of *Mycobacterium bovis*. *J. Bacteriol.* **187**, 6386-6395 (2005).

S123. Krauss, H. *Zoonoses :infectious diseases transmissible from animals to humans* 456pp. (ASM Press, Washington, D.C., 2003).

S124. Mostowy, S. & Behr, M. A. The origin and evolution of *Mycobacterium tuberculosis*. *Clin. Chest Med.* **26**, 207-16, v-vi (2005).

S125. Paradis, S. *et al*. Phylogeny of the Enterobacteriaceae based on genes encoding elongation factor Tu and F-ATPase beta-subunit. *Int. J. Syst. Evol. Microbiol.* **55**, 2013-2025 (2005).

S126. Selander, R. K. *et al*. Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. *Infect. Immun.* **58**, 2262-2275 (1990).

S127. Frankel, G., Newton, S. M., Schoolnik, G. K. & Stocker, B. A. Intragenic recombination in a flagellin gene: characterization of the H1-j gene of *Salmonella typhi*. *EMBO J.* **8**, 3149-3152 (1989).

S128. Roumagnac, P. *et al*. Evolutionary history of *Salmonella typhi*. *Science*. **314**, 1301-4 (2006).

S129. Andersson, S. G., Stothard, D. R., Fuerst, P. & Kurland, C. G. Molecular phylogeny and rearrangement of rRNA genes in *Rickettsia* species. *Mol. Biol. Evol.* **16**, 987-995 (1999).

S130. Lee, J. H. *et al*. Differentiation of rickettsiae by groEL gene analysis. *J. Clin. Microbiol.* **41**, 2952-2960 (2003).

S131. Zhu, Y., Fournier, P. E., Ogata, H. & Raoult, D. Multispacer typing of *Rickettsia prowazekii* enabling epidemiological studies of epidemic typhus. *J. Clin. Microbiol.* **43**, 4708-4712 (2005).

S132. Azad, A. F. & Beard, C. B. Rickettsial pathogens and their arthropod vectors. *Emerg. Infect. Dis.* **4**, 179-186 (1998).

S133. Jongwutiwes, S. *et al*. Mitochondrial genome sequences support ancient population expansion in *Plasmodium vivax*. *Mol. Biol. Evol.* **22**, 1733-1739 (2005).

S134. Escalante, A. A. *et al*. A monkey's tale: the origin of *Plasmodium vivax* as a human malaria parasite. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1980-1985 (2005).

S135. Mu, J. *et al*. Host switch leads to emergence of *Plasmodium vivax* malaria in humans. *Mol. Biol. Evol.* **22**, 1686-1693 (2005).

S136. Shope, R. E. & Meegan, J. M. in *Viral infections of humans: epidemiology and control* (eds Evans, A. S. & Kaslow, R. A.) 151-183 (Plenum Medical Book Co., New York, 1997).

S137. de Thoisy, B., Dussart, P. & Kazanji, M. Wild terrestrial rainforest mammals as potential reservoirs for flaviviruses (yellow fever, dengue 2 and St Louis encephalitis viruses) in French Guiana. *Trans. R. Soc. Trop. Med. Hyg.* **98**, 409-412 (2004).

S138. Cook, S. & Holmes, E. C. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. *Arch. Virol.* **151**, 309-325 (2006).

S139. Burke, D. S. & Monath, T. P. in *Fields' Virology* (eds Fields, B. N., Knipe, D. M., Howley, P. M. & Griffin, D. E.) 1043-1125 (Lippincott Williams & Wilkins, Philadelphia, 2001).

S140. Cook, S. & Holmes, E. C. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. *Arch. Virol.* **151**, 309-325 (2006).

S141. Mutebi, J. P. *et al*. Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the *Flavivirus* genus based on the 3' noncoding region. *J. Virol.* **78**, 9652-9665 (2004).