Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation

Laura C. Woods[a,1], Rebecca J. Gorrell[b,1], Frank Taylor[a,b], Tim Connallon[a], Terry Kwok[b,2], and Michael J. McDonald[a,2]

[a]School of Biological Sciences, Monash University, Monash, VIC 3800, Australia; and [b]Biomedical Discovery Institute, Monash University, Monash, VIC 3800, Australia

Edited by W. Ford Doolittle, Dalhousie University, Halifax, NS, Canada, and approved September 7, 2020 (received for review March 23, 2020)

Horizontal gene transfer (HGT) confers the rapid acquisition of novel traits and is pervasive throughout microbial evolution. Despite the central role of HGT, the evolutionary forces that drive the dynamics of HGT alleles in evolving populations are poorly understood. Here, we show that HGT alters the evolutionary dynamics of genetic variation, so that deleterious genetic variants, including antibiotic resistance genes, can establish in populations without selection. We evolve antibiotic-sensitive populations of the human pathogen *Helicobacter pylori* in an environment without antibiotic but with HGT from an antibiotic-resistant isolate of *H. pylori*. We find that HGT increases the rate of adaptation, with most horizontally transferred genetic variants establishing at a low frequency in the population. When challenged with antibiotic, this low-level variation potentiates adaptation, with HGT populations flourishing in conditions where nonpotentiated populations go extinct. By extending previous models of evolution under HGT, we evaluated the conditions for the establishment and spread of HGT-acquired alleles into recipient populations. We then used our model to estimate parameters of HGT and selection from our experimental evolution data. Together, our findings show how HGT can act as an evolutionary force that facilitates the spread of nonselected genetic variation and expands the adaptive potential of microbial populations.

experimental evolution | horizontal gene transfer | antibiotic resistance

There is a broad consensus that horizontal gene transfer (HGT) from locally adapted populations can provide for rapid adaptation to new niches (1). However, remarkably little is known about the selective forces that act on horizontally transferred genetic variants as they segregate in a population. Previous empirical studies suggest that, similar to de novo mutations, very few HGT events will result in a fitness increase for the host (2–4). The few theoretical models that take HGT into account make contrasting predictions about the fates of these alleles. For instance, a model that treats HGT as similar to de novo genetic variation predicts that only strongly beneficial horizontally transferred genetic variants can establish in microbial populations (5, 6). Conversely, a permissive model that allows for gene flow from individuals or environmental DNA from outside of the population shows that the fixation of neutral and weakly deleterious mutations is possible (7). If this is the case, microbial populations could harbor HGT-derived genetic variation that is predominantly harmful, yet such variation may increase the evolutionary potential for rapid adaptation to environmental change.

The permissive model of HGT is supported by analyses of DNA sequence data from microbiomes, which often contain large amounts of within-species genetic variation. Although the ratio of nonsynonymous to synonymous substitutions suggests that most selection is purifying (8, 9), the selective spread of HGT alleles has been detected in time-resolved microbiome data (9). The prevalence of "soft" selective sweeps (10, 11)—which preserve haplotype diversity surrounding the sweeping allele—suggests that alleles contributing to adaptation may have been present in the population long enough to recombine into a number of genetic backgrounds before eventually being favored by selection (12). In principle, this means that neutral and deleterious alleles that are introduced by HGT may be able to persist in a population until an environmental change causes them to become favored.

Experimental populations of microbes have been used to study the dynamics of genetic variation in evolving populations and contrary to microbiome data, suggest that adaptation is dominated by clonal interference and strong selective sweeps (13, 14). However, studies that incorporate sexual reproduction have revealed that recombination alters the molecular dynamics of genetic variation so that soft selective sweeps are possible (15–17). Incorporating HGT into experimental evolution has proven more of a challenge due to the disruptive effect of inducing competence in most model species used for experimental evolution (18). However, some pioneering evolution experiments with bacterial sex (conjugation) (19, 20) have confirmed that HGT can facilitate adaptation under conditions of strong selection for donor genes. Other evolution experiments with HGT by conjugation show that HGT can also facilitate the spread of genes that decrease fitness (21–23). For instance, Stevenson and coworkers (22) found that a plasmid with a fitness cost of ∼40% was able to fix in experimental

Significance

Bacteria can obtain genes from other bacteria, or the surrounding environment, by horizontal gene transfer (HGT). While it is clear that HGT is very important for microbial populations, it is not understood how HGT changes the rate or mechanisms of adaptation. In this study, we evolve populations of the bacteria *Helicobacter pylori* and use DNA sequencing to track the movement of HGT genes as they spread through the population. We show that HGT can help antibiotic resistance genes establish at a low frequency in a population, even in the absence of the antibiotic. We find that these HGT treatment populations flourish when treated with antibiotics, showing how HGT can potentiate adaptation to future environmental change.

Author contributions: L.C.W., R.J.G., T.K., and M.J.M. designed research; L.C.W., R.J.G., F.T., and T.C. performed research; L.C.W., F.T., and T.C. analyzed data; and L.C.W., R.J.G., T.C., T.K., and M.J.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]L.C.W. and R.J.G. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: terry.kwok@monash.edu or mike.mcdonald@monash.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2005331117/-/DCSupplemental.

First published October 14, 2020.

populations. In this case, "infectious HGT" drove the fixation of the plasmid, which encoded the means for its own transfer to new donors, converting new hosts despite the costs (22, 23). While conjugation is an important mechanism of HGT, many bacterial species, especially pathogens (24), import extracellular DNA through noninfectious, mechanisms (1).

To investigate the dynamics of horizontally transferred genetic variants, we study evolving populations of the gut bacterium *Helicobacter pylori*. *H. pylori* is naturally competent and provides an excellent model for experimental evolution with HGT. DNA that is directly added to cultures of *H. pylori* is taken up by cells (without any experimental manipulation required), and integrated into the bacterial chromosome (25). We track the evolutionary dynamics of horizontally acquired and de novo genetic variants using whole-genome metagenomic sequencing, estimate the fitness effects of these variants, and apply these data to a model that explains how genetic variation is maintained in populations evolving under both selection and HGT (7). We find that HGT increases the range of selective conditions under which genes can spread through a population, allowing deleterious and neutral genetic variants to become established and potentially contribute to adaptation after environmental change.

To test whether antibiotic resistance could spread in populations without antibiotic-mediated selective pressure, we propagated replicate populations of *H. pylori* P12 that are sensitive to the antibiotic metronidazole in metronidazole-free growth media (hereinafter, "antibiotic" refers specifically to metronidazole). Replicate populations were divided into two treatment groups: HGT and non-HGT. The HGT treatment was carried out by regularly adding genetic material from a "donor" strain of *H. pylori* to cultures of the recipient, *H. pylori* P12. The donor strain carries 34 fixed genetic differences with the recipient, including 23 variants in the *rdxA gene*, 1 in the *frxA* gene, and another 10 variants distributed across the genome (*SI Appendix*, Table S1). Mutations in *rdxA* and *frxA* have been shown to increase metronidazole resistance (26, 27). In addition, the *H. pylori* P12 populations each have a number of low-frequency standing genetic variants (*SI Appendix*). The growth media of HGT treatment populations were inoculated with donor DNA at intervals of ~23 generations (Fig. 1*A*). In addition to gene flow from outside the population, both the HGT and non-HGT treatment populations are capable of recombination between individuals in the population during the course of the experiment. Finally, after evolution without antibiotic for 7 wk (~161 generations), we transferred the HGT and non-HGT populations to growth media supplemented with the antibiotic metronidazole.

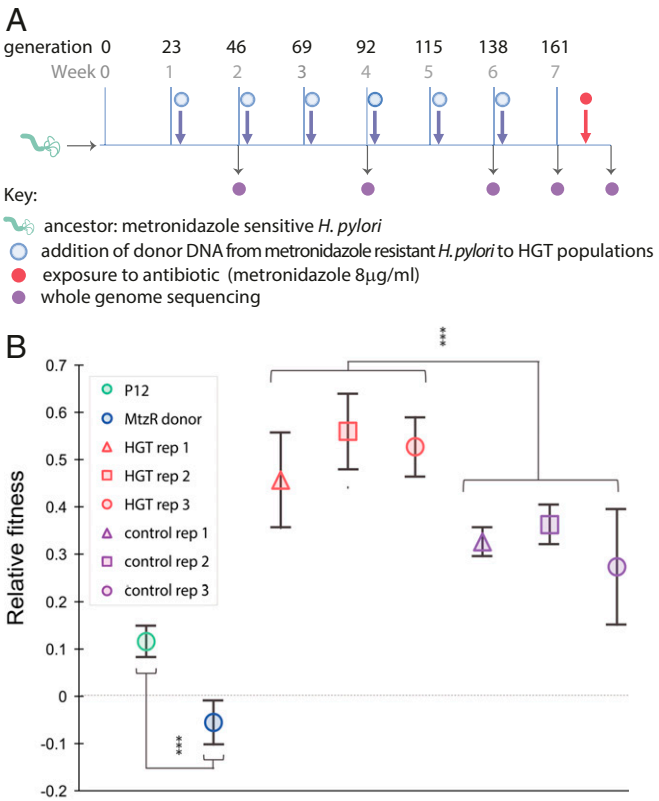

Fig. 1. Experimental populations with HGT evolve a higher fitness than non-HGT populations in antibiotic-free conditions. (*A*) Six replicate populations of *H. pylori* that are sensitive to the antibiotic metronidazole were propagated by 10-fold dilution every 24 h. After seven daily dilution cycles, cell-free gDNA from metronidazole-resistant *H. pylori* was added to the growth media (blue arrows) of the HGT treatment replicates only (HGT rep1, rep2, and rep3). After seven 7-d cycles, HGT and non-HGT controls were treated with antibiotic by plating on agar plates supplemented with metronidazole (8 μg/mL) (*Materials and Methods*). Whole-population samples were sequenced at several of the time points (purple dots) before and after exposure to antibiotic. (*B*) Competitive fitness assays were performed on all replicate populations after ~161 generations of evolution in antibiotic-free media. HGT treatment lines evolved a higher fitness (HGT replicates 1, 2, and 3; pink) relative to non-HGT control lines (control replicates 1, 2, and 3; purple). The metronidazole-sensitive ancestor (P12; green) is fitter than the metronidazole-resistant donor strain (MtzR; blue). Values are relative to a reference *H. pylori* strain with GFP and kanR markers (*Materials and Methods*). Error bars represent the 95% CIs around the mean. ***Significant difference between groups at $P < 0.001$.

Donor Strain Alleles Are Horizontally Transferred to the *H. pylori* Recipient Strain in the Absence of Antibiotic Selection

We carried out competitive fitness assays on all replicate populations after ~161 generations of evolution in antibiotic-free growth media (Fig. 1*B*). We found that all HGT populations evolved a higher fitness than the non-HGT populations (Welch's two-sample *t* test, $t = 5.8923$, $P < 0.001$), although both HGT and non-HGT replicates showed significant increases in fitness compared with the ancestor (HGT populations: Welch's two-sample *t* test, $t = 15.585$, $P < 0.001$; control populations: Welch's two-sample *t* test, $t = 6.3659$, $P < 0.001$).

The higher fitness of the HGT treatment populations compared with the non-HGT populations in the absence of antibiotic selection was surprising because the ancestor was fitter than the antibiotic-resistant donor strain (Welch's two-sample *t* test, $t = 8.3578$, $P < 0.001$) (Fig. 1*B*). It is possible that the addition of DNA to the growth media, which is part of the HGT treatment protocol, could have stimulated a higher rate of exchange between individuals in the HGT populations and increased the rate of adaptation (28). Alternatively, HGT may have resulted in the transfer of beneficial alleles from the donor strain into the evolving host populations. We tested for the presence of donor alleles by sequencing all replicate populations to a minimum 500-fold coverage at multiple time points during evolution to track the frequencies of de novo and HGT-originating mutations (Fig. 2, *SI Appendix*, Fig. S1 and Table S1, and Dataset S1). We found that, with the exception of one SNP (a single nucleotide polymorphism that was found in two out of three of the populations), all alleles from the donor strain, including resistance alleles, were maintained at detectable frequencies in all replicate HGT treatment populations for the length of the experiment (0.989%, 95% CI ± 0.368%) (Fig. 2). One putative beneficial allele from the donor strain, which reached a high frequency in all three HGT populations, was an SNP that repaired a nonsense mutation in the *hopG* gene that we determined was present in the recipient strain. The parallel spread of this allele through all HGT populations suggests that this restoration-of-function

substitution may account for some of the increased fitness in the HGT replicates compared with the controls (Figs. 1*B*, 2, and 3*A*).

Antibiotic Resistance Invades the Recipient Population, without Selection for Antibiotic

To verify the transfer of antibiotic resistance alleles from the donor strain to recipient populations, we performed cell counts to assess the frequency of the metronidazole resistance phenotype over the course of evolution. In addition, mutations in the metronidazole resistance-associated genes *rdxA* and *frxA* were identified by whole-genome sequencing to be at frequencies of around ∼1 to 5% in the HGT populations (Fig. 2 and *SI Appendix*). However, our phenotypic assays suggested much lower frequencies of resistance, at around 0.01% (±0.0074 95% CI) (*SI Appendix*, Table S2). To investigate this discrepancy, we engineered mutations into both *rdxA* and *frxA* and found that mutations in both genes are required for metronidazole resistance (*Materials and Methods*). This shows that a mutation in either gene alone is insufficient to cause resistance and explains the relatively low frequency of metronidazole resistance in the evolving HGT populations observed in our phenotypic assays. Consistent with this, the incidence of double mutants in our evolved cultures matches the expected frequency of double mutants under the independent segregation of alleles with frequencies of ∼1% each. Together, these results confirm that all alleles, including antibiotic resistance alleles, are being transferred from the donor strain into the evolving populations and are persisting there at low frequencies in the absence of antibiotic selection.

Low-Frequency HGT Alleles Rescue the Evolved Populations from Antibiotic Treatment

We next tested whether the low level of metronidazole resistance detected in the evolving HGT populations was sufficient to rescue these populations from exposure to the antibiotic metronidazole. We challenged all replicate populations after 161 generations of evolution in antibiotic-free growth media with lethal concentrations (8 μg/mL) of metronidazole and found that, while the antibiotic drove all non-HGT populations to extinction, HGT populations grew to high densities, consistent with antibiotic resistance (Fig. 2 and *SI Appendix*, Table S3). These observations are consistent with the preexisting antibiotic resistance alleles in HGT populations becoming rapidly fixed (Fig. 2) and thus, rescuing the populations during exposure to metronidazole.

We consider two mechanisms for the spread of HGT-derived alleles through the evolving population. First, genetic variants might only spread from the experimentally added genomic DNA into individual cells, without subsequent transfer between individuals. If this is the only means by which DNA can be transferred horizontally, the chance that the population will be able to survive an antibiotic treatment depends on the time since DNA was last added to the culture. After addition of metronidazole, the selective pressure provided by the antibiotic will promote the rare haplotypes that contain antibiotic resistance alleles, resulting in a "hard" selective sweep.

The second mechanism involves a combination of gene flow from outside the population and the spread of alleles between individuals. Between-individual transfer of alleles promotes the existence of an expanded set of genotypes in the population through recombination both before and after the initial transfer of genetic material from outside the population, allowing HGT-originating mutations to exist on multiple genetic backgrounds. We tested whether one or both mechanisms were at work by sequencing the evolved populations after the administration of the antibiotic. Populations that had been treated with metronidazole show a steep increase in the frequency of resistance-associated alleles in *rdxA* and *frxA*. However, de novo mutations, standing variation, and the HGT alleles outside of the *rdxA* and *frxA* resistance genes did not experience large changes in frequency (Fig. 2). These "soft sweep" dynamics suggest that the metronidazole resistance mutations occurred in many genetic backgrounds and that HGT alleles spread between individuals as well as from the environment.

Altogether, these data show how multiple antibiotic resistance alleles were able to spread through the HGT populations despite not conferring a selective benefit and subsequently facilitated adaptation to an antibiotic challenge. Our data also show that after the antibiotic resistance alleles were driven to a high frequency upon antibiotic exposure, the reservoir of background genetic diversity in each population was maintained. Overall, genetic exchange among individuals in the population limits the negative consequences of linkage between beneficial and

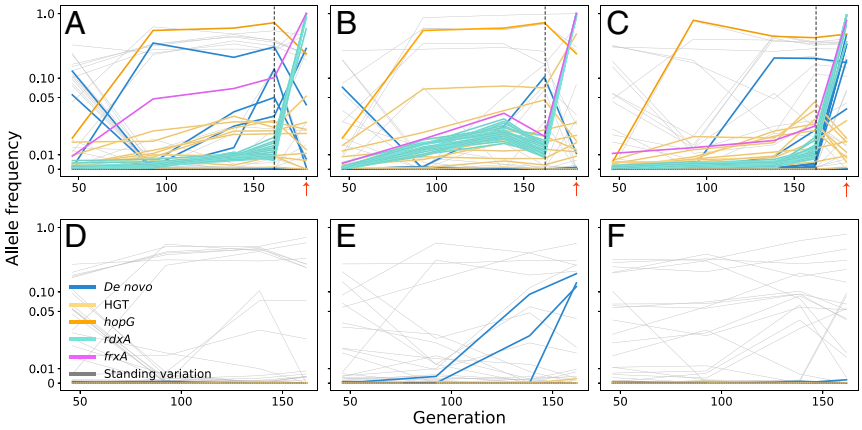

Fig. 2. Trajectories of allele frequencies during evolution, determined by whole-population sequencing. Each line depicts an individual allele segregating in three HGT (*A*–*C*) and three non-HGT control (*D*–*F*) populations plotted on a symlog scale (linear for values below 0.05). The dashed vertical lines after 161 generations show the frequencies of all alleles at the end of the no-antibiotic evolution experiment, before exposure to metronidazole (8 μg/mL). The frequencies at the end of the plot show allele frequencies after the exposure to metronidazole (small red arrows). The HGT alleles are shown in light orange except for the *rdxA* (turquoise), *frxA* (magenta), and *hopG* (bright orange) HGT alleles. Inactivation of the *rdxA* (turquoise) and *frxA* (magenta) genes is required for metronidazole resistance. Light gray lines show standing genetic variants, already present in the recipient population at the beginning of the experiment. The identity and frequency of HGT and de novo mutations are provided in *SI Appendix*, Table S1, and all mutations, including standing genetic variants, are described in *SI Appendix*.

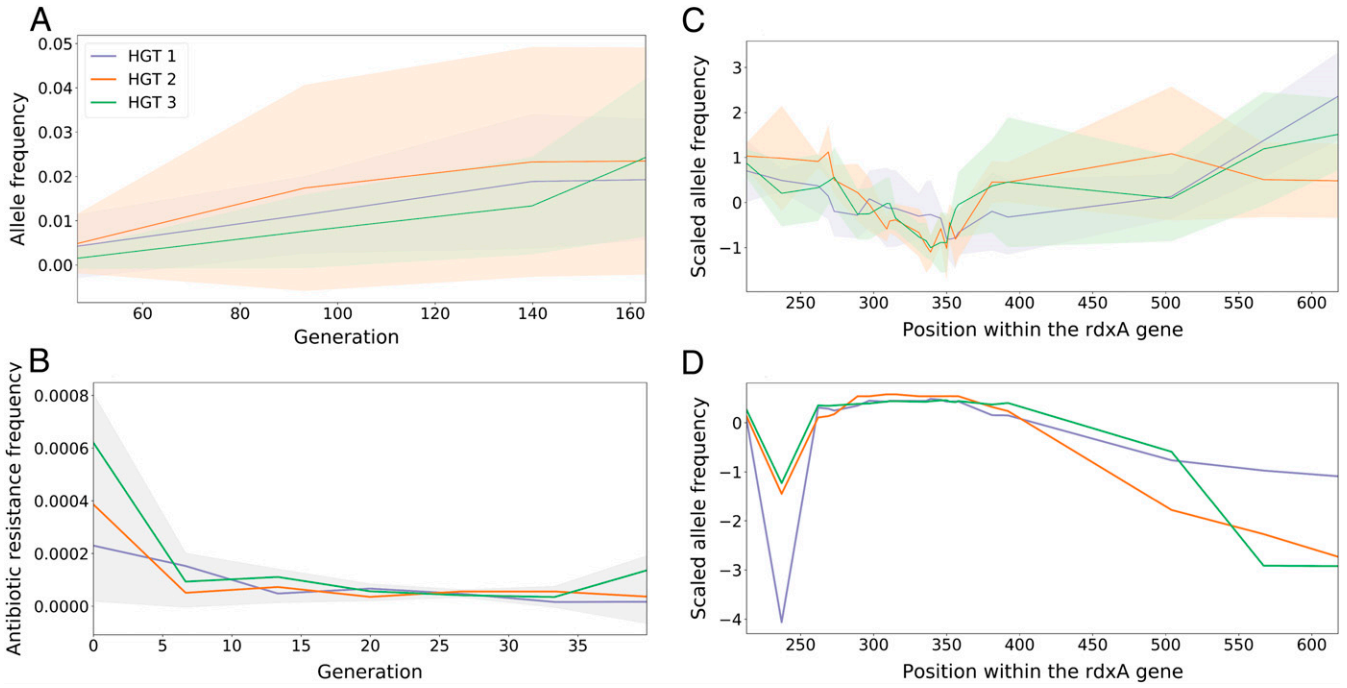

Fig. 3. Estimates of the fitness effects of HGT-derived variants. (*A*) The mean HGT allele trajectory for each HGT treatment population. Means were calculated from the 10 HGT alleles not associated with metronidazole resistance (*SI Appendix*, Table S1). (*B*) The three HGT treatment populations at generation ~161 were propagated for another ~40 generations in experimental conditions (media without antibiotic), without further HGT treatments. Note that before the start of this ~40-generation experiment, the HGT populations had already evolved for ~23 generations (from generation 138 to 161) without the addition of donor DNA. (*C*) The scaled distributions of allele frequencies of *rdxA* alleles are shown across three replicate populations after 161 generations of growth in media without antibiotic. Shaded regions are 95% CIs calculated from the four genome-sequenced time points before the addition of metronidazole (Fig. 1*A*). (*D*) The relative distribution of allele frequencies of *rdxA* alleles after exposure to antibiotic; since this distribution was calculated from a single time point, there are no CIs. Allele frequencies were scaled to account for between-population variation (*Materials and Methods*) (44). The distributions of *rdxA* alleles after addition of metronidazole were significantly different from the distribution before antibiotic (Kolmogorov–Smirnov D criterion = 0.3913, P = 0.001103) (*SI Appendix*, Fig. S2).

harmful mutations and incidentally preserves population diversity that may contribute toward adaptation to future challenges.

Gene Flow from HGT Provides an Evolutionary Force That Influences Allele Frequencies

In all of the evolved HGT populations, the frequency of donor-derived alleles consistently increased despite the absence of any antibiotic-mediated selective pressure (Fig. 3*A*). In experimental populations of microbes without recombination, a constant increase in frequency suggests the action of positive natural selection, either on the allele or on a linked driver allele (13). However, some theoretical results (7) suggest that genetic variants that enter the population by HGT can establish without selection or a genetic driver mechanism (22), although experimental evidence has been lacking. To determine whether gene flow from HGT was responsible for maintaining the high frequencies of these alleles, we sampled from all three HGT populations at generation 161 and then propagated them for 40 additional generations, without further addition of DNA or introduction of antibiotic. The rate of HGT-derived resistance initially declined steeply before leveling off at a lower, nonzero frequency, reduced from 0.041 ± 0.018% (95% CI) to 0.0035 ± 0.0018% (95% CI) (Fig. 3*B*). The ~10-fold decline in the frequency of metronidazole resistance suggests that there is selection against at least one of the alleles required for metronidazole resistance in the absence of the antibiotic. Taken together, the spread of metronidazole resistance alleles into the population when donor DNA is regularly added and the corresponding decline in frequency when donor DNA is not added suggest that HGT from outside the population is contributing to the establishment and maintenance of resistance alleles (Fig. 2). Moreover, these data suggest two evolutionary forces shaping the dynamics of metronidazole resistance alleles in media without antibiotic: gene flow from outside the population and purifying selection.

HGT Alleles with a Wide Range of Fitness Effects Establish in Evolving Populations

Allele frequencies are determined by their fitness effects, the fitness effects of linked alleles, and the rate of recombination. We looked for signatures of selection in a set of 23 synonymous, nonsynonymous, and indel alleles spread across the *rdxA* gene, which is ~600 bp in length. The scaled and centered values for each allele relative to all other alleles in the *rdxA* gene were calculated and averaged across the four sequenced time points prior to selection with antibiotic. We found the distributions to be highly similar across replicate populations, with a distinct valley in allele frequencies near the center of the gene (Fig. 3*C*), suggesting the action of purifying selection against one or more alleles in this region. After selection with metronidazole, the distribution pattern of the *rdxA* allele frequencies was reversed, significantly different from the distribution pattern of the alleles prior to selection (Kolmogorov–Smirnov D criterion = 0.3913, P = 0.0011) (Fig. 3*D* and *SI Appendix*, Fig. S1). This result is consistent with the competitive fitness assay showing that the antibiotic-resistant donor had a low fitness in antibiotic-free conditions (Fig. 1*B*) and the evolution experiment showing a decrease in metronidazole resistance over generations when HGT DNA is not supplied to the media (Fig. 3*B*). Altogether, these observations suggest that the *rdxA* alleles that cause metronidazole resistance have a low fitness in antibiotic-free conditions.

HGT Drives the Establishment of Deleterious Alleles, and the Fixation of Neutral Alleles, under a Wide Range of Conditions

HGT can take two forms in these experimental populations: unidirectional gene flow from outside the population ("invasion") and bidirectional gene flow between individuals within the population ("infection"). We can describe these dynamics using a single-locus model with two alleles developed by Novozhilov and colleagues (7) but extended to allow for bidirectional HGT between individuals (*SI Appendix*), which reflects the mechanism by which *H. pylori* and other naturally competent species undergo HGT for the genetic variants in our experiment. In our model, evolutionary trajectories of donor alleles are determined by gene flow into the population (invasion with rate γ) and HGT rates between individuals with different genotypes (infection with rates θ_a and θ_d representing the relative rates with which HGT converts carriers of ancestral vs. donor alleles, respectively). For a given locus, the effective selection coefficient of the donor allele is $s_e = s + (\theta_a - \theta_d)$, where s is the fitness effect of the donor allele (the selection coefficient). Assuming constancy of the parameters over time, the general solution for donor allele frequency at time t (in generations) is

$$p_t = \frac{(p_0 s_e + \gamma)e^{(s_e+\gamma)t} - \gamma(1-p_0)}{(p_0 s_e + \gamma)e^{(s_e+\gamma)t} + s_e(1-p_0)},$$

where γ is the infection rate and p_0 is the initial frequency of the donor allele ($p_0 = 0$ in the evolution experiment described in this study). The donor allele trajectory for a neutral locus with symmetrical HGT ($s_e = 0$) simplifies to

$$p_t = 1 - e^{-\gamma t}.$$

We generated an estimate of gene flow into the population (γ) by calculating the frequency of HGT events from our DNA sequence data. To do this, we selected a synonymous HGT-derived variant (*HPP12_0753*) (*SI Appendix*, Table S1) as a "neutral" reference. We assumed that the selective forces acting on this synonymous genetic variant were very weak (*Materials and Methods*) and that the trajectory of this allele will be determined by the number of HGT events that introduce it into the population. We inferred the rate of gene flow (γ = 2.46 E-4; 95% CI ± 2.86 E-5) and then applied this value to the general model to estimate fitness effects of the HGT donor alleles segregating in experimentally evolved populations. The analysis suggests that donor alleles with a wide range of selection coefficients are maintained in our experimental populations (Fig. 4*A* and *SI Appendix*, Fig. S3).

Numerical evaluation of the model, using the experimentally inferred γ, shows that empirically detectable deleterious allele frequencies can arise across a broad spectrum of effective selection coefficients (Fig. 4*B*). Although natural rates of gene flow are difficult to estimate and will probably vary widely depending on the mechanism of HGT and the availability of exogenous DNA, the model shows that neutral and weakly deleterious mutations can establish in a population experiencing even very low levels of gene flow (Fig. 4*C*), with more strongly deleterious alleles maintained within higher gene flow (high γ) regimes. An extended model of invasion, HGT, and interactions between alleles show that the single-locus results capture the dynamics when HGT between individuals is high but can otherwise overestimate the frequency trajectories of linked deleterious alleles or those exhibiting synergistic epistasis (*SI Appendix*). Thus, we expect the values of s_e calculated for the tightly linked mutations in *rdxA* could be distorted by the effects of linkage.

As discussed above, in our experimental system, metronidazole resistance was dependent on two loci, *rdxA* and *frxA*. We extended the model to consider two neutral or deleterious loci segregating in a population that become beneficial after the addition of antibiotic. We then carried out a stability analysis for the invasion of single or pairs of donor alleles in a population initially fixed (or nearly fixed) for ancestral, or recipient, alleles. We found that even if the two donor alleles are individually deleterious, but beneficial in combination, the double-donor genotype will spread within the population as long as the fitness benefit of the allele combination exceeds twice the average rate of HGT between individuals, per locus. Thus, while there is a linkage constraint to the evolutionary spread of rare, epistatically beneficial mutations [similar to constraints to the evolution of fitness "peak shifts" in standard models of selection with crossing over (29)], the constraint should be easy to overcome in the context of resistance evolution where selection is strong. Such constraints become imposing under weak epistatic selection.

Conclusion

Our experiments with laboratory populations of *H. pylori* confirm previous theoretical findings (7) that HGT can act as a directional evolutionary force in microbial species, promoting the spread of neutral and weakly deleterious alleles. This has implications for the dynamics of adaptation and the distribution of fitness effects of genetic variants in microbial populations. First, in our laboratory populations, genetic variation introduced and maintained by HGT was able to potentiate adaptation to an antibiotic challenge, whereas naïve populations that lacked HGT from outside the population ultimately perished. Understanding the evolution of antibiotic resistance is a major public health challenge (30), and it is becoming increasingly apparent that antibiotic resistance genes exist outside of antibiotic-selected settings (31–33). The mechanism that we identify in this study could contribute to the broad distribution of resistance genes and the rapid evolutionary response to treatment with antibiotics. Second, we found that unselected genetic variation, including antibiotic resistance genes, could persist in populations long enough to recombine into diverse genetic backgrounds. These data complement findings that HGT variation spreads through microbial populations by soft selective sweeps (9–11, 34) and show how genetic variants that are selected after an environmental change can rapidly fix in a population without the reduction of genetic diversity characteristic of a hard sweep (12).

Altogether, our experiments with the human pathogen *H. pylori* (35) have provided insights into the role of HGT in microbial evolution. Although our results are limited to controlled laboratory experiments with a single species, they point to a need for the further development of population genetics theory specific to microbial evolution with HGT, as well as experimental studies of the dynamics and fitness effects of horizontally acquired genetic variants in evolving populations.

Materials and Methods

Bacterial Strains and Growth Conditions. *H. pylori* P12 wild type (*cag* pathogenicity island+, CagA ABCCC, *vacA* m1s1, metronidazole sensitive) (36) was the ancestral strain for the evolution experiment. The isogenic metronidazole-resistant P12Δ*rdxA* donor strain was derived from P12 wild type through natural transformation with a PCR product containing a multiple cloning site flanked by the 5′ and 3′ ends of *rdxA* (HP0954), which was amplified from plasmid pRdxA (37).

In addition to carrying the insertional inactivation of *rdxA* that confers metronidazole resistance, this strain has diverged during standard laboratory propagation procedures. P12Δ*rdxA* carries 34 mutations that have accumulated during laboratory propagation during selection on metronidazole-containing agar, including mutations in *frxA*. The fluorescent reference strain used for fitness assays was *H. pylori* 7.13 GFP carrying the pTM117-flaA-gfp plasmid (38). *H. pylori* strains were routinely cultured on GC agar (Oxoid) supplemented with 10% (vol/vol) horse serum, 1× vitamin mix, 10 μg/mL vancomycin, and 20 μg/mL nystatin at 37 °C in a microaerobic atmosphere produced using BD GasPak EZ Campy gas generator sachets (Becton Dickinson)

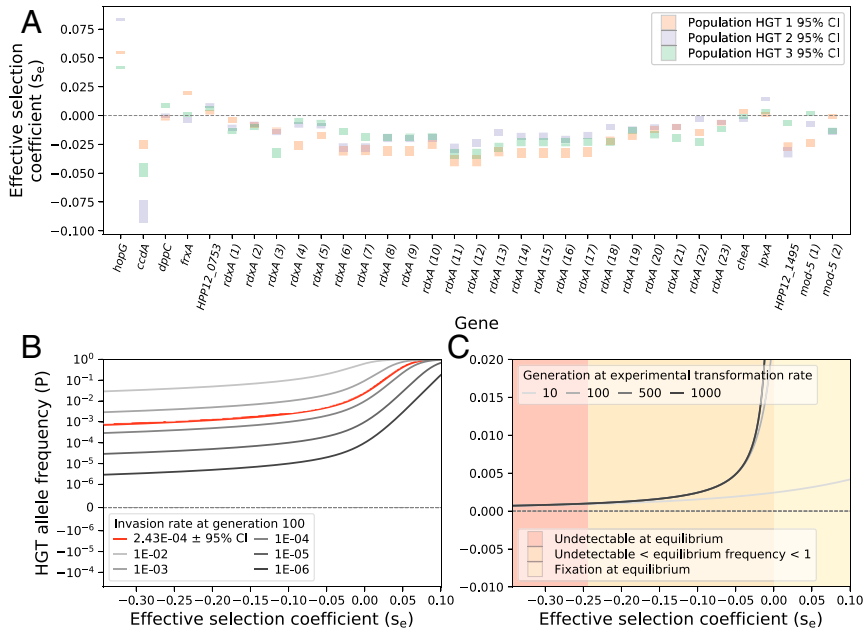

Fig. 4. Maintenance of genetic variation with HGT and fitness estimates from sequence data. *A* shows the estimates of the effective selection coefficient for donor-derived alleles, which were estimated, using the model, from allele frequency trajectories. Calculations were carried out for each of the three independently evolving populations, HGT rep 1 (orange bars), HGT rep 2 (purple bars), and HGT rep 3 (green bars) (*SI Appendix*, Fig. S3). The length of the bars indicates the 95% CI, calculated based on variation in HGT invasion rate (the model parameter γ), estimated from DNA sequence data (*Materials and Methods*). *B* shows the expected frequency of horizontally acquired alleles after 100 generations of evolution for several values of γ. The red curve corresponds to the HGT invasion rate (γ) estimated from our experiments, which give rise, for example, to ∼1% frequencies of deleterious alleles with $s_e \sim -0.025$ after 100 generations. *C* shows the expected frequencies of horizontally acquired alleles after 10, 100, 500, and 1,000 generations, given the HGT invasion rate (γ) estimated from the evolution experiment. The shaded areas correspond to three equilibrium frequency states for donor alleles (undetectably rare, detectable and polymorphic, and fixed).

in anaerobic chambers. Routine liquid *H. pylori* cultures were performed in Brain Heart Infusion (BHI) broth (Oxoid) supplemented with 10% (vol/vol) fetal bovine serum (Serena), 1× vitamin mix, and 10 μg/mL vancomycin and incubated at 37 °C in microaerobic atmosphere with shaking at 120 rpm. GC plates or liquid cultures were further supplemented with metronidazole (GCM; 1 to 16 μg/mL) for selection and routine culture of transformants as specified for individual experiments or 10 μg/mL kanamycin (GCK) for 7.13 GFP culture. The evolution experiment and fitness assays were performed in RPMI media supplemented with 25 mM MES buffer, 5% yeast extract (Oxoid), and 5 μg/mL vancomycin and incubated as a static culture at 37 °C in a 5% CO_2 incubator.

Evolution Experiment. Three replicate populations of *H. pylori* P12 were established in a 24-well plate in 500 μL of growth media per well at a starting bacterial density of optical density (OD_{600nm}) 0.1. Cultures were diluted 1:10 every 24 h (Fig. 1), and daily OD_{600nm} monitoring of each culture before transfer in the first week showed that this transfer dilution resulted in ∼3.3 generations per transfer. After the first week (i.e., seven transfers or ∼23.3 generations), each replicate population was expanded into two independent replicate transfers, and one replicate of each population immediately commenced weekly recurrent transformation using gDNA from the metronidazole-resistant P12Δ*rdxA* donor strain. For this, the freshly transferred cultures were incubated for ∼5 h in order for the *H. pylori* to develop natural competence for DNA uptake, after which ∼1 μg donor gDNA (concentration estimated using a Nanodrop reader) was added to the growth media of each HGT population. These populations are referred to as "HGT populations." The second replicate of each population was mock transformed in parallel using an equivalent volume of DNA-free buffer (i.e., 10 mM Tris, 1 mM EDTA) instead of gDNA, and they are referred to as "non-HGT populations." This pattern of HGT was repeated weekly and performed at every seventh transfer over a period of 6 wk (Fig. 1*A*). The antibiotic challenge performed after 161 generations was carried out by exposing all HGT and non-HGT control treatment populations on agar plates with 8 μg/mL metronidazole, and the cultures that survived were recovered and stored. Evolved populations were collected for storage throughout the course of the evolution experiment on the day preceding HGT (i.e., generations 46, 92, 138, and 161 and after antibiotic treatment). For this, the populations were first amplified to sufficient density for storage by growing 300 μL of each replicate population as a liquid overlay on GC agar for 24 h at 37 °C in 5% CO_2. The resulting bacterial growth was then resuspended in 1 mL BHI + 30% (vol/vol) glycerol for storage at −80 °C.

Phenotypic and Genotypic Monitoring of Metronidazole Resistance in Evolving *H. pylori* Populations. The *H. pylori* populations were monitored for the total viable bacterial cell count and proportion of metronidazole-resistant cells throughout the course of evolution, including the evolution experiment shown in Fig. 3*B*, where the HGT treatment populations at generations ∼161 were maintained for a further 40 generations without addition of donor DNA. The frequency of metronidazole resistance was monitored by spread plating 10-fold serially diluted cultures at 16 to 20 h and 7 d after each HGT event on both nonselective GC agar and GCM_8 agar containing 8 μg/mL metronidazole. Colonies were counted after 4 d of incubation.

The final antibiotic-naïve populations (i.e., generation 161) were also assessed for their extent of metronidazole resistance above that of the minimum inhibitory concentration for *H. pylori* P12 (0.19 μg/mL metronidazole), as well as above the resistance break point (>8 μg/mL metronidazole) by comparing the numbers of colonies that grew on GC, $GCM_{0.19}$, GCM_1, GCM_4, and GCM_{16} agar plates for each of the HGT and non-HGT populations, where the subscript indicates the concentration of metronidazole in micrograms per milliliter. Colonies were counted after 4 d of incubation. From these plates, 36 metronidazole-resistant clones were randomly selected for sequencing of their *rdxA* and *frxA* genes (oligo sequences in *SI Appendix*, Table S4). Sequencing data were analyzed to determine whether metronidazole resistance required mutation of one or both of these genes under conditions of low to high metronidazole concentrations.

Fitness of Evolved Populations. *H. pylori* strains (P12 evolved populations, P12 ancestor, P12Δ*rdxA* donor, and the fluorescent 7.13 reference strain) were grown in liquid culture overnight at 37 °C with shaking at 120 rpm. *H. pylori* 7.13 was used as the reference strain for fitness assays since it will not exchange DNA with the strain being measured. RPMI media were inoculated with both the 7.13 GFP fluorescent reference strain and a test strain at a 1:1 ratio at a starting OD_{600} of 0.2. Four replicates for each assay were propagated as described for the evolution experiment. On days 0, 1, 3, and 5,

dilutions of the grown liquid cultures were plated on duplicate GC and GCK agar plates and grown under standard growth conditions. Visualization of colonies was performed using the LAS-3000 Intelligent Darkbox Gel Imager (Fujifilm) with appropriate GFP filters. Selection coefficients for each of the strains of interest relative to the reference strain were calculated by taking the natural logarithm of the ratio of the two competing strains before and after evolution (18).

DNA Extraction and Bacterial Genome Sequencing. All bacterial gDNA was prepared from fresh plate-grown bacteria using the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) using the manufacturer's protocols. DNA was sent to GeneWiz for library preparation and sequencing on the Illumina MiSeq next-generation sequencing platform to generate 2 Gb of data for each paired-end 150-bp read set. GeneWiz trimmed adapter sequences from Illumina reads prior to data delivery.

Whole-Genome Sequence Analysis of Evolving Populations. Reads returned from GeneWiz were uploaded to a high-performance computing cluster running on Linux CentOS for processing and analysis. Given the amount of read data returned and the relatively small genome size of *H. pylori*, the average per-base depth exceeded 500 reads. This permitted the further processing of reads using a per-base Phred score threshold of 30 for filtering and a minimum length after trimming of 10 bp using BBDuk (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/). Bowtie2 alignment of the processed reads and polymorphic variant calling was performed using the breseq pipeline with the lowest minimum coverage thresholds for variant detection (39). To filter spurious sequence changes from downstream analysis, insertions and deletions in repetitive motifs were excluded in the variant calling program, and variants must have been detected in at least two time points unless they occurred either at the beginning or final time point, in which case they must reached a frequency threshold of 0.05. Coverage must be within two SDs of the mean, except as described in the case of *rdxA* mutations. On top of these filtering steps, for variants to be designated as HGT, they must have occurred in at least two HGT populations at detectable frequencies, no more than once in any control line sequences (to account for read errors), and be detectable at a frequency of at least 0.05 in the donor strain. For classification as a de novo mutation, variants must occur exclusively in a single population and rise to a frequency of at least 0.02. Because variant calling programs do not always correctly recognize that a variant has occurred in a homopolymeric or repetitive sequence, final manual verification of HGT-generated and de novo variants was carried out by examination of the read alignment with the Interactive Genome Viewer (40, 41). Any variants that were identifiable but did not meet criteria for either de novo or HGT designation were considered part of the standing genetic variation.

As breseq is unable to recover large novel sequence changes, an alternative "allele-binning" strategy was employed to resolve such variants, which were known to have occurred in the *rdxA* gene. A de novo assembly of the metronidazole-resistant donor strain was produced using SPAdes (42), and the region corresponding to the *rdxA* gene as well as the 150 bp flanking the ends of the gene was extracted. The mutant sequence was concatenated to the corresponding wild-type sequence separated by a 2,000-bp ambiguous nucleotide ("N") string to produce "bins" against which the reads of all evolved sequences were aligned and coverage was determined to find overall frequencies of the mutant alleles. Due to the recombinogenic potential of *H. pylori*, even if a read was "binned" to one sequence, it was possible that it contained the allele at a particular locus of the other sequence, especially in the case of well-spaced SNPs. Therefore, the relevant loci were analyzed for variants within both bins. Misaligned reads were removed from analysis, and coverage values were modified to account for their removal.

Construction of *rdxA*$^-$/*frxA*$^-$ Double Mutants. PCR products spanning *rdxA* and *frxA* were amplified from metronidazole-resistant P12Δ*rdxA* and used to transform metronidazole-sensitive *H. pylori* P12 (oligo sequences in *SI Appendix*, Table S4). The *rdxA*$^-$ allele was supplied as genetic material for transformation, either alone or together with *frxA*$^-$. In both cases, successful transformants were selected on GCM_8 plates, although single clones were patched and scaled up for liquid culture in nonselective media to prevent further mutagenesis resulting from exposure to metronidazole (43). The *rdxA* and *frxA* regions of all clones were PCR amplified, and the sequences were verified to confirm the necessity of both genes to impart metronidazole resistance.

Calculation of Scaled Proportions for *rdxA* Alleles. Alleles in the *rdxA* locus were found at a range of frequencies in each population at different time points through the evolution experiment. We wanted to quantify the relationship between individual alleles within the *rdxA* gene. To compare the relative frequency of *rdxA* alleles before and after the addition of antibiotics, we generated standardized variables from *rdxA* allele frequencies across all time points and replicates. This was carried out using the R "scale" function (44). Briefly, each allele site is designated as a "column" where allele frequencies are converted to a scaled value in two steps. First, to "center" allele frequencies, the column mean is subtracted from the allele frequency. Then, to scale allele frequencies, the scaled allele frequency is divided by the column SD. After scaling, the relative frequency of individual *rdxA* alleles across the gene can be compared with the distribution of *rdxA* allele frequencies in another population or with the distribution of *rdxA* allele frequencies after selection on antibiotic. The frequencies from multiple time points during the evolution experiment were combined to generate CIs. The distribution of *rdxA* allele frequencies from after antibiotic treatment was compared with the distribution from the combined preantibiotic time points using a Kolmogorov–Smirnov test.

Application of the Mathematical Models to Experimental Evolution Data. From our theoretical model, the transformation rate of a population is related to the frequency of neutral donor alleles by the following identity:

$$\gamma = -\frac{1}{t}\ln(1 - p_t),$$

where γ is the transformation rate, t is the number of generations since the onset of HGT, and p_t is the frequency of the (neutral) donor allele at time t. The above expression applies when the donor allele is initially absent from the population (as in the experiments described here). Frequency trajectories of donor alleles are approximately linear when p_t is small, in which case $\gamma \approx p_t/t$ for neutral donor alleles during the timescale of our experiments. Given the linear scaling between p_t and γt, the slope of the linear regression of neutral donor SNP frequency on time (in units of generations) provides an estimate of the transformation rate in our experimental populations. Moreover, the 95% CIs for estimates of γ can be calculated directly from CIs of the slope of the linear regression model.

We used the experimental frequency trajectories for a single synonymous SNP that was unlinked to any of the other variants in the experiment (and therefore, unlikely to have been influenced by hitchhiking with nearby SNPs under positive or purifying selection) to estimate the transformation rate in our experimental populations. Transformation rates and their 95% CIs were based on linear regression of frequency estimates on the time points at which frequencies were estimated; transformation rates were estimated using data from each of three biological replicates, independently (*SI Appendix*, Fig. S3), as well as from the allele frequency averages across replicates (Fig. 4). The four estimates were close to one another, suggesting that allele frequency changes conform approximately to our deterministic model.

Given the transformation rate estimated from the neutral SNP frequency trajectory, we estimated selection coefficients for putatively selected donor SNPs by numerically solving for s_e in the general solution for the allele frequency under arbitrary selection:

$$\frac{\hat{p}_t = \hat{\gamma}e^{(s_e+\hat{\gamma})t} - \hat{\gamma}}{\hat{\gamma}e^{(s_e+\hat{\gamma})t} + s_e,}$$

where $\hat{\gamma}$ is the transformation rate estimate from the neutral SNP and $\hat{p}_t$ is the estimated frequency of a given donor allele at generation t. In practice, we used allele frequencies estimated after $t = 161$ generations of experimental evolution without the antibiotic (initial donor allele frequencies were zero at generation $t = 0$). Calculations were carried out using the point estimates for γ from each replicate lineage and the estimate of γ based on the average neutral allele frequencies. Upper and lower CIs for the estimates of selection are based on the 95% CI for γ from the linear regression models (described immediately above).

Other Statistical Analyses and Data Visualization. Statistical analyses were performed in base R (33), including the dgof library (45), and in Python (Python Software Foundation) using the NumPy/SciPy (46) and Pandas (47) packages. Figures were produced using the native R library (44) and the Matplotlib package in Python (48).

Data Availability. Raw sequencing reads used to generate the data in Fig. 2 have been deposited in GenBank (accession no. PRJNA666798) (49). Custom scripts used for the simulations are available at GitHub (https://github.com/woodlaur189).

ACKNOWLEDGMENTS. T.C. is supported by funds from the Australian Research Council (ARC). T.K. and M.J.M. were supported by National Health and Medical Research Council Ideas Grant APP1186140 and a Monash University Science-Medicine Research Seed Fund. M.J.M. was supported by ARC Discovery Grant DP180102161 and ARC Future Fellowship FT170100441. We thank Imogen Scott for technical assistance and two anonymous reviewers for valuable comments on the manuscript.

1. S. M. Soucy, J. Huang, J. P. Gogarten, Horizontal gene transfer: Building the web of life. *Nat. Rev. Genet.* **16**, 472–482 (2015).
2. A. Knoppel, P. A. Lind, U. Lustig, J. Nasvall, D. I. Andersson, Minor fitness costs in an experimental model of horizontal gene transfer in bacteria. *Mol. Biol. Evol.* **31**, 1220–1227 (2014).
3. D. A. Baltrus, Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* **28**, 489–495 (2013).
4. R. Sorek *et al*., Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318**, 1449–1452 (2007).
5. O. G. Berg, C. G. Kurland, Evolution of microbial genomes: Sequence acquisition and loss. *Mol. Biol. Evol.* **19**, 2265–2276 (2002).
6. B. R. Levin *et al*., The population genetics of antibiotic resistance. *Clin. Infect. Dis.* **24**, S9–S16 (1997).
7. A. S. Novozhilov, G. P. Karev, E. V. Koonin, Mathematical modeling of evolution of horizontally transferred genes. *Mol. Biol. Evol.* **22**, 1721–1732 (2005).
8. S. Schloissnig *et al*., Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–50 (2013).
9. N. R. Garud, B. H. Good, O. Hallatschek, K. S. Pollard, Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS Biol.* **17**, e3000102 (2019).
10. M. L. Bendall *et al*., Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME J.* **10**, 1589–1601 (2016).
11. J. Barroso-Batista *et al*., The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLoS Genet.* **10**, e1004182 (2014).
12. P. W. Messer, D. A. Petrov, Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol. Evol.* **28**, 659–669 (2013).
13. G. I. Lang *et al*., Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571–574 (2013).
14. B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, M. M. Desai, The dynamics of molecular evolution over 60,000 generations. *Nature* **551**, 45–50 (2017).
15. M. J. McDonald, D. P. Rice, M. M. Desai, Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* **531**, 233–236 (2016).
16. K. Kosheleva, M. M. Desai, Recombination alters the dynamics of adaptation on standing variation in laboratory yeast populations. *Mol. Biol. Evol.* **35**, 180–201 (2018).
17. M. K. Burke, G. Liti, A. D. Long, Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Mol. Biol. Evol.* **31**, 3228–3239 (2014).
18. M. J. McDonald, Microbial experimental evolution: A proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* **20**, e46992 (2019).
19. V. G. Peabody, H. Li, K. C. Kao, Sexual recombination and increased mutation rate expedite evolution of *Escherichia coli* in varied fitness landscapes. *Nat. Commun.* **8**, 2112 (2017).
20. H. Y. Chu, K. Sprouffske, A. Wagner, Assessing the benefits of horizontal gene transfer by laboratory evolution and genome sequencing. *BMC Evol. Biol.* **18**, 54 (2018).
21. R. Maddamsetti, R. E. Lenski, Analysis of bacterial genomes from an evolution experiment with horizontal gene transfer shows that recombination can sometimes overwhelm selection. *PLoS Genet.* **14**, e1007199 (2018).
22. C. Stevenson, J. P. Hall, E. Harrison, A. Wood, M. A. Brockhurst, Gene mobility promotes the spread of resistance in bacterial populations. *ISME J.* **11**, 1930–1932 (2017).
23. A. J. Lopatkin *et al*., Persistence and reversal of plasmid-mediated antibiotic resistance. *Nat. Commun.* **8**, 1689 (2017).
24. P. Seitz, M. Blokesch, Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental gram-negative bacteria. *FEMS Microbiol. Rev.* **37**, 336–363 (2013).
25. S. Bubendorfer *et al*., Genome-wide analysis of chromosomal import patterns after natural transformation of *Helicobacter pylori*. *Nat. Commun.* **7**, 11995 (2016).
26. A. Goodwin *et al*., Metronidazole resistance in *Helicobacter pylori* is due to null mutations in a gene (rdxA) that encodes an oxygen-insensitive NADPH nitroreductase. *Mol. Microbiol.* **28**, 383–393 (1998).
27. J. Y. Jeong *et al*., Sequential inactivation of rdxA (HP0954) and frxA (HP0642) nitroreductase genes causes moderate and high-level metronidazole resistance in *Helicobacter pylori*. *J. Bacteriol.* **182**, 5082–5090 (2000).
28. M. S. Dorer, J. Fero, N. R. Salama, DNA damage triggers genetic exchange in *Helicobacter pylori*. *PLoS Pathog.* **6**, e1001026 (2010).
29. J. F. Crow, M. Kimura, Evolution in sexual and asexual populations. *Am. Nat.* **99**, 439–450 (1965).
30. M. A. Brockhurst *et al*., Assessing evolutionary risks of resistance for new antimicrobial therapies. *Nat. Ecol. Evol.* **3**, 515–517 (2019).
31. P. Dong, H. Wang, T. Fang, Y. Wang, Q. Ye, Assessment of extracellular antibiotic resistance genes (eARGs) in typical environmental samples and the transforming ability of eARG. *Environ. Int.* **125**, 90–96 (2019).
32. F. Benz *et al*., Clinical extended-spectrum beta-lactamase antibiotic resistance plasmids have diverse transfer rates and can spread in the absence of antibiotic selection. https://doi.org/10.1101/796243. (3 March 2020).
33. E. Bakkeren *et al*., *Salmonella* persisters promote the spread of antibiotic resistance plasmids in the gut. *Nature* **573**, 276–280 (2019).
34. N. J. Croucher *et al*., Evidence for soft selective sweeps in the evolution of pneumococcal multidrug resistance and vaccine escape. *Genome Biol. Evol.* **6**, 1589–1602 (2014).
35. B. N. Dang, D. Y. Graham, *Helicobacter pylori* infection and antibiotic resistance: A WHO high priority? *Nat. Rev. Gastroenterol. Hepatol.* **14**, 383–384 (2017).
36. W. Fischer *et al*., Strain-specific genes of *Helicobacter pylori*: Genome evolution driven by a novel type IV secretion system and genomic island transfer. *Nucleic Acids Res.* **38**, 6089–6101 (2010).
37. L. C. Smeets, N. Arens, A. A. van Zwet, C. M. Vandenbroucke-Grauls, J. G. Kusters, Transfer of the metronidazole resistance (*RDXA*) gene in *Helicobacter pylori* (abstract). *Gastroenterology* **118**, A677 (2000).
38. Y. N. Srikhanta *et al*., Methylomic and phenotypic analysis of the ModH5 phasevarion of *Helicobacter pylori*. *Sci. Rep.* **7**, 16140 (2017).
39. D. E. Deatherage, J. E. Barrick, Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).
40. J. T. Robinson *et al*., Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
41. H. Thorvaldsdottir, J. T. Robinson, J. P. Mesirov, Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
42. A. Bankevich *et al*., SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
43. G. Sisson *et al*., Metronidazole activation is mutagenic and causes DNA fragmentation in *Helicobacter pylori* and in *Escherichia coli* containing a cloned *H. pylori* RdxA(+) (Nitroreductase) gene. *J. Bacteriol.* **182**, 5091–5096 (2000).
44. R Development Core Team, *A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, Vienna, Austria, 2016).
45. T. B. E. Arnold, W. John, Nonparametric goodness-of-fit tests for discrete null distributions. *R J.* **3**, 34–39 (2011).
46. T. E. Oliphant, Python for scientific computing. *Comput. Sci. Eng.* **9**, 10–20 (2007).
47. W. McKinney *et al*., "Data structures for statistical computing in python" in *Proceedings of the 9th Python in Science Conference*, (SciPy, Austin, TX, 2010), Vol. 445, pp. 51–56.
48. J. D. Hunter, Matplotlib 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
49. L. C. Woods, R. J. Gorrell, F. Taylor, T. Connallon, T. Kwok, M. J. McDonald, Data files pertaining to "Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation." https://www.ncbi.nlm.nih.gov/bioproject/666798. Deposited 1 October 2020.