Disrupted Human–Pathogen Co-Evolution: A Model for Disease

Nuri Kodaman
Dartmouth College

Rafal S. Sobota
Dartmouth College

Robertino Mera
Vanderbilt University

Barbara G. Schneider
Vanderbilt University

Scott M. Williams
Dartmouth College

Follow this and additional works at: https://digitalcommons.dartmouth.edu/facoa

Part of the Diseases Commons, Infectious Disease Commons, and the Medical Genetics Commons

Recommended Citation
Kodaman, Nuri; Sobota, Rafal S.; Mera, Robertino; Schneider, Barbara G.; and Williams, Scott M., "Disrupted Human–Pathogen Co-Evolution: A Model for Disease" (2014). Open Dartmouth: Faculty Open Access Articles. 1306.
https://digitalcommons.dartmouth.edu/facoa/1306

This Article is brought to you for free and open access by Dartmouth Digital Commons. It has been accepted for inclusion in Open Dartmouth: Faculty Open Access Articles by an authorized administrator of Dartmouth Digital Commons. For more information, please contact dartmouthdigitalcommons@groups.dartmouth.edu.
Disrupted human–pathogen co-evolution: a model for disease

Nuri Kodaman1,2, Rafal S. Sobota1,2, Robertino Mera3, Barbara G. Schneider3 and Scott M. Williams1*  

1 Department of Genetics, Geisel School of Medicine, Dartmouth College, Hanover, NH, USA  
2 Department of Molecular Physiology and Biophysics, Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN, USA  
3 Division of Gastroenterology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA  

INTRODUCTION

Human response to infectious agents is known to be highly heritable, but identifying the genetic variants responsible for differences in disease susceptibility has proven difficult. Pathogenic variation has, in some cases, become a better predictor of disease outcome, but it too does not sufficiently predict whether a given individual or class of individuals will present with disease. Thus far, genetic epidemiological studies of infectious disease have typically sought to explain the inter-individual variation in disease phenotypes by assessing genetic factors in humans or pathogens alone, under the implicit assumption that these factors have effects that are essentially independent of each other. Here, we argue that genome-by-genome interactions between host and pathogen are likely to play a major role in infectious disease etiology, and as such, should be incorporated into genetic epidemiological models. In short, insofar as host and pathogen jointly determine disease phenotypes, no genetic variant in either should be considered harmful without taking the context of the other into account.

The term “interaction” has two related but distinct meanings in the context of infectious disease, one molecular, and one statistical. Here we refer mainly to the statistical meaning of the term. At the individual level, all aspects of pathogenesis involve molecular interactions of varying importance, e.g., between a pathogenic epitope and a host receptor. Such interactions can be detected statistically, however, only when multiple variants exist in a population and when specific pairings lead to different effects. In some cases, pathogenic variants may function independently of host variation, and vice versa. However, because many pathogens have co-existed with their human hosts for millennia and have likely co-evolved with them, we argue here that statistical interactions, where appropriately sought, will often be found, with profound biomedical implications.

A major goal in infectious disease research is to identify the human and pathogenic genetic variants that explain differences in microbial pathogenesis. However, neither pathogenic strain nor human genetic variation in isolation has proven adequate to explain the heterogeneity of disease pathology. We suggest that disrupted co-evolution between a pathogen and its human host can explain variation in disease outcomes, and that genome-by-genome interactions should therefore be incorporated into genetic models of disease caused by infectious agents. Genetic epidemiological studies that fail to take both the pathogen and host into account can lead to false and misleading conclusions about disease etiology. We discuss our model in the context of three pathogens, Helicobacter pylori, Mycobacterium tuberculosis and human papillomavirus, and generalize the conditions under which it may be applicable.

Recent advances in genomics have provided both the impetus and the means to evaluate human–pathogen co-evolutionary hypotheses directly. Whole-genome sequencing of many pathogenic species has substantially improved the resolution with which we classify strains, and facilitated the detection of potentially virulent genetic variants. A clearer picture of microbial evolution has also emerged, marked by selective mechanisms such as rapid gene gain/loss and horizontal gene transfer (Pallen and Wren, 2007). Overlaid human genetic variation onto this emerging evolutionary picture of microbial diversity offers the potential to make the pathogenic process more transparent.

The past few decades have also seen an explosion in studies seeking to identify human susceptibility loci for infectious diseases (Rowell et al., 2012). Candidate gene and family based linkage studies have identified several common polymorphisms with clinical significance at the population level, such as the CCR5 deletion that protects against HIV (Samson et al., 1996; Picard et al., 2006; Casanova and Abel, 2007). However, most human susceptibility is in fact polygenic, with individual polymorphisms conferring small marginal effects (Hill, 2001). Where infectious disease phenotypes deviate from the “one susceptibility locus – one infection” model, elucidating the genetic architecture underlying inter-individual variation has proven elusive.

While genome-wide association studies (GWAS) may be better designed to accommodate multifactorial phenotypes, those performed thus far on infectious diseases have typically been less informative than GWAS performed on complex non-communicable diseases (Jallow et al., 2009; Hill, 2012; Ko and Urban, 2013). A major challenge facing the GWAS of infectious disease has been the recruitment of a sufficient number of cases and matched controls to achieve adequate statistical power (Hill, 2012; Ko and Urban, 2013). Another potential drawback, and the...
one that concerns us here, is the fact that many infectious disease phenotypes depend on complex interactions between host and pathogen genomes. In such cases, the pooling together of human samples infected with even subtly different pathogenic strains can obscure genetic associations (Hill, 2012; Ko and Urban, 2013). A problem common to all GWAS is that the statistical effect sizes of biologically meaningful polymorphisms are often too small to pass significance thresholds after correction for multiple testing. This problem is exacerbated, however, when human polymorphisms (or networks of polymorphisms) (Wilfert and Schmid-Hempel, 2008) confer variable, or even opposite effects in the context of different pathogenic strains within the same study cohort. In this regard, it is perhaps telling that the most successful GWAS performed on infectious disease susceptibility to date have been on leprosy; the signal-to-noise ratios in these association studies may be higher because Mycobacterium leprae exhibits substantially less genetic heterogeneity than many other pathogens (Monot et al., 2009; Hill, 2012).

There is in fact strong empirical and theoretical justification for the hypothesis that the effects of susceptibility and virulence alleles in the respective gene pools of humans and pathogens are often contingent upon each other. The evolution of virulence is a dynamic process, easily perturbed by extrinsic variables over space and time, and therefore unlikely to follow the same trajectory in every population. For example, a spike in the density of hosts available for transmission can select for increased virulence, by reducing the cost of lethal harm (Anderson and May, 1982). If a pathogen is transmitted vertically (parent to child), the genetic factors that affect pathogenicity are “co-inherited” by host and pathogen, often promoting commensalism (Frank, 1996; Messenger et al., 1999). Even in these cases, the adventitious introduction of a microbial competitor can induce a commensal species to evolve a defensive toxin that harms the host, if only incidentally (Blaser and Kirschner, 2007; Frank and Schmid-Hempel, 2008). The evolution of defenses against pathogenic harm must also navigate fitness tradeoffs that vary with population, including tradeoffs pertaining to the correlated nature of complex traits (Lambrechts et al., 2006). As pathogens evolve rapidly, exerting strong selective pressures on different human populations, host phenotypes will respond in the ad hoc manner typical of evolution, limited by the available genetic variation at hand (Jacob, 1977). Whether the result is a steady-state equilibrium due to a perpetual “arms race” or a commensal detente, the same genes and pathways are unlikely to be involved in every population. As a consequence, when humans and pathogens migrate to new environments or admix, the ensuing disruption of co-evolutionary equilibria and loss of complementarity between host and pathogen genotypes may yield unpredictable and potentially deleterious biomedical consequences.

Our emphasis on the significance of mismatched traits is consistent with the genetic mosaic theory of co-evolution, which aims to account for why virtually all co-evolutionary interactions observed in natural populations show spatial variation in outcomes (Thompson et al., 2002; Thompson, 2014). The theory posits that co-evolution occurs in the context of geographically distinct “selection mosaics,” each characterized by a unique genetic and environmental profile, where environmental variables can include both biotic and abiotic factors. Every selection mosaic progresses toward its own co-evolutionary equilibrium, while gene flow between selection mosaics ensures that patterns of maladaptation will be common and detectable where properly studied (Thompson et al., 2002; Ridenhour and Nuismer, 2007).

Despite the likely etiological importance of human–pathogen co-evolution, attempts at empirical confirmation have been rare. Indeed, “proof” of co-evolution poses a formidable challenge, requiring a demonstration of increased reproductive fitness in each species driven by reciprocal changes in two genomes over time (Woolhouse et al., 2002). Although these criteria have been met in laboratory studies and in some natural populations (Lenski and Levin, 1985; Little, 2002; Little et al., 2006), a similarly rigorous assessment of human–pathogen co-evolution must accommodate long generation times and the genetic and phenotypic complexity of the human traits under selection. Nonetheless, substantial phenomenological evidence consistent with human–pathogen co-evolution now exists, including evidence of spatial patterns of parallel genetic variation between species, and of correlated functional changes at the molecular level (Kraaijeveld et al., 1998; Lively and Dybdahl, 2000; Funk et al., 2000; Woolhouse et al., 2002). The collection of high-density genomic data in paired human–pathogen samples and improvements in phenotypic data, as well as advances in pathogen genomics, should soon enable more explicit tests of the concept.

Our aim here is to summarize the growing body of evidence in favor of the hypothesis that genetic interactions driven by host and pathogen co-evolution can have significant implications for genetic epidemiological studies and biomedicine. While this is not a novel hypothesis, it remains understudied. We also underscore how recent advances in genomic technology provide new opportunities to test for genome-by-genome interactions, and offer suggestions on how to incorporate them into more accurate genetic models of disease.

**HELCOBACTER PYLORI**

Studies of Helicobacter pylori provide perhaps the best evidence in favor of human–pathogen co-evolution, and distinctly illustrate the power of the modern genetic toolkit to investigate it. *H. pylori* chronically infects the gastric epithelia of half the world’s population, causing peptic ulcers in 10–20% of those infected, and distal gastric carcinoma in ~1% (Peek and Blaser, 2002; Jemal et al., 2011). The majority of individuals infected, however, suffer only from superficial gastritis in adulthood, while likely gaining protection against diseases such as esophageal cancer and reflux esophagitis, and more controversially, childhood asthma and diarrhea (Rothenbacher et al., 2000; Vaizi et al., 2000; Blaser et al., 2008). That *H. pylori* should have a largely innocuous and potentially symbiotic relationship with its host follows from co-evolutionary theory, based on its vertical mode of transmission, its long-term colonization of a single host, and its ~50,000 year association with *Homo sapiens* (Rothenbacher et al., 2002; Moodley et al., 2012). Why a fraction of individuals develop life-threatening clinical disease, on the other hand, requires explanation, with one possibility being the disruption of long-standing co-evolutionary relationships.
Although *H. pylori*-mediated diseases often advance to the clinical stage in late adulthood, their onset typically occurs during reproductive years (Correa et al., 1976; Susser and Stein, 2002). Importantly, a disease need not have an especially large selection coefficient to shape allele frequency distributions in populations, especially over thousands of years (Ewald and Cochran, 2000). In fact, the historical fitness load of peptic ulcers, obtained by multiplying prevalence by selection coefficient, has been estimated to be similar to those for infectious diseases such as meningitis and rubella (Cochran et al., 2000). Also consistent with co-evolutionary theory is the fact that *H. pylori*-mediated gastric diseases occur disproportionately in men (Susser and Stein, 2002; Engel et al., 2003); *H. pylori* is usually, but not necessarily, transmitted by the mother, such that female fitness has likely exerted a stronger constraint against *H. pylori* virulence.

Some *H. pylori* virulence factors appear to increase the risk of serious clinical outcome regardless of host genotype. The *cag* pathogenicity island, present in some strains, encodes a type IV secretion system, and *VacA* encodes a pore-forming cytotoxin. Both have been implicated as carcinogenic risk factors, though neither is a necessary nor sufficient one (Wroblewski et al., 2010). Other virulence factors released by *H. pylori* include urease, which facilitates neutralization of the otherwise forbidding acidity of the gastric mucosa; *NAP*, which enables iron uptake; and arginase, which helps *H. pylori* subvert host macrophages. These, like most *H. pylori* virulence factors, operate to create a basal inflammatory state without generating an excessive immune response. Serious clinical disease reflects a disturbance of this balance (Baldari et al., 2005; Blaser and Kirschner, 2007; Salama et al., 2013). The maintenance of this balance also depends partly on human genetic factors (Lichtenstein et al., 2000; Chiba et al., 2006; Mayerle et al., 2013a). Candidate gene studies on *H. pylori*-mediated diseases have implicated several gene polymorphisms that appear to affect risk, most notably in the interleukin-1 (IL-1) family of cytokines (Schneider et al., 2008). Recently, two GWAS assessing susceptibility to gastric cancer and *H. pylori* infection identified SNPs with odds ratios ranging from 1.3 to 1.4, mostly of uncertain biological function (Shi et al., 2011; El-Omar, 2013; Mayerle et al., 2013b, Table 1). These polymorphisms account for only a small proportion of the estimated heritability of disease phenotypes.

Studies of human or *H. pylori* genetics in isolation have generally failed to explain why populations with similar rates of *H. pylori* infection exhibit strikingly different susceptibilities to gastric cancer. For example, in many African and South Asian countries, the low incidences of gastric cancer in the presence of almost universal rates of *H. pylori* infection remain a source of much speculation, and have been referred to collectively as the "African enigma" and the "Asian enigma" (Holcombe, 1992; Campbell et al., 2001; Ghoshal et al., 2007). In Latin America, where *H. pylori* strains native to Amerindian populations have been largely displaced by European strains (Correa and Piazuelo, 2012), the predominantly Amerindian populations living at high altitudes suffer disproportionately from gastric cancer relative to other populations with similar infection rates (de Sablet et al., 2011; Torres et al., 2013). These and other points of evidence raise the possibility that the pathogenicity of a given *H. pylori* strain may vary with human genomic variation, and that some individuals may be better adapted to their infecting strains than others.

Modern genomic techniques have made the assessment of such hypotheses feasible. Over the past two decades, a comprehensive phylogeography of *H. pylori* has been constructed using multilocus sequence typing (MLST), a procedure by which polymorphisms in fragments from housekeeping genes are used to characterize bacterial isolates (Maiden et al., 1998). Analyses of samples from around the world have revealed a strong concordance between *H. pylori* phylogenetic clusters and the geographical locations from which they are derived (Falush et al., 2003; Moodley and Linz, 2009; Moodley et al., 2009). Ancestral *H. pylori* sequences inferred using MLST data also correspond to geographically defined human populations (Falush et al., 2003; Moodley et al., 2012). The typical modern *H. pylori* chromosome is now understood to be an amalgam of fragments from multiple ancestral sequences, a consequence of *H. pylori*’s high recombining tendency (Suerbaum et al., 1998; Falush et al., 2003). The genome of an *H. pylori* isolate can thus be quantitatively resolved into ancestral proportions, which correlate with proportions of human ancestry in admixed populations (Kodaman et al., 2014). In some cases, the ancestries of *H. pylori* isolates outperform human mitochondria in differentiating ethnic groups (Wirth et al., 2004).

These shared patterns of ancestry are unlikely to have arisen merely from parallel divergence due to founder effects or neutral drift. Certainly, the well-documented evolvability of functional loci within *H. pylori* strains, even within single individuals over a 6 year span, argues for the importance of adaptive microevolution (Israel et al., 2001; Dorer et al., 2009). Furthermore, at least 25% of known genes, including genes involved in mucosal adherence and the evasion of host immunity, are absent in some *H. pylori* strains isolated from different ethnic groups (Salama et al., 2000; Gressmann et al., 2005). In at least one case, variants of an *H. pylori* gene (*babA2*) encode adhesion proteins that exhibit host-specific effects, a hallmark of co-evolution. BabA binds to blood group antigens, triggering the release of proinflammatory cytokines. Notably, Amerindians, who almost all carry blood group O, harbor strains with a BabA variant that has up to a 1500-fold greater binding affinity to blood group O (Aspholm-Hurtig et al., 2004).

If we conclude from these patterns of genetic covariation that co-evolution between humans and *H. pylori* has occurred and that it has promoted commensalism, then we may ask whether individuals who develop serious clinical disease have inherited mutually ill-adapted sets of host and pathogen alleles. Under this hypothesis, we should expect to find significant interactions between specific pairs of host and pathogen loci in disease models. Toward this end, candidate pairs of loci can be tested based on biochemical evidence of protein–protein interactions, such as those between the heparin/heparin-B and the Lewis(b) antigen, its epithelial receptor (Backstrom et al., 2004). However, the effect size of any single two-locus interaction may be relatively small, as gastric disease etiology is phenotypically heterogeneous, and likely to be influenced by a large number of human and *H. pylori* genes (El-Omar, 2013). Thus, characterizing the relevant loci in a biologically meaningful way will ultimately require a systems biological approach.
We recently took a broad-based view to assess the impact of human – *H. pylori* co-evolution on gastric disease, using ancestry estimates from both humans and their *H. pylori* isolates in the absence of knowledge of specific interacting loci (Kodaman et al., 2014). Our study participants were recruited from two Colombian populations with highly different rates of gastric cancer, despite a nearly universal prevalence of *H. pylori* infection in both. We found that the low-risk human, coastal population was of admixed African, European, and Amerindian ancestry, whereas the high-risk, Andean population was mainly of Amerindian ancestry, with a minority of European ancestry. Severity of gastric disease correlated with the proportion of African *H. pylori* ancestry in patients with primarily Amerindian ancestry. On the other hand, patients with a large proportion of African human ancestry infected by African *H. pylori* strains had the best prognoses, consistent with ancestral coadaptation, and likely pertinent to the “African enigma.” The interaction between Amerindian human ancestry and African *H. pylori* ancestry accounted for the difference in disease risk between mountain and coastal populations, whereas even the well-known virulence factor, CagA, did not. These findings are thus consistent with the idea that neither human nor *H. pylori* genetic variation confers susceptibility or virulence per se, but only in context (Figure 1).

These findings also bring to light how understanding co-evolutionary interactions can inform and improve public health measures. It has been suggested that because *H. pylori* dominates the gastric microbiome in infected persons and has been shown to confer some beneficial effects, large-scale antibiotic eradication programs may not be warranted (Bik et al., 2006; Hung and Wong, 2009). Simply estimating ancestry from human samples and *H. pylori* isolates may help to identify individuals at greatest risk for gastric cancer, for whom antibiotic treatment may be most appropriate.

**MYCOBACTERIUM TUBERCULOSIS COMPLEX**

Another interesting candidate to study from a co-evolutionary perspective is *Mycobacterium tuberculosis* (*Mtb* and closely related
species, believed to have co-existed with anatomically modern humans for \( \sim 70,000 \) years (Comas et al., 2013). Since the advent of antibiotics, tuberculosis (TB) has ceased to be as common a cause of human mortality as it once was, but it remains among the most deadly infectious diseases worldwide, with immunocompromised individuals at particularly high risk (Dye and Williams, 2010; Fenner et al., 2013). As with \textit{H. pylori}, the majority of Mtb infections do not develop into clinical disease: 90% of cases are asymptomatic with only latent infection. However, 10% of individuals with latent infections develop TB over their lifetime, for mostly unknown reasons (Barry et al., 2009).

In contrast to \textit{H. pylori}, Mtb is transmitted horizontally, and must cause active disease to be transmitted (e.g., via coughing or sneezing). Because Mtb transmission increases with virulence, evolutionary theory predicts that strong selective pressures should favor increased virulence until the number of transmissions per infected host reaches a fitness-reducing limit (Knolle, 1989; Frank and Schmid-Hempel, 2008). Such a limit necessarily depends on population-specific parameters, of which host density is probably the most important (Comas et al., 2013). Thus, the limited pathogenicity and chronicity of Mtb likely reflect its historical adaptation to isolated, low-density human populations. These historical conditions remain relevant in part because Mtb reproduces clonally and without lateral gene transfer; evolution only through point mutations and irreversible gene deletions limits a pathogen’s ability to shift virulence strategies rapidly in response to changing population parameters (Achtman, 2008; Galagan, 2014).

Before advances in genotyping technology improved strain classification, the apparent genetic homogeneity of Mtb led investigators to believe that variation in disease outcome depended primarily on environmental and human genetic factors (Galagan, 2014). Twin and adoption studies provided compelling evidence for the involvement of human genetic variation as a risk modifier (Comstock, 1978). The most recent analyses have calculated the heritable component of Mtb-related immune response phenotypes to range from 30 to 71% (Moller and Hoal, 2010). These findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have motivated a large number of linkage and candidate gene association studies seeking to identify relevant susceptibility findings have moti...
the infectivity of *M. africanum*, but not other strains, in Ghanaian populations (Herb et al., 2008; Thye et al., 2011).

Despite being an ancient strain with ample opportunity to spread beyond West Africa, *M. africanum* has not done so, possibly indicating host-specific adaptation (de Jong et al., 2010; Gagneux, 2012). Other MtB lineages also appear to associate preferentially with particular human populations, though not as exclusively. A study of ethnically diverse, US-born patients in San Francisco showed that such preferential associations with MtB lineages persisted even in a cosmopolitan setting (Gagneux et al., 2006). Interestingly, when TB transmission in non-sympatric populations did occur, patients were significantly more likely to be immunocompromised, indicating that non-sympatric MtB lineages may require some degree of host immunosuppression to compete with sympatric lineages. Mechanisms of MtB immune evasion, therefore, may have been shaped by population-specific variation in human immune response.

While the above discussion has focused mainly on pulmonary TB, we note here that extra-pulmonary TB, a less common and more severe form of disease, may be especially amenable to analyses guided by co-evolutionary hypotheses. This form of the disease leads more quickly to fatality and results in fewer trans-missions with even in a cosmopolitan setting (Gagneux et al., 2006). Interestingly, when TB transmission in non-sympatric populations did occur, patients were significantly more likely to be immunocompromised, indicating that non-sympatric MtB lineages may require some degree of host immunosuppression to compete with sympatric lineages. Mechanisms of MtB immune evasion, therefore, may have been shaped by population-specific variation in human immune response.

While the above discussion has focused mainly on pulmonary TB, we note here that extra-pulmonary TB, a less common and more severe form of disease, may be especially amenable to analyses guided by co-evolutionary hypotheses. This form of the disease leads more quickly to fatality and results in fewer transmissions with even in a cosmopolitan setting (Gagneux et al., 2006). Interestingly, when TB transmission in non-sympatric populations did occur, patients were significantly more likely to be immunocompromised, indicating that non-sympatric MtB lineages may require some degree of host immunosuppression to compete with sympatric lineages. Mechanisms of MtB immune evasion, therefore, may have been shaped by population-specific variation in human immune response.

**HUMAN PAPILLOMAVIRUS**

**Human papillomavirus (HPV)** is the most common sexually transmitted infectious agent in the world, and the second most common infectious cause of cancer after *H. pylori* (de Martel et al., 2012). Cervical cancer is the major source of mortality associated with HPV, but the virus also causes cancers of the anus, vagina, penis, and oropharynx (zur Hausen, 1989; zur Hausen, 1991; Carter et al., 2001; de Martel et al., 2012). Although over 100 types of papillomaviruses infect humans, only a fraction of them are carcinogenic (Bernard et al., 2010). Infection with two specific types, HPV 16 and HPV 18, account for approximately 70% of cervical cancer cases worldwide, with the remainder of cases largely attributable to 14 other types (Bernard et al., 2010). Nevertheless, the great majority of infections with even carcinogenic HPV types are ultimately benign, demonstrating that HPV infection, although necessary, is not sufficient to cause of cervical cancer (Schiffman et al., 2005; Plummer et al., 2007).

Papillomaviruses (PVs) are notable for their slow rate of evolutionary divergence relative to other pathogens – only an order of magnitude higher than humans, in the case of HPV (Ong et al., 1993; Rector et al., 2007; Shah et al., 2010). This is commonly attributed to their use of high-fidelity host replication mechanisms (Van Doorslaer, 2013). A slow evolutionary rate precludes rapid adaptation to new hosts, and PV strains correspondingly show little evidence of interspecies transmission or related horizontal gene transfer (Herbst et al., 2009; Shah et al., 2010; Van Doorslaer, 2013). All carcinogenic types of HPV belong to a single genus of papillomaviruses that diverged from a common ancestor about 75 million years ago, predating the primate lineage (Rector et al., 2007; Van Doorslaer, 2013). By the emergence of *H. sapiens*, the common ancestor of HPV 16 and HPV 18 had diverged into separate species, and in fact HPV 16 and HPV 18 had already diverged from all other HPV types within their respective species clades (Lewin, 1993; Ong et al., 1993). Given this combination of early divergence, slow evolution, and strict host specialization, we would expect variants within HPV types independently to have similar phylogeographic patterns to that of *H. sapiens*. Global data collected for the two most frequently sexually transmitted types, HPV 16 and 18, reflect such a pattern (Bernard, 1994). The subtypes and variants of HPV 16 cluster into five major branches of a phylogenetic tree: European (E), Asian/American (AA), East Asian (As), and two African (Af1 and Af2) (Ho et al., 1993; Ong et al., 1993). Subtypes and variants of HPV-18 clustering into three major branches: African (Af), European (E), and Asian + American Indian (As+AI) (Ong et al., 1993).

Biochemical and bioinformatic analyses indicate that HPV evolution has not been entirely neutral. Viral genes expressed early during a PV infection, for example, appear to have evolved at different rates than those expressed late (Garcia-Vallve et al., 2005; Rector et al., 2007). Although most PV genes show signs of strong purifying selection, the exceptions appear to be important (DeFilippis et al., 2002; Chen et al., 2005; Carvajal-Rodriguez, 2008). Two genes under diversifying selection, E6 and E7, are essential for viral replication. They induce cell cycle progression in host cells, and encode proteins that, in the high-risk HPVs, are oncogenic (White et al., 1994; Doorbar, 2006; Klingelhutz and Roman, 2012). Of note, E6 and E7 interfere with the human tumor suppressor proteins, pRB and p53 (Dyson et al., 1989; Huibregtse et al., 1993a,b; Storey et al., 1998; Munger et al., 2004; Doorbar, 2006). In turn, polymorphisms in the human p53 gene were shown to modulate the tumorigenicity of HPV 16 and 18 (Storey et al., 1998). Patients homozygous for the p53Arg mutation were seven times more likely to develop cervical cancer than individuals with 1 or 2 p53Pro alleles (Storey et al., 1998). Other human polymorphisms, such as those in the genes *RPS* and *TYMS*, influence HPV transmissibility. In a study of high-risk HPV infections in Nigerian women, variants in these genes were shown to modulate risk of infection with HPV 16 and 18. Despite the effects described above, genetic variation in neither the host nor the pathogen has been successful in explaining most heritable risk of HPV-associated disease, when considered in isolation (Magnusson et al., 2000; Hildesheim and Wang, 2002; Wheeler, 2008; Chen et al., 2013; Shi et al., 2013, Table 1).

Because the integration of the HPV genome within the human genome is permanent, death of the host ends all possibility of viral multiplication and transmission. Even strains that damage the health of the host sufficiently to reduce human-to-human sexual contact can suffer a competitive disadvantage. Therefore, both host and pathogen should cooperate to prevent severe disease. As with *H. pylori* and MTB, there is some empirical evidence supporting the idea that humans and HPV types co-evolved to limit tumorigenesis, and that evolutionarily mismatched strains may be driving severe clinical outcomes. A study of high-grade
cervical intraepithelial neoplasia (CIN) and invasive cervical cancer in an Italian cohort of Caucasian women demonstrated that non-European variants of HPV16, A1 and AA, were found at an increased frequency in invasive lesions (Tornesello et al., 2004). A separate study of mostly Caucasian (81%) female university students in the United States showed that those infected with non-European HPV 16 variants were 6.5 times more likely to develop high-grade CIN than those with European variants (Xi et al., 1997). The same study demonstrated a similar HPV 16-related risk profile (4.5 relative risk) in a predominantly Caucasian (79%) population of women presenting at a sexually transmitted disease clinic (Xi et al., 1997). Finally, at the molecular level, there is some evidence that variants of the HPV 16 E6 protein, described above, may be better adapted for replication within specific hosts (DeFilippis et al., 2002).

DISCUSSION

Taken together, the three examples above illustrate how co-evolution can promote a reduction in antagonism between pathogen and host, and in doing so leave discernible signatures on the genomes of both species. If, as we argue here, the disruption of historical co-evolutionary relationships can explain many differences in disease outcomes, knowledge of the conditions under which such relationships arise and dissolve will be helpful in defining genetic architecture of disease etiology. The applicability of this model depends, to a large extent, on the degree of integration between host and pathogen genomes, which can take many forms.

A long-standing association between humans and pathogens may be a necessary factor for cross-genomic integration, as with the three pathogens we have discussed. In contrast, many infectious diseases that occur epidemically are caused by zoonotic pathogens for which the human host is an evolutionary dead end, such as Salmonella enterica and Borrelia burgdorferi (Sokurenko et al., 2006; Falush, 2009). Other pathogens have had limited occasion to co-evolve with humans, because they cause disease primarily on an opportunistic basis (e.g., Streptococcus pneumonia or Clostridium difficile) or over a broad range of hosts (e.g., Toxoplasma gondii) (Ajzenberg et al., 2004; Sokurenko et al., 2006). The epidemic outbreaks caused by these pathogens may leave detectable signatures on the human genome, but reciprocal evolution in the pathogen need not occur.

For human-specific pathogens that cause endemic diseases and are not recent, the likelihood that severe disease is the outcome of a co-evolutionary mismatch should increase with the overlap between host and pathogen fitness. The pathogenicity of vertically transmitted pathogens, for example, should decrease over time, because such pathogens often depend on host survival (and possibly reproduction) for transmission. However, a strong overlap between host and pathogen fitness can also exist in the absence of vertical transmission. A horizontally transmitted pathogen, such as HPV, can evolve to be largely benign insofar as it depends on a healthy host for transmission.

When a pathogen’s fitness depends on its ability to cause damage to its human host, as with Mtb, attenuated antagonism becomes a special case, and its disruption becomes more difficult to detect and requires more evidence to confirm. While Mtb strains that increase the duration of a transmissible state will generally have a competitive advantage, the optimal duration can be expected to vary based on many population-level parameters, such as host density. This probably explains why modern Mtb lineages that are more common in high-density urban populations exhibit greater virulence. On the other hand, if horizontal transfer is confined to small, isolated populations, it may be considered effectively vertical. With such pathogens, a better understanding of the co-evolutionary history will be necessary to infer whether severe disease is caused by disrupted co-evolution or by another factor, such as infection by a universally more virulent strain or an opportunistic infection in an immunosuppressed patient.

The life history of the pathogen is also important in assessing the possibility and nature of co-evolution. A pathogen typically faces a tradeoff between fecundity and longevity. Increased fecundity within a host increases the probability (or rate) of transmission, but may negatively affect host lifespan or mobility (Frank and Schmid-Hempel, 2008). Therefore, a pathogen’s position on the continuum between greater fecundity and increased longevity will often reflect the degree to which its fitness depends on the health of the host. The case of HPV is somewhat of an exception in this regard. Host immune responses can induce diverse strategies, creating HPV types that are highly fecund, or less fecund with few virions per host. Whereas highly fecund types are more likely to transmit, they are also more likely to induce a vigorous immune response leading to clearance. Low fecundity types on the other hand, are more likely to persist as subclinical infections that can lead to prolonged inflammation and eventually cancer (DeFilippis et al., 2002). However, human populations that co-evolved with specific variants of these persistent types may be less likely to develop cancer, as described above.

Another factor influencing the applicability of the model we propose is a pathogen’s recombinogenicity. In theory, a pathogen that recombines freely is more likely to be panmictic, and hence less likely to co-evolve with a particular human host population (Bull et al., 1991). In fact, epidemic disease outbreaks often follow recombination events, and the pathogens responsible for the epidemics often appear superficially clonal, likely reflecting the rapid proliferation of especially successful recombinant strains (Grigg et al., 2001; Heitman, 2006). A case in point is Neisseria meningitides (Falush, 2009), as well as the eukaryotic parasites Toxoplasma gondii and Plasmodium falciparum, which though able to recombine sexually, exhibit surprisingly limited genetic diversity (Grigg et al., 2001). On the other hand, the strict clonality of Mtb and HPV has likely favored co-evolution, leading to reduced antagonism, while recombination in H. pylori can disrupt the co-evolutionary relationship favored by vertical transmission.

Recombination can also occur via horizontal gene transfer, as among species within the microbiome (Smillie et al., 2011; Ravel et al., 2011; Liu et al., 2012). This would suggest that co-evolution might be a relatively weak force in shaping microbiota genetic variation. However, data possibly supporting human–microbiome co-evolution exist; for example, the strongest correlate of an individual’s microbiota identity is ethnicity (Benson et al., 2010; Human Microbiome Project Consortium, 2012). The extent to which this correlation is driven by mutual genetic factors is unclear.

www.frontiersin.org August 2014 | Volume 5 | Article 290 | 7
as recurring environmental exposure and frequent vertical transmission may also account for most, if not all of it (Turnbaugh et al., 2009). Assessing whether the genomes of the microbiome and humans are integrated will be a key area of research, as it relates to co-evolution and disease risk (McFall-Ngai et al., 2013).

CONCLUSION
While the prospect of introducing co-evolutionary interactions into genetic epidemiology models may appear to add a new layer of complexity to an already difficult problem, a co-evolutionary perspective should help us construct more precise and accurate hypotheses, improving our ability to find real and reproducible results. Importantly, co-evolved genes will not be neutral in either species, which may make their identification easier. Although many methods exist to find loci that are candidates to have evolved under selection (Aguileta et al., 2009; Karlsson et al., 2014), and these methods can assess the strength, timing, and direction of selection (e.g., balancing or positive), they are not at present well adapted to the study of joint patterns of selection.

If the ultimate goal is to find interacting genes that have co-evolved to be benign and are subsequently disrupted in disease, we will need to identify differential patterns of concerted selection in paired human and pathogenic loci from different populations. The limiting factor to the development of appropriate methods toward this end has probably been the lack of prospectively collected paired genetic data for humans and pathogens. Once these data are available, existing methods to detect epistasis within a species can be adapted for cross-species pathogens. Once these data are available, existing methods to detect epistasis within a species can be adapted for cross-species pathogens. Once these data are available, existing methods to detect epistasis within a species can be adapted for cross-species pathogens. Once these data are available, existing methods to detect epistasis within a species can be adapted for cross-species pathogens. Once these data are available, existing methods to detect epistasis within a species can be adapted for cross-species pathogens.

Finally, we should note that the ultimate impact of this approach may extend beyond infectious diseases to what are traditionally considered non-communicable diseases. For example, we now recognize that both gastric and cervical cancers, as well as atherosclerosis, may have origins in infection (Libby et al., 2002; Porta et al., 2011). The number of such examples will certainly expand.

ACKNOWLEDGMENTS
This study was supported by the National Center for Research Resources, grant UL1 RR024975-01, which is now at the National Center for Advancing Translational Sciences; National Cancer Institute Grant P01 CA28842, the Vanderbilt-Ingram Cancer Center, the Wendy Dio family and the TJ Martell Foundation. Scott M. Williams was partially supported by P20 GM103534.

REFERENCES
Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R. E., et al. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65. doi: 10.1038/nature11632
Achtman, M. (2008). Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62, 53–70. doi: 10.1146/annurev.micro.62.081307.162832
Aguileta, G., Refregier, G., Yockteng, R., Fournier, E., and Giraud, T. (2009). Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. Infect. Genet. Evol. 9, 656–670. doi: 10.1016/j.meegid.2009.03.010
Aizenberg, D., Banuls, A. L., Su, C., Dumetre, A., Demar, M., Carme, B., et al. (2004). Genetic diversity, clonality and sexuality in Toxoplasma gondii. Int. J. Parasitol. 34, 1185–1196. doi: 10.1016/j.ijpara.2004.06.007
Anderson, R. M., and May, R. M. (1982). Coevolution of hosts and parasites. Parasitology 85, 411–426. doi: 10.1017/S0031182000055360
Aasheim-Hurtig, M., Dalide, G., Laehmann, M., Kalia, A., Ilver, D., Vikström, S., et al. (2004). Functional adaptation of BabA, the H. pylori ABO blood group antigen binding adhesin. Science 305, 519–522. doi: 10.1126/science.1098801
Backstrom, A., Lundberg, C., Kersulyte, D., Berg, D. E., Boren, T., and Amqvist, A. (2004). Metastability of Helicobacter pylori bab adhesin genes and dynamics in Lewis b antigen binding. Proc. Natl. Acad. Sci. U.S.A. 101, 16923–16928. doi: 10.1073/pnas.040417101
Baldari, C. T., Lanzavecchia, A., and Telford, J. L. (2005). Immune subversion by Helicobacter pylori. Trends Immunol. 26, 199–207. doi: 10.1016/j.it.2005.01.007
Barreiro, L. B., Neyrolles, O., Bubb, C. L., Tailleux, L., Quach, H., McElreavey, K., et al. (2006). Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. PLoS Med. 3:e20. doi: 10.1371/journal.pmed.0030020
Barry, C. E. III, Boshoff, H. I., Dartois, V., Dick, T., Ehr, S., Flynn, J., et al. (2009). The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nat. Rev. Microbiol. 7, 845–855. doi: 10.1038/nrmicro2236
Bartha, I., Carlson, J. M., Brumme, C. J., McLaren, P. J., Brumme, Z. L., John, M., et al. (2013). A genome-to-genome analysis of associations between human genetic variation, HIV-1 sequence diversity, and viral control. Elife (Cambridge) 2:e01123. doi: 10.7554/eLife.01123
Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S., Kim, J., et al. (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc. Natl. Acad. Sci. U.S.A. 107, 18933–18938. doi: 10.1073/pnas.1007028107
Bernard, H. U. (1994). Coevolution of papillomaviruses with human populations. Trends Microbiol. 2, 140–143. doi: 10.1016/0966-842X(94)90602-5
Bernard, H. U., Burk, R. D., Chen, Z., van Doorslaer, K., zur Hausen, H., and de Villiers, E. M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology 401, 70–79. doi: 10.1016/j.virology.2010.02.002
Bik, E. M., Eckburg, P. B., Gill, S. R., Nelson, K. E., Pordum, E. A., Francois, F., et al. (2006). Molecular analysis of the bacterial microbiota in the human stomach. Proc. Natl. Acad. Sci. U.S.A. 103, 732–737. doi: 10.1073/pnas.0506655103
Blaser, M. J., and Kirschner, D. (2007). Does Helicobacter pylori protect against asthma and allergy? Gut 57, 561–567. doi: 10.1136/gut.2007.133462
Bull, J. J., Molineux, I. J., and Rice, W. R. (1991). Selection of benevolence in a host–parasite system. Evolution 45, 875–882. doi: 10.2307/2409695
Campbell, D. I., Warren, R. F., Thomas, J. E., Figura, N., Telford, J. L., and Sullivan, P. B. (2001). The African enigma: low prevalence of gastric atrophy, high prevalence of chronic inflammation in West African adults and children. Helicobacter 6, 263–267. doi: 10.1046/j.1083-4389.2001.00047.x
Carter, J. I., Madeleine, M. M., Shera, K., Schwartz, S. M., Cushing-Hagen, K. L., Wipf, G. C., et al. (2001). Human papillomavirus 16 and 18 L1 serology compared across anogenital cancer sites. Cancer Res. 61, 1934–1940.
Carvajal-Rodriguez, A. (2008). Detecting recombination and diversifying selection in human alpha-papillomavirus. Infect Genet. Evol. 8, 689–692. doi: 10.1016/j.meegid.2008.07.002
Casanova, J. L., and Abel, L. (2007). Human genetics of infectious diseases: a unified theory. EMBO J. 26, 915–922. doi: 10.1038/sj.emboj.7601558
Cavalli-Sforza, L. L. (2005). The Human Genome Diversity Project: past, present and future. Nat. Rev. Genet. 6, 333–340. doi: 10.1038/nrg1596
de Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., et al. (2012). Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* 13, 607–615. doi: 10.1016/S1470-2045(12)70137-7

de Jong, B. C., Hill, P. C., Aiken, A., Awine, T., Antonio, M., Adetifa, I. M., et al. (2008). Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. *J. Infect. Dis.* 198, 1037–1043. doi: 10.1086/591504

de Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., et al. (2012). Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* 13, 607–615. doi: 10.1016/S1470-2045(12)70137-7

de Sablet, T., Piazza, M. B., Loro, T., Bortoloni, M. C., Salzano, F. M., Pericchi, L. R., Zambrano-Guzmán, O., et al. (2008). Amerindian Helicobacter pylori strains go extinct, as European strains expand their host range. *PLoS ONE* 3:e3307. doi: 10.1371/journal.pone.0003307

Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clin. Sci. (Lond.)* 110, 525–541. doi: 10.1042/CS20050369

Dorer, M. S., Talario, S., and Salama, N. R. (2009). Helicobacter pylori's unconventional role in health and disease. *PLoS Pathog.* 5:e1000544. doi: 10.1371/journal.ppat.1000544

Dye, C., and Williams, B. G. (2010). The population dynamics and control of tuberculosis. *Science* 328, 856–861. doi: 10.1126/science.1185449

Dyson, N., Howley, P. M., Munger, K., and Harlow, E. (1989). The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243, 934–937. doi: 10.1126/science.2537322

El-Omar, E. M. (2013). *Helicobacter pylori* susceptibility in the GWAS era. *JAMA* 309, 1939–1940. doi: 10.1001/jama.2013.35590

Engel, L. S., Chow, W. H., Vaughan, T. L., Gammon, M. D., Risch, H. A., Stanford, J. L., et al. (2003). Population attributable risks of esophageal and gastric cancers. *J. Natl. Cancer Inst.* 95, 1404–1413. doi: 10.1093/jnci/dgd474

Everard, P. W., and Cochran, G. M. (2000). *Chlamydia pneumoniae* and cardiovascular disease: an evolutionary perspective on infectious causation and antibiotic treatment. *J. Infect. Dis.* 181(Suppl. 3), S394–S401. doi: 10.1086/315602

Falush, D. (2009). Toward the use of genomics to study microevolutionary change in bacteria. *PLoS Genet.* 5:e1000627. doi: 10.1371/journal.pgen.1000627

Falush, D., Wirth, T., Linn, B., Pritchard, J. K., Stephens, M., Kidd, M., et al. (2003). Traces of human migrations in *Helicobacter pylori* populations. *Science* 299, 1582–1585. doi: 10.1126/science.1080571

Frank, S. A. (1996). Models of parasite virulence. *Q. Rev. Biol.* 71, 37–78. doi: 10.1086/419267

Frank, S. A., and Schmid-Hempel, P. (2008). Mechanisms of pathogenesis and the evolution of parasite virulence. *J. Evol. Biol.* 21, 396–404. doi: 10.1111/j.1420-9101.2007.01480.x

Funk, D. J., Helbling, L., Wernigreen, J. I., and Moran, N. A. (2000). Intraspecific phylogenetic congruence among multiple symbiont genomes. *Proc. Biol. Sci.* 267, 2517–2521. doi: 10.1098/rspb.2000.1314

Gagneux, S. (2012). Host–pathogen coevolution in human tuberculosis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 850–859. doi: 10.1098/rstb.2011.0316

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B. C., Narayanam, S., et al. (2006). Variable host–pathogen compatibility in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2869–2873. doi: 10.1073/pnas.0511240103

Galagan, J. E. (2014). Genomic insights into tuberculosis. *Nat. Rev. Genet.* 15, 307–320. doi: 10.1038/nrg3664

Gao, L., Tao, Y., Zhang, L., and Jin, Q. (2010). Vitamin D receptor genetic polymorphisms and tuberculosis updated systematic review and meta-analysis. *Int. J. Tuberc. Lung Dis.* 14, 15–23.

Garcia-Vallve, S., Alonso, A., and Bravo, I. G. (2005). Papillomaviruses: different genes have different histories. *Trends Microbiol.* 13, 514–521. doi: 10.1016/j.tim.2005.09.003

Ghoshal, U. C., Tripathi, S., and Ghoshal, U. (2007). The Indian enigma of frequent *H. pylori* infection but infrequent gastric cancer: is the magic key in Indian diet, host’s genetic make up, or friendly bug? *Ann. J. Gastroenterol.* 102, 2113–2114. doi: 10.1111/j.1572-0241.2007.01324.x

Glynn, J. R., Whiteley, J., Bifani, P. J., Kremer, K., and van Soolingen, D. (2002). Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg. Infect. Dis.* 8, 843–849. doi: 10.3201/eid0805.020020

Gressmann, H., Linz, B., Ghai, R., Pleissner, K. P., Schlabach, R., Yamaoka, Y., et al. (2005). Gain and loss of multiple genes during the evolution of *Helicobacter pylori*. *PLoS Genet.* 1:e45. doi: 10.1371/journal.pgen.0010043

Grigg, M. E., Bonneysoy, F., Hehl, A. B., Suzuki, Y., and Boothroyd, J. C. (2001). Success and virulence in toxoplasma as the result of sexual recombination between two distinct ancestries. *Science* 294, 161–165. doi: 10.1126/science.1061888

Heitman, J. (2006). Sexual reproduction and the evolution of microbial pathogens. *Curr. Biol.* 16, R711–R725. doi: 10.1016/j.cub.2006.07.064

Herb, F., Thye, T., Niemann, S., Browne, E. N., Chinbubha, M. A., Gyapong, I., et al. (2008). ALOX variants associated with susceptibility to human pulmonary tuberculosis. *Hum. Mol. Genet.* 17, 1052–1066. doi: 10.1093/hmg/ddm378

Herbst, L. H., Lenz, J., Van Doorslaer, K., Chen, Z., Stacy, B. A., Wellenhof, J. F. Jr., et al. (2009). Genomic characterization of two novel rettillian papillomaviruses, *Chelonina mydas papillomavirus 1* and *Careta caretta papillomavirus 1*. *Virology* 383, 131–135. doi: 10.1016/j.virol.2008.09.022

Hershberg, R., Lipatov, S., Small, P. M., Sheffer, H., Niemann, S., Homolka, S., et al. (2008). High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol.* 6:e310. doi: 10.1371/journal.pbio.0060311

Hildesheim, A., and Wang, S. S. (2002). Host and viral genetics and risk of cervical cancer: a review. *Virus Res.* 89, 229–246. doi: 10.1016/S0168-1702(02)00191-0

Hill, A. V. (2001). The genomes and genetics of human infectious disease susceptibility. *Annu. Rev. Genomics Hum. Genet.* 2, 373–400. doi: 10.1146/annurev.genom.2.1.373
ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. J. Virol. 67, 6424–6431.
Pallen, M. J., and Wren, B. W. (2007). Bacterial pathogenomics. Nature 449, 835–842. doi: 10.1038/nature06248
Parwati, I., van Creveld, R., and van Soolingen, D. (2010). Possible underly-
ing mechanisms for successful emergence of the Mycobacterium tuberculosis-B. pseudotuberculosis genotype strains. Lancet Infect. Dis. 10, 103–111. doi: 10.1016/s1473-3099(09)70330-5
Peek, R. M. Jr., and Blaser, M. J. (2002). Helicobacter pylori and gastrointestinal tract adenocarcinomas. Nat. Rev. Cancer 2, 28–37. doi: 10.1038/nrc7023
Picard, C., Casanova, J. L., and Abel, I. (2006). Mendelian traits that confer predis-
position or resistance to specific infections in humans. Curr. Opin. Immunol. 18, 383–390. doi: 10.1016/j.coi.2006.05.005
Plummer, M., Schiffman, M., Castle, P. E., Maucort-Boulch, D., Wheeler, C. M., and HIV LAM Group. (2007). A 2-year prospective study of human papillomavirus persistence among women with a cytological diagnosis of atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion. J. Infect. Dis. 195, 1582–1589. doi: 10.1086/516784
Porta, C., Riboldi, E., and Sica, A. (2011). Mechanisms linking pathogen-
associated inflammation and cancer. Cancer Lett. 305, 250–262. doi: 10.1016/j.canlet.2010.07.012
Portevin, D., Gagneux, S., Comas, I., and Young, D. (2011). Human macrophage responses to clinical isolates from the Mycobacterium tuberculosis complex. Dis.

mune response. Nat. Genet. 44, 257–259. doi: 10.1038/ng.1080
Rothenbach, D., Winkler, M., Gonser, T., Adler, G., and Brenner, H. (2002). Inverse relation-
ship between gastric colonization of Helicobacter pylori and diarrheal illnesses in children: results of a population-based cross-sectional study. J. Infect. Dis. 182, 1446–1449. doi: 10.1086/351887
Rothenbach, D., Blaser, M. J., Bode, G., and Brenner, H. (2000). Inverse relation-
ship between gastric colonization of Helicobacter pylori and diarrheal illnesses in children: results of a population-based cross-sectional study. J. Infect. Dis. 182, 1446–1449. doi: 10.1086/351887
Ridouenhor, B. J., and Nuissier, S. L. (2007). Polygenic traits and parasitic local adap-
tation. Evolution 61, 368–376. doi: 10.1111/j.1558-5646.2007.00209.x
Rowell, J. L., Dowling, N. F., Yu, W., Yesupriya, A., Zhang, L., and Gwinn, M. (2014). A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. Nature 434, 81–87. doi: 10.1038/nature14608
Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., et al. (2013). Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. MBio 4, e00250-13. doi: 10.1128/microbiol.00250-13
R DEVELOPMENT TEAM. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
Rowell, J. L., Dowling, N. F., Yu, W., Yesupriya, A., Zhang, L., and Gwinn, M. (2012). A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. Nature 434, 81–87. doi: 10.1038/nature14608
Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., et al. (2013). Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. MBio 4, e00250-13. doi: 10.1128/microbiol.00250-13
R Development Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
R Development Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., et al. (2013). Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. MBio 4, e00250-13. doi: 10.1128/microbiol.00250-13
Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., et al. (2013). Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice. MBio 4, e00250-13. doi: 10.1128/microbiol.00250-13
R Development Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
R Development Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
NOS2A and TLR2 in African-Americans and Caucasians. *Int. J. Tuberc. Lung Dis.* 13, 1068–1076.

Wheeler, C. M. (2008). Natural history of human papillomavirus infections, cytologic and histologic abnormalities, and cancer. *Obstet. Gynecol. Clin. North Am.* 35, 519–536; vii. doi: 10.1016/j.ogc.2008.09.006

White, A. E., Livano, E. M., and Tlsty, T. D. (1994). Differential disruption of genomic integrity and cell cycle regulation in normal human fibroblasts by the HPV oncoproteins. *Genes Dev.* 8, 666–677. doi: 10.1101/gad.8.6.666

Wilfert, L., and Schmid-Hempel, P. (2008). The genetic architecture of susceptibility to parasites. *BMC Evol. Biol.* 8:187. doi: 10.1186/1471-2148-8-187

Wirth, T., Wang, X., Linz, B., Novick, R. P., Lums, J. K., Blaser, M., et al. (2004). Distinguishing human ethnic groups by means of sequences from *Helicobacter pylori*: lessons from Ladakh. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4746–4751. doi: 10.1073/pnas.0306629101

Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B., and Levin, R. B. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat. Genet.* 32, 569–577. doi: 10.1038/ng1202-569

Wroblewski, L. E., Peek, R. M. Jr., and Wilson, K. T. (2010). *Helicobacter pylori* and gastric cancer: factors that modulate disease risk. *Clin. Microbiol. Rev.* 23, 713–739. doi: 10.1128/CMR.00011-10

Xi, L. F., Koutsky, L. A., Galloway, D. A., Kuyper, J., Hughes, J. P., Wheeler, M., et al. (1997). Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.* 89, 796–802. doi: 10.1093/jnci/89.11.796

zur Hausen, H. (1989). Papillomavirus in anogenital cancer: the dilemma of epidemiologic approaches. *J. Natl. Cancer Inst.* 81, 1680–1682. doi: 10.1093/jnci/81.22.1680

zur Hausen, H. (1991). Viruses in human cancers. *Science* 254, 1167–1173. doi: 10.1126/science.1659743

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 June 2014; paper pending published: 30 June 2014; accepted: 05 August 2014; published online: 25 August 2014.

Citation: Kodaman N, Sobota RS, Mera R, Schneider BG and Williams SM (2014) Disrupted human–pathogen co-evolution: a model for disease. *Front. Genet.* 5:290. doi: 10.3389/fgene.2014.00290

This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics.

Copyright © 2014 Kodaman, Sobota, Mera, Schneider and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.