Evolutionary Transition from Pathogenicity to Commensalism: Global Regulator Mutations Mediate Fitness Gains through Virulence Attenuation

Gunther Jansen*,[1] Lena L. Crummenerl,[1] Felix Gilbert,[1] Timm Mohr,[1] Roxana Pfefferkorn,[1] Robert Thänert,[1] Philip Rosenstiel,[2] and Hinrich Schulenburg[1]

[1]Evolutionary Ecology and Genetics, Zoological Institute, University of Kiel, Kiel, Germany
[2]Molecular Cell Biology, Institute for Clinical Molecular Biology, University of Kiel, Kiel, Germany
*Corresponding author: E-mail: gjansen@zoologie.uni-kiel.de.
Associate editor: John True

Abstract

Symbiotic interactions are indispensable for metazoan function, but their origin and evolution remain elusive. We use a controlled evolution experiment to demonstrate the emergence of novel commensal interactions between *Pseudomonas aeruginosa*, an initially pathogenic bacterium, and a metazoan host, *Caenorhabditis elegans*. We show that commensalism evolves through loss of virulence, because it provides bacteria with a double fitness advantage: Increased within-host fitness and a larger host population to infect. Commensalism arises irrespective of host immune status, as the adaptive path in immunocompromised *C. elegans* knockouts does not differ from that in wild type. Dissection of temporal dynamics of genomic adaptation for 125 bacterial populations reveals highly parallel evolution of incipient commensalism across independent biological replicates. Adaptation is mainly achieved through frame shift mutations in the global regulator *lasR* and nonsynonymous point mutations in the polymerase gene *rpoB* that arise early in evolution. Genetic knockouts of *lasR* not only corroborate its role in virulence attenuation but also show that further mutations are necessary for the fully commensal phenotype. The evolutionary transition from pathogenicity to commensalism as we observe here is facilitated by mutations in global regulators such as *lasR*, because few genetic changes cause pleiotropic effects across the genome with large phenotypic effects. Finally, we found that nucleotide diversity increased more quickly in bacteria adapting to immunocompromised hosts than in those adapting to immunocompetent hosts. Nevertheless, the outcome of evolution was comparable across host types. Commensalism can thus evolve independently of host immune state solely as a side-effect of bacterial adaptation to novel hosts.

Key words: symbiosis, experimental evolution, *Pseudomonas*, immunocompromised, evolutionary transition, commensalism, global regulator, lasR.

Introduction

The recognition that interactions with microbes are essential for animal and plant function (Rakoff-Nahoum et al. 2004; Cash et al. 2006; Dethlefsen et al. 2007) has led to the emerging understanding of organisms as evolved symbiotic complexes (Gilbert et al. 2012; Douglas 2014). Yet it is not understood how well-known symbioses such as between leaf-cutter ants and fungi (Mueller et al. 2005; Schultz and Brady 2008), fig wasps and trees (Machado et al. 2001; Herre et al. 2008) or humans and their microbiota (Human Microbiome Project Consortium 2012) originated. Insights are traditionally inferred from comparative studies on closely related species groups such as *Wolbachia*, which shows the full spectrum of symbiosis, including parasitism, where one partner receives fitness benefits at the expense of the other, commensalism, which brings asymmetrical advantage to one partner at no cost to the other, and mutualism, which favors both (McGraw et al. 2002; Fry et al. 2004). Nevertheless, empirical evidence for the origin of symbiosis remains equivocal because confounding factors in the history of these associations cannot be excluded and, therefore, cause and effect relationships cannot easily be teased apart. Moreover, testing hypotheses on the role of the observed differences in the origin of symbiosis is usually impossible lacking the now extinct original populations.

Evolution experiments of incipient symbiosis between unfamiliar partners offer an alternative (Kawecki et al. 2012). The use of simple, genetically tractable model organisms with short generation times allows the study of evolving biological interactions in real-time, and supports dissection of mechanisms and mutations underlying adaptations to the biological interaction. Moreover, replaying evolution in independent biological replicates may provide insights in repeatability and idiosyncrasy of symbiosis evolution.

Experiments with incipient symbioses between unfamiliar bacteria, between Archaea and bacteria or between algae and yeasts have illustrated that evolving beneficial interactions often involves optimizing metabolite transfer rates (Hillesland and Stahl 2010) and may occur spontaneously when environmental conditions enforce metabolic

Open Access

interdependencies (Hom and Murray 2014), especially in spatially structured environments (Harcombe 2010; Hillesland and Stahl 2010). However, the genetic mechanisms underlying adaptation to novel interactions remain elusive. Moreover, the insights cited above were obtained from microcosms consisting of unicellular partners with metabolic dependencies engineered into their genomes or environments (Hillesland and Stahl 2010; Hom and Murray 2014), which therefore cannot explain incipient interactions between microbes and multicellular organisms with functional immune systems. Immune systems have recently been recognized as central to mediating the establishment and persistence of symbiosis (McFall-Ngai et al. 2013) and may have evolved as gate keepers that stringently control beneficial interaction partners and exclude pathogens. Compromised host immunity may thus result in altered adaptive evolution of interacting bacteria, different evolutionary dynamics, and/or changes in bacterial life history.

We asked how bacterial populations that infect a novel metazoan host adapt to the incipient interaction, using both immunocompetent and immunocompromised host strains. We used *Pseudomonas aeruginosa* PA14, a facultative human pathogen that can orally infect our model metazoan host, *Caenorhabditis elegans*, and establish a lethal gut infection (fig. 1; Tan et al. 1999). We serially passaged PA14 in populations of *C. elegans* across 13 transfers. To dissect the influence of host immunity, we used immunocompetent wild type (WT) and two defined immunocompromised *C. elegans* mutants that are nonfunctional in the p38 mitogen-activated protein kinase (MAPK) or the insulin-like receptor (ILR) immunity pathways. Deleterious mutations in one of these pathways shut down most of the *C. elegans* innate defense against pathogenic infection, particularly against *P. aeruginosa*, and thus render the mutants highly susceptible to infection (for a more in-depth discussion of the rationale behind our choice, see Discussion, and for reviews, see Schulenburg et al. 2008; Engelmann and Pujol 2010; Pukkila-Worley and Ausubel 2012; Cohen and Troemel 2015).

To investigate how bacterial populations adapted to the three host types during the evolution experiment, we analyzed the dynamics of phenotypic adaptation, that is, changes in PA14-induced host mortality, bacterial within-host fitness (infection loads), and fitness of hosts exposed to evolved bacteria. We then sequenced the full genomes of 125 bacterial populations across three time points of the evolution experiment to investigate the temporal dynamics of genetic adaptation underlying the phenotypes.

Results

Evolution Experiment

To investigate how PA14 adapt to immunocompetent and immunocompromised hosts, we performed an evolution experiment consisting of 13 serial passages of PA14 in populations of 300 nematodes (fig. 2A). As hosts we included the N2 Bristol laboratory strain of *C. elegans* and two immunocompromised mutant lines that carried a knockout in one of the central immunity genes of *C. elegans*: *daf-16* in the ILR
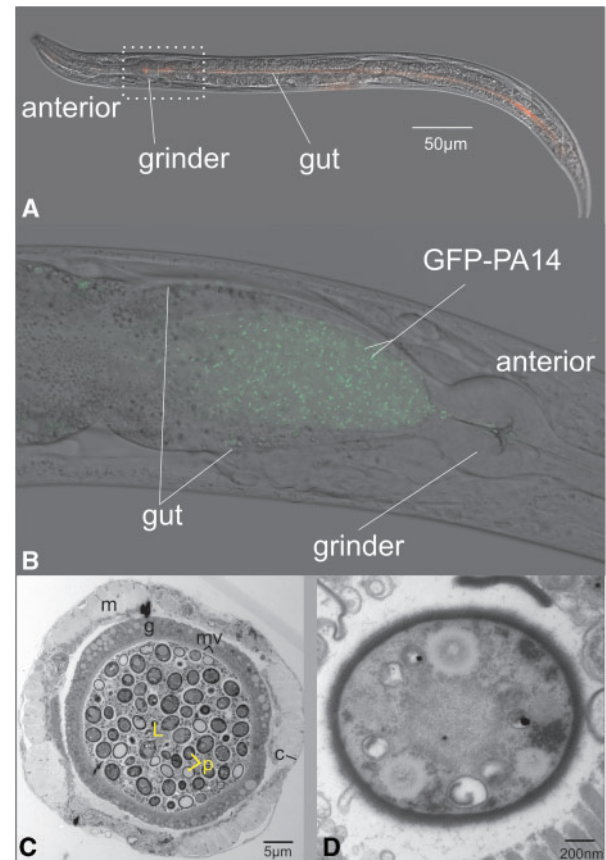

FIG. 1. (A) DIC microscopy image of healthy *Caenorhabditis elegans* nematode. Note narrow gut (colored red). (B) Representative image of nematode infected with fluorescently labeled *Pseudomonas aeruginosa* PA14. The region depicted corresponds to the square indicated in (A). *Caenorhabditis elegans* naturally feeds on bacteria and breaks up bacterial cells using the grinder. PA14 survives this process and colonizes the worm gut. Infection leads to distention of the gut lumen and, eventually, death of the nematode. (C) Transmission electron micrograph of cross-section of gut of infected worm. The host gut (g) is inflated by the large number of *Pseudomonas* cells (p) in its lumen (L) and presses the host organs against the cuticula (c). m, nematode muscle; mv, gut microvilli; p, *Pseudomonas* cells. (D) Detail of *Pseudomonas* cell in close proximity to host microvilli (visible on bottom right).

pathway and *nsy-1* in the p38 MAPK kinase pathway. These mutants are more susceptible to infection with PA14 and are killed more quickly than N2 worms on ancestral PA14 (supplementary fig. S1, Supplementary Material online). The experiment included 15 biologically independent replicates per treatment, thus totaling 45 evolving PA14 populations.

At the start of each transfer, bacterial population sizes were standardized and nematodes were infected through the natural oral route. Using auxotrophic media and stringent filtrations we excluded noninteracting environmental bacteria and imposed strong selection on *P. aeruginosa* to maintain close interactions with their novel hosts. At the end of each infection round, we extracted bacteria from all infected hosts for serial transfer and froze a subsample for subsequent phenotypic and genetic investigations. After population size standardization a new batch of host nematodes freshly grown up from stocks was infected for a new round.
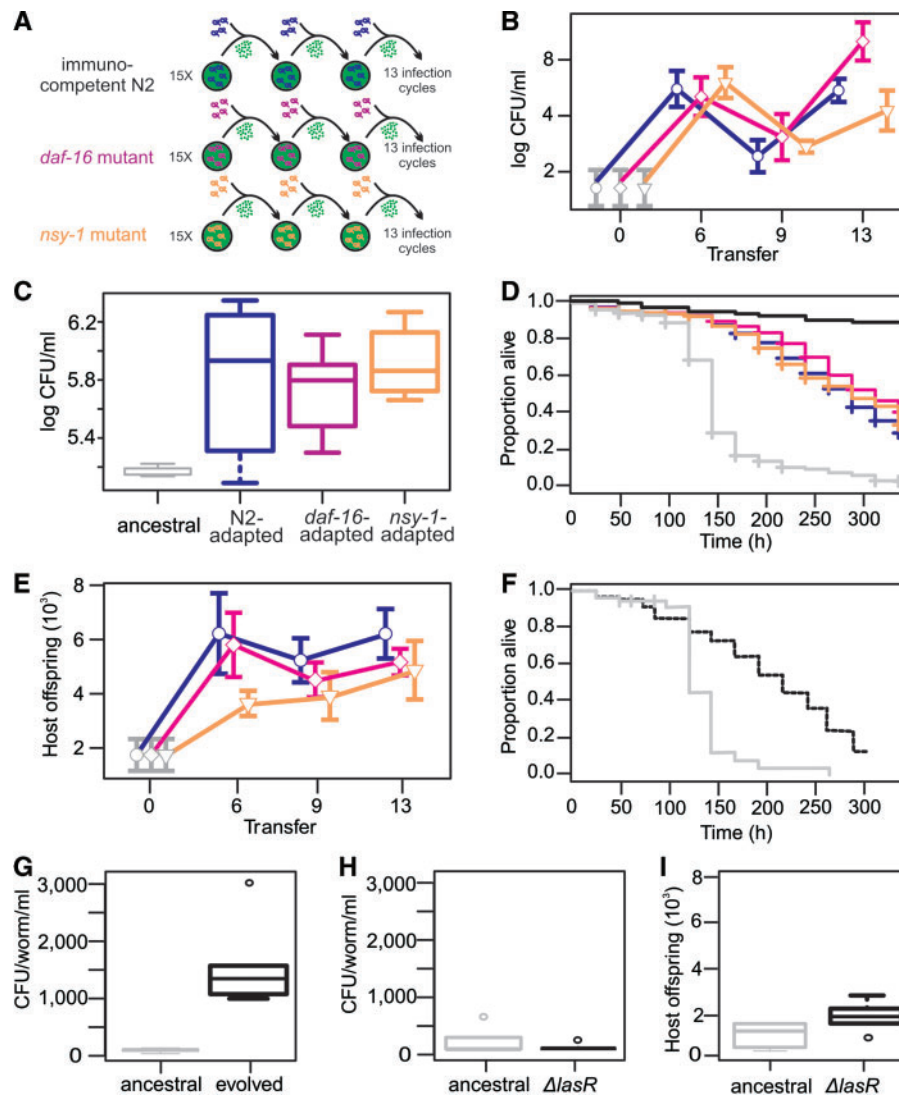
FIG. 2. Overview of experimental design and phenotypic assays. (A) Experimental design. The evolution experiment consisted of three host lines, the WT N2 strain and the two immmunocompromised strains mutated in *daf-16* and *nsy-1*. Populations of 300 nematodes fed on bacterial lawns consisting of a mixture of PA14 and heat-killed *Escherichia coli* on auxotrophic agar. The bacterivorous worms feed on bacteria and become infected. After 4 days, bacteria were extracted from all worms and a new round of infection was initialized. The experiment was repeated for 13 transfers, which corresponds to approximately 30 host generations. (B) Bacterial fitness of evolved PA14 quantified from 300 infected worms plus offspring (15 biological replicates). (C) CFU counts for bacterial suspensions extracted as in (B), but from starting host populations of exactly 30 worms and excluding offspring. 3×5 randomly chosen replicate populations. (D) Kaplan–Meier survival plot of *Caenorhabditis elegans* exposed to ancestral and evolved PA14. For each PA14 population, the survival of 30 N2 nematodes was scored in daily intervals. Each curve represents the average N2 survival across 15 biologically independent PA14 populations within each of the three treatments at the end of the experiment. (E) Host fitness, quantified as absolute number of offspring of three parental worms exposed to ancestral and evolved PA14 for 3×15 replicates. (F) Kaplan–Meier survival plots of 30 *C. elegans* exposed to ancestral PA14 (gray) and the $\Delta lasR$ mutant (black dashed), replicated three times. (G) CFU count for ancestral and 15 randomly chosen evolved PA14 extracted from 30 infected worms and analyzed using the manual method (see Materials and Methods for details). (H) Infection load and (I) host fitness measurements for ancestral PA14 and $\Delta lasR$ mutant ($N = 5$). Gray, ancestral PA14; black; uninfected control (survival on lab food *Escherichia* coli OP50); blue, N2-adapted PA14; magenta, *daf-16*-adapted PA14; orange, *nsy-1*-adapted PA14.

Phenotypic Evolution

Using the material frozen during the evolution experiment, we compared PA14 phenotypes from the beginning of the experiment (ancestor), and transfers 6, 9 and 13. We measured absolute bacterial and host fitness counted directly as offspring numbers, and proxied virulence using host survival assays.

To investigate the evolution of virulence, we exposed 30 *C. elegans* hosts to the evolved bacteria (for each treatment

and transfers 6, 9, and 13) and tracked the survival of each individual nematode across time. Although ancestral PA14 kill *C. elegans* within hours to days (supplementary fig. S1, Supplementary Material online), we found that *C. elegans* survived significantly longer on evolved bacteria by the end of the experiment (log-rank test, $P < 2 \times 10^{-16}$ for all treatments groups and all time points compared with ancestral PA14; fig. 2D). However, host survival was still lower than that of uninfected worms (log-rank test, $P < 2 \times 10^{-16}$ for all

treatments and time points, multiple comparisons to uninfected control). The survival assays suggest that PA14 rapidly evolved an attenuated virulence phenotype that resulted in lowered host mortality in both immunocompetent and immunocompromised hosts.

Next, we quantified bacterial fitness (infection load), as the total number of viable bacterial cells isolated from infected host populations, using the same extraction protocol and conditions as during the evolution experiment. This way, we ensured our measurements captured bacterial fitness in a way that was relevant for the selective regime during the experiment. Our results show that bacterial fitness increased quickly and strongly over time in all treatments relative to the WT (fig. 2B; Wilcoxon signed-rank test, $P = 1.58 \times 10^{-4}$, 3.56×10^{-3}, 2.59×10^{-4} for N2-, daf-16- and nsy-1-adapted PA14 of transfer 6, respectively; ns, ns, $P = 0.0212$ for transfer 9; $P = 2.59 \times 10^{-4}$, 6.58×10^{-4}, 7.63×10^{-3} for transfer 13). To exclude that these results could be explained by unequal host population sizes among treatments (see below), we repeated the fitness assay for transfer 13 with exactly 30 hosts and confirmed the outcome (fig. 2C; Wilcoxon signed-rank test, $P = 0.032$, 3.57×10^{-3}, 2.45×10^{-3} for N2-, daf-16- and nsy-1-adapted PA14, respectively). To further exclude the possibility that fitness increases were due to adaptation to other steps in the experimental protocol such as biofilm formation or other bacterial phenotypes, we manually quantified bacterial fitness at individual host levels (see Materials and Methods for details). Also this approach led us to conclude that infection loads were significantly increased at the end of the experiment (fig. 2G; Wilcoxon signed-rank test, $P = 7.94 \times 10^{-3}$). These carefully performed assays uniformly confirmed that evolved bacteria reached higher densities within infected hosts than ancestral PA14, showing that they evolved higher within-host fitness.

Finally, we measured host fitness, as the number of offspring individual nematodes produce during infection with PA14. We found that evolved PA14 supported larger host population sizes (fig. 2E; Wilcoxon signed-rank test, $P = 7.61 \times 10^{-3}$, 8.55×10^{-3}, 2.51×10^{-3} for N2-, daf-16- and nsy-1-adapted PA14 of transfer 6, respectively; for transfer 9: 6.54×10^{-4}, 6.58×10^{-4}, 0.015; and for transfer 13: 1.57×10^{-3}, 6.58×10^{-4}, 0.014). Because the nonevolving worms were taken from a frozen stock at each transfer, the latter can be entirely contributed to the effects of the evolved bacteria.

In conclusion, our phenotypic assays showed that bacteria adapted to the interaction with *C. elegans* hosts by reducing virulence (lowering host mortality), increasing bacterial within-host fitness, and supporting larger host populations with more offspring. Taken together, these results suggest that evolved PA14 received a double fitness benefit from virulence attenuation: They increased their cell numbers in individual worms and supported a larger number of hosts to infect. Thus, all of the 45 initially pathogenic PA14 populations became commensals and did so independent of immune state of the hosts.

Genomic Analyses

To investigate which molecular mechanisms had been selected for, and to what extent similar genetic functions evolved in independent replicates, we sequenced complete PA14 genomes for 125 populations using Illumina HiSeq 2000 Paired-End sequencing at 300× coverage. We obtained sequences for the ancestral PA14 strain and for all 15 biological replicate populations for each treatment at three stages: Early, intermediate, and the last transfer (transfers 6, 9, and 13; some replicate populations were lost during library preparation; see fig. 3B).

We corrected the published PA14 reference backbone for the variants that were detected in our ancestral PA14 strain (supplementary table S1, Supplementary Material online) and mapped the remaining samples to our ancestral reference genome with bwa (Li and Durbin 2010). Using a homemade genomic pipeline combining frequentist and heuristic methods, we analyzed single nucleotide polymorphisms (SNP), indels, structural variants (SV), and copy number variants (CNVs). We could reliably identify SNPs at a minimal variant frequency of 0.05, except when read depth fell below 50. Our analyses revealed that a few key mutations, including a frame shift insertion in *lasR*, a frame shift deletion in *rhlR* and nonsynonymous point mutations in *rpoB* arose early (by transfer 6) in the majority of genomes within all treatments, and usually remained fixed throughout the experiment (figs. 3B, 4A and B; supplementary table S2, Supplementary Material online). Additional mutations were mostly found at the final time point of the experiment, often within the *lasR/rhlR* regulon (figs. 3A, 4A and B). At transfers 9 and 13, the *nsy-1*-adapted bacteria had significantly more mutations than the other treatments (Wilcoxon signed-rank test, $P = 7 \times 10^{-4}$ and $P = 0.0229$, respectively; fig. 4B and C). At the last transfer, bacteria adapting to immunocompetent (N2) hosts acquired idiosyncratic mutations that were fixed in only one or few replicate populations, whereas immunocompromised-adapted PA14, especially in the *nsy-1* host, sustained more diverse populations containing multiple low-frequency genotypes (5–20% frequency within each population genome) in a more limited set of genes across independent replicates. Many of the mutations in the *nsy-1*-adapted PA14 were synonymous and intergenic, which suggests that they may not have been adaptive. These increased numbers of mutations may be the result of a potential mutator genotype, but we did not find mutations in any of the genes commonly associated with hypermutators (e.g., *mutL*, *mutS*, *mutM*, *uvrD*).

To statistically evaluate differences between treatments and within treatments across time, we further calculated population nucleotide diversity π in 5-kb sliding windows across each genome with 50% overlap. A comparable method was recently used to identify a new mviN domain-containing gene that is deleted in avirulent, experimentally coevolved *Bacillus thuringiensis* populations (Masri et al. 2015). Using a sliding Wilcoxon signed-rank test, we tested which of these genomic regions showed a significantly increased average nucleotide diversity, comparing either 1) 15 replicate populations of N2,
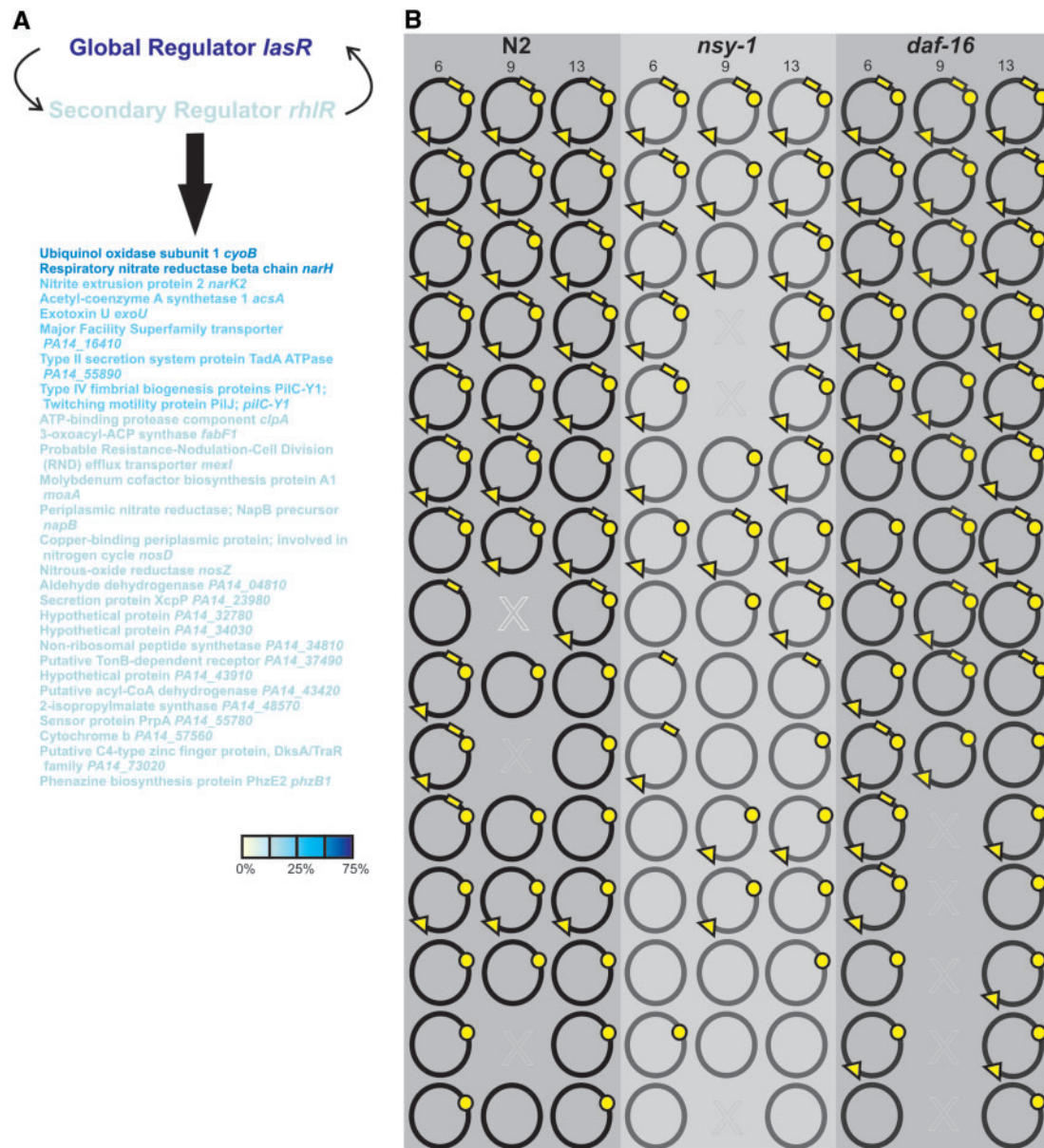
Fig. 3. Evolution of the *lasR* operon. (A) Schematic overview of the *lasR–rhlR* regulon. LasR and RhlR are autoinducers that activate and regulate each other and together control expression of approximately 200 genes across the PA14 genome. Listed genes belong to the regulon and have been affected by at least one mutation in at least two replicates in the experiment. Darker colors indicate that more replicates carried mutation(s) in that gene. Percentages of 45 replicate population genomes at transfer 13. (B) Overview of all 124 (minus ancestor) sequenced PA14 genomes adapting to one of three host types for each of three time points. Genomes of replicates sequenced at different times of the experiment are in the same position across transfers. Symbols indicate presence of key mutations in *lasR* (triangles), *rhlR* (polygons), *rpoB* (circles), and *ftsY* (rectangles). Crosses are replicates lost during library preparation.

15 replicate populations of *nsy-1*-adapted, and 15 replicate populations of *daf-16*-adapted populations at a particular time point (thus three tests, one for each time point); or 2) 15 replicates of N2-adapted populations at time point 6, 15 replicates of N2-adapted populations at time point 9, and 15 replicates of N2-adapted populations at time point 13 (thus again three tests, one for N2, one for the *nsy-1*-adapted and one for the *daf-16*-adapated populations). We found 130, 52 and 104 regions in the N2-adapted, *daf-16*-adapted and *nsy-1*-adapted bacteria, respectively, that showed significantly increased nucleotide diversity across time (sliding Wilcoxon

signed-rank test, $P < 0.05$ with false discovery rate [FDR] corrections; fig. 5 and supplementary table S3, Supplementary Material online). Thus, adapting PA14 populations collected increasing numbers of mostly synonymous or intergenic polymorphisms during the experiment, which mostly occurred in a few regions of the genomes and which were similar across independent biological replicates. This suggests that particular genomic regions were under selection, because the expectation from neutral evolution is more or less evenly distributed, nonfunctional mutations across the genome that by chance are different among independent biological
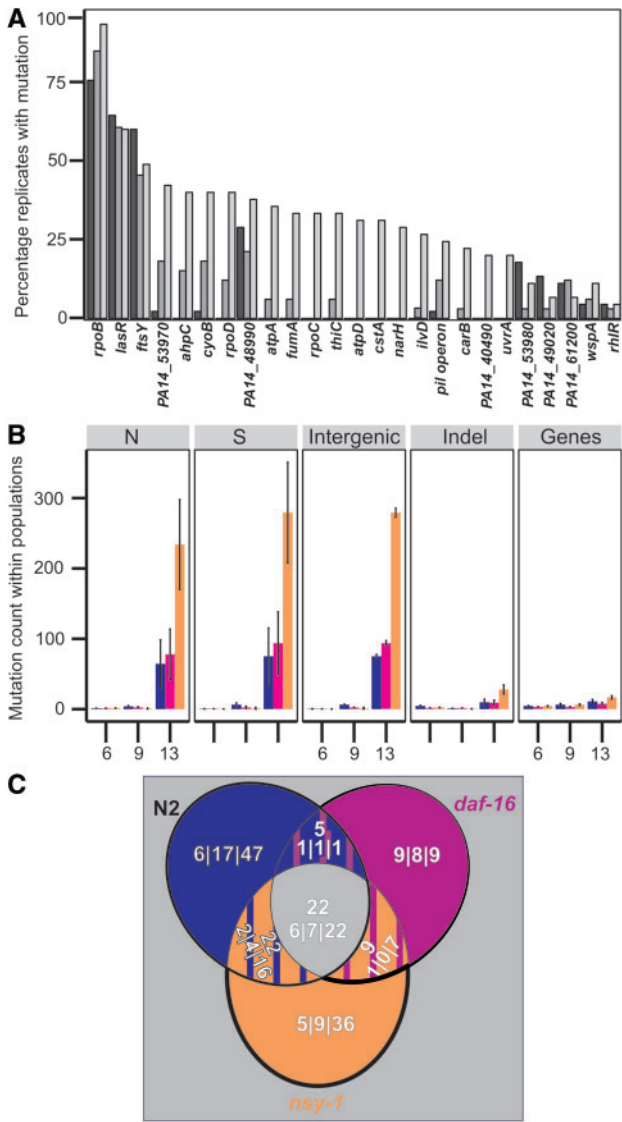
FIG. 4. Overview of mutations found in evolved genomes. Fifteen replicate genomes sequenced per treatment at transfers 6, 9, 13 (124 total; for transfer 9, 3, 3 and 5 population genomes are missing for the N2, *daf-16* and *nsy-1* treatments, respectively). (*A*) Histogram depicting percentage of replicate population genomes across treatments carrying at least one mutation (of any type) in a given gene. Black, transfer 6; dark gray, transfer 9; light gray, transfer 13. (*B*) Absolute number of mutations per sequenced population genome, averaged over 15 independent biological replicates per treatment per time point (34 for transfer 9) (mean ± SE). N, nonsynonymous; S, synonymous substitutions; intergenic, mutations in intergenic regions; genes, number of genes with at least one mutation. (*C*) Venn diagram summarizing genes affected by SNP and indels in the sequenced genomes. Numbers separated by bars represent number of genes found in transfers 6, 9, and 13. Numbers above indicate the number of shared genes irrespective of the transfer (genes mutated in parallel in different treatments, but not necessarily at the same transfer). Blue, N2-adapted PA14; magenta, *daf-16*-adapted PA14; orange, *nsy-1*-adapted PA14.

replicates. Functional analyses (see below) confirm that at least some additional mutations must have been beneficial. Moreover, nucleotide diversities were only significantly different between treatment groups for 103 regions at transfer 6,
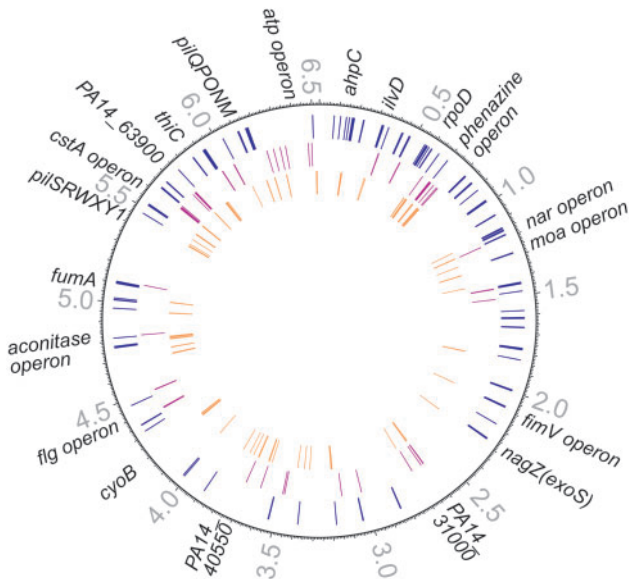

FIG. 5. Schematic overview of genomic regions with significantly different nucleotide diversity π across transfers in 5-kb sliding window with 50% overlap across the genome (based on all replicate population genomes for all time points). Outer, ring WT; middle, *daf-16*-adapted PA14, inner, *nsy-1*-adapted PA14. Each bar represents a 5-kb region which evolved significantly more nucleotide diversity. Numbers correspond to PA14 genome positions; names of genes in 20 most significant regions are provided.

but not at transfers 9 and 13 (supplementary table S4, Supplementary Material online). This suggests that PA14 populations adapting to the immunocompromised hosts reached high nucleotide diversities earlier in evolution than the ones adapting to immunocompetent hosts.

Functional term and gene expression category enrichment analysis of regions becoming more diverse over time suggested overrepresentation of genes involved in transcriptional regulation, chemotaxis/motility, amino acid/nitrogen metabolism, and transmembrane transport (table 1). Similar analyses of gene sets unique to each treatment did not yield additional functional categories, suggesting that PA14 adaptation was functionally similar in immunocompetent and immunocompromised hosts, but may have been faster in immunocompromised hosts.

Functional-Genetic Confirmation of *lasR*

Because of the similarity of both phenotypic and genomic results across treatments, and because of the high frequency of *lasR* mutations in most of the biologically independent replicates (fig. 4A), we focused on this gene for further analysis. To test whether the disabling frame shift mutation in *lasR* could explain the phenotypes we observed in the evolved PA14 (reduced virulence, increased bacterial within-host fitness and higher host fitness), we tested a $\Delta lasR$ loss-of-function mutant under our experimental conditions. Indeed, host mortality was significantly reduced when nematodes were infected with $\Delta lasR$ compared with the ancestral PA14 (fig. 2F; log-rank test, $P < 10^{-10}$), which confirmed the loss of virulence phenotype we found for the evolved PA14

Table 1. Results of DAVID Enrichment Analysis of Gene Lists Displaying Functional Categories of Significantly Enriched Clusters of Genes.

Gene List[a]	Functional Cluster	Enrichment Score
All PA14	Ribonucleotide/ATP binding	3.239
	Regulation of protein complex disassembly	2.583
	Nitrogen compound/nucleoside/nucleotide biosynthesis	2.422
	Hydrolase/carbon-oxygen lyase activity	1.569
	Copper ion binding	1.564
	Amine biosynthesis	1.297
	Amino acid derivative metabolism	1.284
	Intracellular/cytoplasm	1.268
	Transmembrane active transport	1.171
	Transcription/RNA metabolism	1.129
	Dehydrogenase	1.117
N2-adapted PA14	Ribonucleotide/ATP binding	2.109
	Glycine/serine/threonine metabolism	1.694
	ATP-ase couple transmembrane transport	1.530
	Cellular amino acid derivative biosynthesis	1.527
	Oxidative phosphorylation/hydrogen-proton transport	1.300
	Nitrogen metabolism	1.150
	Purine/pyrimidine metabolism	1.147
daf-16-adapted PA14	Nitrogen compound/nucleoside/nucleotide biosynthesis	2.921
	Ribonucleotide/ATP binding	2.247
	Chemotaxis/bacterial behavior	1.867
	Oxidative phosphorylation/hydrogen-proton transport	1.187
nsy-1-adapted PA14	Ribonucelotide/ATP binding	2.755
	Pyrimidine metabolism	1.909
	Nitrogen compound/nucleoside/nucleotide biosynthesis	1.360
	Oxidative phosphorylation/hydrogen-proton transport	1.160
	Aldehyde dehydrogenase	1.144

[a]Genes carrying at least one mutation in treatment.

bacteria. However, when we assayed bacterial (fig. 2H) and host fitness (fig. 2I), we could not find the increases observed in the evolved lines, suggesting that the lasR mutation may have been necessary, but not sufficient to explain the evolved phenotypes.

Discussion

To investigate how bacteria adapt to novel interactions with metazoan hosts, we serially passaged P. aeruginosa PA14 infecting metazoan C. elegans hosts for 13 host generations. To test the influence of host immune state on the evolutionary trajectories of adapting bacteria, we included the standard N2 laboratory strain and two immunocompromised host lines that carried knockout mutations in the daf-16 (ILR pathway) or nsy-1 (p38 MAPK pathway) genes that are centrally involved in C. elegans immune defense against bacterial pathogens, including P. aeruginosa.

We found remarkable parallel evolution across evolving populations and treatments, both on the phenotypic and genomic levels. Across treatments and 45 independent biological replicates, most PA14 populations rapidly evolved an attenuated phenotype causing much reduced host mortality. Loss of virulence was paired with increased bacterial fitness within individual host bodies. Moreover, hosts produced more offspring on evolved bacteria than on the pathogenic

ancestor. Importantly, hosts were not evolved during the experiment, such that these effects could be contributed to bacterial adaptation.

The similarity of adaptive paths among immunocompromised and immunocompetent treatments is rather surprising, because compared with healthy hosts, immunocompromised hosts are more susceptible to infection, have a reduced life span, shed more pathogens, and become infected earlier and thus have a longer infectious period. The generality of our results crucially depends on how representative our two chosen mutants were for immunocompromise in C. elegans.

The p38 MAPK pathway is now recognized as the major defense pathway against both Gram-negative and Gram-positive pathogens, and particularly important during P. aeruginosa infection (Troemel et al. 2006; Schulenburg et al. 2008; Engelmann and Pujol 2010; Cohen and Troemel 2015). An early forward genetic screen for hypersensitivity against Pseudomonas infection already identified the p38 MAPK pathway as the major defense pathway (Kim et al. 2002). It relays signals through NSY-1, a MAP kinase, and SEK-1, a MAP kinase kinase, through the transcription factor PMK-1. Upon infection with P. aeruginosa, PMK-1 differentially regulates expression of not only antimicrobials such as C-type lectins, ShK toxins, and CUB-like genes but also detoxifying factors

such as glutathione-S-transferases, acetyl-transferases, and UDP-glucuronosyl transferases (Troemel et al. 2006). After PMK-1-mediated activation of SKN-1, glutathione-S-transferases detoxify reactive oxygen species produced by the dual oxidase Ce-Duox1/BLI-3 that are formed as an initial defense to *P. aeruginosa* infection (Hoeven et al. 2011; Papp et al. 2012). The crucial involvement of this pathway in immune defense to many pathogens, particularly *Pseudomonas*, makes mutants in this pathway prime candidates for our purposes of creating immunocompromised hosts.

The ILR pathway acts in parallel to the p38 MAPK pathway, (the two have almost no targets in common), and is involved in innate immunity, longevity, and stress resistance (Troemel et al. 2006; Cohen and Troemel 2015). Insuline-like peptides activate the DAF-2/IGFR receptor, which results in activation of a kinase cascade that eventually leads to phosphorylation of the forkhead transcription factor DAF-16/FoxO. Activated DAF-16 migrates from the cytoplasm into the cell nucleus, where it upregulates expression of a large set of genes involved in resistance to stress, immunity, growth, and metabolism. *Daf-2* mutants are long-lived and resist infection by Gram-negative and Gram-positive pathogens better than WT nematodes (Garsin et al. 2003). This effect depends on intact p38 MAPK signaling (Troemel et al. 2006). The ILR signaling cascade plays an essential role in defense against *Pseudomonas* in the intestine (Evans et al. 2008). *Pseudomonas aeruginosa* (but not *Enterococcus faecalis* or *Salmonella typhimurium*) is able to activate the ILR signaling cascade in *C. elegans* through the nematodal *ins-7* gene, which leads to inhibition of *daf-16* and thus suppression of antimicrobial effector genes such as lysines (*lys-7*), saposins (*spp-1*), and thaumatin homologues (*thn-2*) in *C. elegans* (Evans et al. 2008). This is a specific mechanism of pathogenisis that involves host immune suppression and that depends on the *Pseudomonas gacA*, *lasR*, and *rhlR* quorum sensing genes. The ILR pathway and the *daf-16* mutant we used in this experiment are thus very well characterized in *C. elegans*, making this mutant a prime candidate for our purposes of studying immunocompromised model hosts. For these reasons, and based upon the pilot studies which confirmed increased susceptibility to PA14 of these mutants compared with immunocompetent *C. elegans*, we can firmly conclude that damaged immunity did not play a substantial role during longer-term adaptation of PA14 to *C. elegans* hosts.

Our phenotypic results show that virulence escalation is not always the optimal evolutionary route after niche expansion to new hosts. The virulence-transmission trade-off theory holds that virulence is optimal whenever it maximizes resource use within the host (and thus often host damage) but does not kill the host before efficient transmission (Anderson and May 1982). Thus, pathogens with environmental reservoirs may afford higher virulence than directly or vertically transmitted ones. Similarly, zoonotic pathogens, usually low-virulence animal symbionts, may be extremely virulent in humans because they pay no evolutionary cost for killing unfamiliar hosts. Interestingly, most serial passage experiments yield increasingly virulent parasites (Ebert 1998; Barclay et al. 2014). However, in natural populations, even

initially highly virulent pathogens often evolve reduced virulence (e.g., myxoma virus in Australian and European rabbits, at least until very effective rabbit immunity forced virulence up again; Kerr et al. 2012). This apparent discrepancy is usually interpreted within the trade-off framework as a lack of evolutionary costs for host killing when transmission is artificial (essentially mimicking zoonosis). In our experiment, maximized host resource use could in principle be accomplished through either increased pathogenicity, for example, early killing without compromising the host cuticle for transmission, or an entirely avirulent strategy. Both of these would have ensured passage of higher numbers of cells to the next round.

We observed the avirulent strategy. Our selection regime favored bacteria in both dead and alive hosts, as free-roaming cells were removed before transfer. Therefore, virulent PA14 could only have been transferred if they transmitted in time to host offspring or left dead host bodies intact. Very high virulence would have caused host extinctions or could have destroyed host bodies, precluding transfer. As in many natural populations, our setup thus selected against virulence that was excessive enough to endanger transmission. Nevertheless, PA14 cells did not become prudent pathogens that transmitted in time or saprophytic bacteria that leave the cuticle intact to ensure transmission. They became avirulent commensals. We demonstrated that this is the consequence of fitness advantages the evolved PA14 received from being commensals. Commensalism both increases bacterial cell production within single hosts and supports larger host populations, which allows transfer of even higher bacterial numbers. Reduced host damage thus provides a two-pronged bacterial fitness benefit, and pathogens become evolutionarily successful commensals.

To identify the genetic changes underlying the evolutionary transition from pathogenicity to commensalism, we analyzed 125 PA14 population genomes. We found that a few crucial mutations in *lasR* (and *rhlR* in genomes lacking *lasR* mutations) and *rpoB* evolved early in most replicates and remained fixed for the duration of the experiment in most genomes. Mutations in *rpoB*, which encodes the beta subunit of the DNA-directed RNA polymerase, are best known for their role in rifampin-resistance, but also occur during adaptation to the cystic fibrosis (CF) lung, particularly in combination with *lasR* lesions (Smith et al. 2006). Alterations to the polymerase and its sigma factors (encoded by *rpoC* and *rpoD*, which also carried mutations in some evolved genomes) influence the efficiency of expression through altered stability of transcription initiation complexes (Zhou and Jin 1998), and may ultimately enhance growth and reproduction at the cost of motility and stress resistance (Gummesson et al. 2009). Interestingly, these characteristics correspond to those identified by the functional term analysis of the population genomic results, suggesting that *rpo* mutations may have played a role in optimizing growth in the new host, for example, through a shift to novel amino acids or nitrogen-rich compounds.

The other candidate genes, *lasR* and *rhlR*, are part of the canonical quorum sensing locus that centrally regulates a

broad set of PA14 genes involved in virulence, secretion, and biofilm formation (a summary of genes within the regulon that were mutated in the evolved PA14 genomes is provided in fig. 3A). Several studies, including a forward-genetic screen, confirmed that loss of *lasR* function causes an avirulent phenotype in *C. elegans* (Lee et al. 2006; Feinbaum et al. 2012). Interestingly, besides *lasR*, several of the 170 other genes identified by Feinbaum *et al.* also carried mutations in the evolved PA14 genomes in this study. However, we could not find statistical overrepresentation of these genes in our data set, suggesting that more genes and functions were required for adaptation to novel hosts (see next paragraph). Nevertheless, it is intriguing that reduced-function *lasR* mutations evolved spontaneously during adaptation to the *C. elegans* host, which is also observed during clinical transitions of *Pseudomonas* infections in CF patients to the less virulent chronic state (D'Argenio et al. 2007; Hoffman et al. 2009). In CF patients, inactivation of *lasR* confers a growth advantage with amino acids and other nitrogen-rich compounds (D'Argenio et al. 2007), which in our study could correlate with the growth effects of *rpo* mutations.

Our functional genetic analysis of a Δ*lasR* mutant further showed that the commensal phenotype could not be fully explained by a knockdown of *lasR*, but also required additional mutations. Few mutations within a global regulator such as *lasR* can cause large pleiotropic phenotypic shifts and are more likely to evolve than combinations of mutations across a large array of effector genes in disparate regions of the genome. Indeed, pleiotropic adaptive mutations in global regulatory genes have often been implicated, both experimentally and in natural settings, in major evolutionary transitions, and have been shown to occur early in evolution (Sumby et al. 2006; Hunter 2008; Wang et al. 2010; Conrad et al. 2011). The evolutionary path to avirulent chronic infection through the *lasR* regulator may be particularly important, because the pattern we observed in the evolved PA14 bears striking resemblance to what is found in *Pseudomonas* isolates of CF patients. Lesions in few core genes, including *lasR*, are commonly found across unrelated patients, whereas many additional genes are potential targets of selection, but only found rarely in any given infection (Smith et al. 2006; Marvig et al. 2014). Moreover, some of the functions associated with *Pseudomonas* adaptation to other hosts, including CF patients, are similar to what we found here. Examples are avirulence, increased growth, altered metabolism, and regulatory network remodeling (Luzar and Montie 1985; Mahenthiralingam et al. 1994; Damkiær et al. 2013; Marvig et al. 2014). All in all, there is ample evidence suggesting that debilitating mutations in *lasR* are pivotal for initial *Pseudomonas* host adaptation, but that further functional alterations are required.

Mutations in *lasR* may be evolutionary stepping stones toward fully host-adapted genotypes in one of two, not necessarily mutually exclusive, ways. First, *lasR* mutation could be followed by compensatory mutations that alleviate initial fitness costs associated with pleiotropic effects on expression levels of nonadaptive genes. Reversely, subsequent mutations may be too costly to evolve in the WT background, such that

lasR mutations are required first to reduce the production of costly virulence factors or to decrease stressing exposure to the host immune system. Which of these two possibilities is the case here remains a matter of speculation, but the high number of synonymous and intergenic polymorphisms found at later time points of our experiment suggests that compensatory evolution may have been substantial. Nevertheless, additional studies are needed to specifically address this intriguing question.

Finally, we showed parallel phenotypic and genomic evolution between *Pseudomonas* adapted to immunocompetent and immunocompromised hosts. This contrasts with the results of a recent study in a very different biological system, where mice-infecting malaria parasites became more virulent to immunocompetent mice after adaptation to immunocompromised hosts (Barclay et al. 2014). Interestingly, malaria parasites also reached higher population densities in immunocompromised relative to immunocompetent hosts, again opposite to our findings. We found that bacteria adapting to immunocompromised hosts were rather similar, even genetically, to those adapted to immunocompetent hosts. However, the former reached increased nucleotide diversities earlier in evolution that the ones adapting to the immunocompetent hosts, suggesting that immunocompromise accelerated evolution of multiple genotypes, but did not affect its outcome. Our study thus illustrates that bacterial adaptation to complex interactions with immunocompromised hosts is difficult to predict. Therefore, more functional-genetic and field studies of incipient pathogenic interactions, potentially combined with theoretical modeling, are needed to improve our understanding of factors that may predict whether pathogens take the pathogenic or commensal evolutionary route after initial contact with novel host populations.

Conclusion

We conclude that when maladapted pathogens rely on the host for resources, their close association may select for low-damage but persistent coexistence rather than overexploitation. Initially pathogenic bacteria may thus quickly evolve an attenuated, commensal phenotype causing chronic infections as a by-product of selfish fitness benefits (Sachs et al. 2004). In our experiment, the evolutionary transition from pathogenicity to commensalism was initiated by mutations in the global regulator *lasR* and in the polymerase gene *rpoB* that evolved in parallel across multiple biologically independent replicates. Global regulators are a likely target for fast evolutionary changes because few mutations may result in complex, large-scale phenotypic changes. As such, they may be stepping stones that initiate adaptation to novel host environments.

Materials and Methods

Interaction Partners

All experiments were conducted using the standard N2 laboratory strain of the nematode *C. elegans*, and are ideally suited for analysis of host–pathogen evolution experiments (Schulte et al. 2010; Morran et al. 2011; Masri et al. 2013). As

immunocompromised hosts in our experiments we chose strains carrying loss-of-function mutations in the ILR and the p38 MAPK immune signaling pathways, both in the N2 background. The ILR signaling pathway controls innate immunity, stress response, and longevity through the transcription factor *daf-16*, a FOXO homolog that also regulates expression of antimicrobial factors such as Lys-7, saposins and thaumatin homologs. ILR is therefore directly involved in host defense against pathogen infection. Knockouts of the *daf-2* gene (a protein upstream from *daf-16*) in the ILR pathway are long-lived and resist infection to Gram-negative and Gram-positive pathogens(Garsin et al. 2003). Moreover, infection of *C. elegans* with *P. aeruginosa* leads to upregulation of ILR-controlled genes, such as saposins and lysins (Kurz and Tan 2004). The strain DR26 (*daf-16*(m26)I) carries a nonsynonymous mutation in *daf-16* and is therefore more susceptible to infection with *P. aeruginosa* than the immunocompetent N2 strain.

The p38 MAPK pathway is an evolutionary conserved pathway that constitutes the major and indispensable defense against pathogen infection both in the gut and on the epidermis, especially to PA14 (Troemel et al. 2006; Schulenburg et al. 2008; Engelmann and Pujol 2010; Cohen and Troemel 2015). The cascade controls expression of both antimicrobial factors such as C-type lectins and ShK toxins and detoxifying enzymes such as glucosyl-S-transferases (involved in neutralizing reactive oxygen species [ROS]) (Troemel et al. 2006). It has been shown that *C. elegans* mutants carrying knockouts in the *nsy-1* or *sek-1* genes are hypersensitive to *Pseudomonas* infection (Kim et al. 2002), making it a prime candidate for our experiments. We chose the VC390 (*nsy-1*(ok593)II) strain, which has a truncated, nonfunctional *nsy-1* that normally codes for a MAP kinase kinase kinase orthologous to the mammalian ASK family kinases involved in, for example, chemotaxis and pathogen response.

The mutant strains were obtained from the *C. elegans* Genetics Centre (CGC), which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Stocks were maintained in the lab under standard conditions with *Escherichia coli* OP50 as food source. As pathogen we chose the Gram-negative *P. aeruginosa* PA14. We performed pilot studies to confirm that both strains are more susceptible to infection and die earlier than the N2 WT Bristol lab strain (see supplementary fig. S1, Supplementary Material online).

Caenorhabditis elegans are bacterivorous nematodes and can be easily infected with PA14 through the oral route. We chose our conditions so that the WT worms were killed within 4 days, the duration of our infection cycle. For each host strain (N2, and the mutants in *nsy-1* and *daf-16*), 15 independent replicate populations consisting of 300 worms were fed with PA14. The subsequent experiment was performed in a full-factorial, randomized design. The bacteria were serially passaged through nonevolving hosts for 13 infection cycles (fig. 2A). To ensure immunocompromised treatments would not go extinct in the first round before evolution could have happened, the experiment was performed at 20° rather than the usual 25° (Tan et al. 1999).

As such, worms die within 4 days instead of within 2. Plates were stored at constant humidity and at 20 °C without light for 4 days, during which time the bacterivorous worms were feeding on bacteria, thereby getting infected with PA14. The concentration of the inoculate was optimized such that at the end of an infection cycle most bacteria were eaten by the worms (no bacterial lawns were visible), whereas worms were not starving. To further ensure selection against bacteria adapting to the medium or clinging to the exterior of the worms, the worm suspensions were filtered thrice using high-speed centrifugation over a small-pore filter at 2,500 rpm, each time resuspending the infected worms in 400 µl sterile buffer. To allow the bacteria inside the worm bodies to increase in number, the worm pellet was resuspended in phosphate-buffered saline (PBS) containing silica beads (for later homogenization of worm tissue) for 12 h; a small sample was taken from the liquid to ensure no free-living bacteria were transferred. Subsequently, the worms were quickly and briefly frozen at −80 °C to kill the worms and prepare them for extraction of bacteria. After thawing, the worm tissues were homogenized and the resulting suspension plated out onto minimal media supplemented with peptone to grow sufficient bacteria for a new round of infection. After 48 h the bacteria were washed off the plates with 2 ml of PBS. Part of the bacterial suspension was frozen in 86% glycerol at −80 °C for future analysis; the other part was used to start the next round of infection.

Synchronization Protocol

To ensure all worms were in exactly the same stage of development at the start of each round of the evolution experiment or at the start of the phenotypic tests, worms were synchronized. Plates containing proliferating populations containing many gravid hermaphrodites were washed off with sterile buffer and mixed with 1 ml 5 M NaOH, 1 ml bleach, and 3 ml of water. The solution dissolves worm tissue, and sterilizes the eggs. The surviving unhatched eggs were finally seeded onto fresh plates with *E. coli*. After 2–3 days at 20 °C all worms were thus synchronized at the L4 stage.

Phenotypic Assays

At the start of the phenotypic assays, 15 µl of evolved bacteria were taken from −80 °C stocks and plated out on Luria-Bertani (LB) media in a randomized approach. After 48 h of growth, the bacteria were washed off using buffer and aliquots were frozen at −20 °C for subsequent phenotypic tests. All assays were performed under exactly the same conditions as the experiment to prevent unwanted genetic changes in the material. All 15 biological replicate populations of all 3 treatments were analyzed for transfers 6, 9, and 13 and were always encoded during the assays such that experimenters were blind to treatment group allocation. Unless indicated otherwise, all tests were performed on all 15 biological replicates for each analyzed transfer (6, 9, 13) and each treatment (N2-adapted, *daf-16*-adapted, and *nsy-1*-adapted) and fully randomized. To ensure we analyzed phenotypic effects that were only caused by bacterial adaptations (and not by

differences in host genotype), all phenotypic assays described here were based on experiments using the N2 WT strain as host.

Survival Assays

Survival assays measure host survival and life span during exposure to bacteria. Each evolved bacterial population was evaluated against the N2 host and against each immunocompromised strain in a full-factorial, randomized design. Controls consisted of exposures to the laboratory food E. coli OP50 and to the ancestral pathogen in three replicates. The test bacteria were mixed with heat-killed OP50 and plated out. Thirty L4 synchronized worm larvae were added to the plates with dried bacterial lawns, and the number of living and dead worms was scored daily.

Population Growth Assay

To quantify host fitness during exposure to bacteria, five synchronized L4 worms were added to peptone-free agar plates inoculated with bacterial suspensions (PA14 and heat-killed E. coli OP50) and kept at 20 °C for 2 days. The total worm population was washed off and fixed in formaldehyde for subsequent analysis. Three random samples of 10 μl were taken from each sample and the number of worms was counted under a binocular. The mean was calculated after counting each sample thrice, and extrapolated to the total number of worms per tube.

Within-Host Replication

The within-host replication assay approximates bacterial within-host fitness by quantifying the number of bacterial cells extracted from infected host bodies. We repeated this assay several times in its basic form and also using slight variations in setup and in how worms were cleaned from external bacteria. First (fig. 2B), we used the extraction method exactly as at each transfer during the main evolution experiment (this is a measure of the population-level infection load). In total, 300 L4 synchronized worms were pipetted onto a dried bacterial pellet on peptide free medium (PFM) plates. After 4 days, worms were washed off, washed with buffer and filtered twice using high-speed filtration over small-pore filters, and kept overnight in buffer containing silica beads. Twelve hours later the worms were shock-frozen at −80 °C, and tissues were homogenized using a Genogrinder at 1,200 beats/min for 3 min and the resulting suspension plated out in three dilutions onto fresh selective plates (nematode growth medium [NGM] + rifampicin 10 μg/ml). After 2 days, the number of colony forming units (CFU) was counted. Second (fig. 2C), to account for higher host fitness on evolved bacteria, we repeated the experiment by picking exactly 30 worms onto and off the plate. This yielded similar results. Third (fig. 2G), to address concerns about possible adaptation of bacteria to the high-throughput filtration, we repeated the second experiment for fewer replicates, but this time carefully washed the external surface of worms by picking them in and out of sterile saline buffer five times before extracting bacteria. We quantified the number of bacteria present in the final eluate outside of the worms to ensure our extract only consisted of bacteria associated with the worms.

Statistical Analyses

All phenotypic tests were analyzed using standard frequentist techniques in R (such as T-tests, ANOVA [analysis of variance], Kruskal–Wallis, Wilcoxon signed-rank test). Comparisons between treatments across time were analyzed using generalized linear models (GLMs, including logistic regressions) with the phenotypic variable (host or bacterial fitness) as dependent and time and treatment as independent variables. Survival data were analyzed using the Kaplan–Meier approach provided by the survival package in R and the significance of differential survival between treatments was tested using the log-rank test. Figures were created using the ggplot2 package in R.

Genomic Analyses

To identify the genetic changes underlying host adaptation and evolution of commensalism, we sequenced all 15 replicates for all three treatments from time points 6, 9, and 13. To identify changes between the published genome of P. aeruginosa_UCBPP_PA14_uid57977 and our lab strain, we also sequenced our ancestral PA14 strain.

DNA Extraction

Whole-genome DNA extraction was performed directly from bacterial populations that were frozen at −80 °C using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations for Gram-negative bacteria. The 136 samples were sequenced at the Institute for Clinical Microbiology, Kiel University Hospital, using the Illumina HiSeq paired-end sequencing technology (Bentley et al. 2008) with an insert size of 150 bp at 30 × coverage. Due to technical issues 12 of 45 samples of transfer 9 were lost during library preparation (three N2-adapted, three daf-16-adapted, and six nsy-1-adapted populations).

Error Correction and Quality Filtering

The first step in our analysis pipeline was to correct sequencing errors and to remove unreliable reads using Quake 0.3.0 (kmer size 15) (Kelley et al. 2010) and SeqPrep (github.com/jstjohn/SeqPrep).

Mapping

After quality control we mapped the samples to the published P. aeruginosa_UCBPP_PA14_uid57977 reference genome (Lee et al. 2006), which was downloaded from the NCBI ftp server (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/

Pseudomonas_aeruginosa_UCBPP_PA14_uid57977, last accessed July 28, 2015). First, the reference genome was saved in the fasta format and indexed using bwa. Mapping was then performed using the paired-end bwa sample module (Li and Durbin 2010). The resulting sam files were sorted according to genome position, reads not mapped into a proper pair or aligned with a mapping quality below 20 were

filtered out using samtools (Li et al. 2009). Finally, the sam files were converted to (compressed) bam files, indexed using samtools, and visually inspected for low-quality areas with a locally installed version of Integrated Genome Viewer (IGV, Broad Institute; www.broadinstitute.org/software/igv/, last accessed July 28, 2015). Using the samtools flagstat option, basic statistics were calculated and per-base coverage was counted using genomeCoverageBed within the Bedtools package (Quinlan and Hall 2010). Coverage was further analyzed in R and Bioconductor using the HilbertVis and ShortRead packages.

Variant Detection

To reliably call SNPs and SV, duplicate reads were removed using MarDuplicates in Picardtools (http://picard.source-forge.net, last accessed July 28, 2015). SNPs and small insertions or deletions (indels) were then detected using a frequentist approach that finds polymorphisms above a user-determined minimal allele frequency. Using SNVer (Wei et al. 2011), we called SNPs above a threshold frequency of 0.05 and only when they were present on more than two reads. We also called SNPs using a heuristic method that infers polymorphisms based on read depth (min. 50 reads), base quality (min. 15), and variant frequency (min 0.05) in VarScan 2 (Koboldt et al. 2012). For the identification of larger indels and other SV, we used Pindel (Ye et al. 2009) to exploit paired-end information to find regions where only one of the pair could be mapped to the reference. Pindel attempts to break the unmapped read into two and map the resulting shorter fragments to the reference again. This way, it can detect the breakpoints of indels. Further, CNVs were detected using a statistical analysis of binned coverage information in CNVnator (Abyzov et al. 2011). We optimized the bin size to 100 bp.

The pipeline was encoded into serial bash scripts and run on the high performance computing cluster (rzcluster) of the University of Kiel. All of the resulting output files were filtered for duplicates and for ancestral variants, collated using newly developed Perl and bash scripts and annotated using snpEff (Cingolani et al. 2012). The counts of variants and genes carrying variants were quantified and statistically summarized in R.

Population Genomics

We further analyzed population genetic parameters for each of the 125 sequenced bacterial samples, including Tajima's D (TD), nucleotide diversity π, and Watson's θ using the software Popoolation (Kofler et al. 2011). We used sliding windows of size 5 and 10 kb with a 50% overlap.

The estimates for TD and π were then normalized for each genome, corrected for nucleotide diversity of the ancestral clone. Regions with a significantly different π or TD among treatments and/or transfers were identified using Kruskal–Wallis signed-rank tests with FDR corrections for multiple testing. The resulting regions with significant P values are provided in supplementary tables S3 and S4, Supplementary Material online.

We additionally performed multiple regression analyses between phenotypic results and genomic estimates. The

phenotypes under consideration were bacterial fitness, host fitness, and the LT50 (the average time point where 50% of the population has died within a treatment, calculated from fitted Kaplan–Meier survival models). As genomic variables we used per sample number of variants, per sample number of SNPs, per sample number of affected genes. We also performed the same analyses in a sliding window approach, where the model attempted to correlate phenotypes and estimates of π for genomic regions in a 5-kb window with 50% overlap. We performed sets of GLMs using one phenotype as dependent and π values and remaining phenotypes as dependent values and the treatments and/or time points as grouping variables. Model searching was performed using both step-up (building more complex model from the simplest model) and step-down (simplifying the most inclusive model) approaches, using the Akaike Information Criterion for model choice. We repeated the analysis with and without allowing interaction effects. We could not find strong correlations between pheno- and genotypes, probably because most genomes carried distinct sets of variants that could result in similar phenotypes and phenotypes are the result of epistatic mutations and/or because the *lasR* mutation was responsible for most phenotypic changes which was similar across treatments (and would thus not be identified by comparisons between treatments).

Functional Term Analysis

The final list of genes for each sample was collated for each treatment group and for each time point and the gene lists were subjected to gene enrichment analysis using DAVID 6.7 (http://david.abcc.ncifcrf.gov/, last accessed July 28, 2015) (Huang et al. 2008, 2009). Gene lists were first compared with the background of the *P. aeruginosa* _UCBPP_ PA14_uid57977 genome using DAVID. Functional annotation clustering provides the user with the ability to focus on co-occurrence of functional annotations of genes rather than on genes themselves, and therefore allows an analysis of biologically meaningful functional modules. As annotations we used a combination of GO terms, protein–protein interactions, protein functional domains, disease associations, bio-pathways, sequence features, homology, gene functional summaries, gene tissue expression, and manual annotations from the literature encoded in DAVID. DAVID calculates the significance of each enrichment term, including corrections for multiple testing, using Fisher's exact tests and provides an overall enrichment score for each group. The higher the latter value, the stronger the statistical support that that functional group is enriched in the user's gene list.

We also used EASE, the free, stand-alone software package from DAVID bioinformatics resources (http://david.abcc.ncifcrf.gov/), to identify significant overrepresentation between different gene lists and to find the potential functional biological processes involved in our data sets. EASE allows customized gene lists analysis and can assess the overlap between two different gene lists in term of adjusted P values (rather than against the whole genome as a background, as is the case in DAVID). A gene set database was constructed

from published data (NCBI Gene Expression Omnibus and EMBL ArrayExpress). We also manually compiled a list of genes that have been functionally linked to virulence in the literature. Based on this data set, we performed the EASE analysis and selected the results with a Bonferroni-adjusted P value below 0.05.

Functional Analysis of a *lasR* Mutant

Frame shift indels in *lasR*, or, in a few replicates, in *rhlR*, were found in the majority of replicates across all treatments. Because the phenotypic results were similar across the treatments, *lasR* may be crucial for initial adaptation of PA14 populations to the *C. elegans* host environment. This operon is the central quorum sensing locus and controls virulence and biofilm formation in *P. aeruginosa*. A PA14 strain carrying a transposon knockout in *lasR* has previously been implicated in attenuation of virulence in a forward genetic screen (Feinbaum et al. 2012). We therefore obtained a PA14 *lasR* mutant and performed the survival, infection load, and host fitness assays using this bacterial mutant, directly comparing it with the ancestral PA14.

Supplementary Material

Supplementary figure S1 and tables S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

References

Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* 21:974–984.

Anderson RM, May RM. 1982. Coevolution of hosts and parasites. *Parasitology* 85:411–426.

Barclay VC, Kennedy DA, Weaver VC, Sim D, Lloyd-Smith James O, Read AF. 2014. The effect of immunodeficiency on the evolution of virulence: an experimental test with the rodent malaria *Plasmodium chabaudi. Am Nat.* 184:S47–S57.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53–59.

Cash HL, Whitham CV, Behrendt CL, Hooper LV. 2006. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* 313:1126–1130.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* 6:80–92.

Cohen LB, Troemel ER. 2015. Microbial pathogenesis and host defense in the nematode *C. elegans. Curr Opin Microbiol.* 23:94–101.

Conrad TM, Lewis NE, Palsson BØ. 2011. Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol.* 7:509.

Damkiær S, Yang L, Molin S, Jelsbak L. 2013. Evolutionary remodeling of global regulatory networks during long-term bacterial adaptation to human hosts. *Proc Natl Acad Sci U S A.* 110:7766–7771.

D'Argenio DA, Wu M, Hoffman LR, Kulasekara HD, Déziel E, Smith EE, Nguyen H, Ernst RK, Larson Freeman TJ, Spencer DH, et al. 2007. Growth phenotypes of *Pseudomonas aeruginosa* lasR mutants adapted to the airways of cystic fibrosis patients. *Mol Microbiol.* 64:512–533.

Dethlefsen L, McFall-Ngai M, Relman DA. 2007. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* 449:811–818.

Douglas AE. 2014. Symbiosis as a general principle in eukaryotic evolution. *Cold Spring Harb Perspect Biol.* 6:a016113.

Ebert D. 1998. Experimental evolution of parasites. *Science* 282:1432–1436.

Engelmann I, Pujol N. 2010. Innate immunity in *C. elegans. Adv Exp Med Biol.* 708:105–121.

Evans EA, Kawli T, Tan M-W. 2008. *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans. PLoS Pathog.* 4:e1000175.

Feinbaum RL, Urbach JM, Liberati NT, Djonovic S, Adonizio A, Carvunis A-R, Ausubel FM. 2012. Genome-wide identification of *Pseudomonas aeruginosa* virulence-related genes using a *Caenorhabditis elegans* infection model. *PLoS Pathog.* 8:e1002813.

Fry AJ, Palmer MR, Rand DM. 2004. Variable fitness effects of Wolbachia infection in *Drosophila melanogaster. Heredity* 93:379–389.

Garsin DA, Villanueva JM, Begun J, Kim DH, Sifri CD, Calderwood SB, Ruvkun G, Ausubel FM. 2003. Long-lived *C. elegans* daf-2 mutants are resistant to bacterial pathogens. *Science* 300:1921–1921.

Gilbert SF, Sapp J, Tauber AI. 2012. A symbiotic view of life: we have never been individuals. *Q Rev Biol.* 87:325–341.

Gummesson B, Magnusson LU, Lovmar M, Kvint K, Persson Ö, Ballesteros M, Farewell A, Nyström T. 2009. Increased RNA polymerase availability directs resources towards growth at the expense of maintenance. *EMBO J.* 28:2209–2219.

Harcombe W. 2010. Novel cooperation experimentally evolved between species. *Evolution* 64:2166–2172.

Herre EA, Jandér KC, Machado CA. 2008. Evolutionary ecology of figs and their associates: recent progress and outstanding puzzles. *Annu Rev Ecol Evol Syst.* 39:439–458.

Hillesland KL, Stahl DA. 2010. Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc Natl Acad Sci U S A.* 107:2124–2129.

Hoeven RVD, McCallum KC, Cruz MR, Garsin DA. 2011. Ce-Duox1/BLI-3 generated reactive oxygen species trigger protective SKN-1 activity

via p38 MAPK signaling during infection in *C. elegans*. *PLoS Pathog.* 7:e1002453.

Hoffman LR, Kulasekara HD, Emerson J, Houston LS, Burns JL, Ramsey BW, Miller SI. 2009. *Pseudomonas aeruginosa* lasR mutants are associated with cystic fibrosis lung disease progression. *J Cyst Fibros.* 8:66–70.

Hom EFY, Murray AW. 2014. Niche engineering demonstrates a latent capacity for fungal-algal mutualism. *Science* 345:94–98.

Huang DW, Sherman BT, Lempicki RA. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 4:44–57.

Huang DW, Sherman BT, Lempicki RA. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37:1–13.

Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–214.

Hunter P. 2008. The great leap forward. Major evolutionary jumps might be caused by changes in gene regulation rather than the emergence of new genes. *EMBO Rep.* 9:608–611.

Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012. Experimental evolution. *Trends Ecol Evol.* 27:547–560.

Kelley DR, Schatz MC, Salzberg SL. 2010. Quake: quality-aware detection and correction of sequencing errors. *Genome Biol.* 11:R116.

Kerr PJ, Ghedin E, DePasse JV, Fitch A, Cattadori IM, Hudson PJ, Tscharke DC, Read AF, Holmes EC. 2012. Evolutionary history and attenuation of myxoma virus on two continents. *PLoS Pathog* 8:e1002950.

Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, Inoue H, Tanaka-Hino M, Hisamoto N, Matsumoto K, Tan M-W, et al. 2002. A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* 297:623–626.

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22:568–576.

Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C, Schlötterer C. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One* 6:e15925.

Kurz CL, Tan M-W. 2004. Regulation of aging and innate immunity in *C. elegans*. *Aging Cell* 3:185–193.

Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL, Miyata S, Diggins LT, He J, Saucier M, Déziel E, et al. 2006. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol.* 7:R90.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26:589–595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Luzar MA, Montie TC. 1985. Avirulence and altered physiological properties of cystic fibrosis strains of *Pseudomonas aeruginosa*. *Infect Immun.* 50:572–576.

Machado CA, Jousselin E, Kjellberg F, Compton SG, Herre EA. 2001. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. *Proc R Soc Lond B Biol Sci.* 268:685–694.

Mahenthiralingam E, Campbell ME, Speert DP. 1994. Nonmotility and phagocytic resistance of *Pseudomonas aeruginosa* isolates from chronically colonized patients with cystic fibrosis. *Infect Immun.* 62:596–605.

Marvig RL, Sommer LM, Molin S, Johansen HK. 2014. Convergent evolution and adaptation of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nat Genet.* 47:57–64.

Masri L, Branca A, Sheppard A, Papkou A, Laehnemann D, Guenther PS, Prahl S, Saebelfeld M, Hollensteiner J, Liesegang H, et al. 2015. Host-pathogen coevolution: the selective advantage of *Bacillus thuringiensis* virulence and its Cry toxin genes. *PLoS Biol.* 13:e1002169.

Masri L, Schulte RD, Timmermeyer N, Thanisch S, Crummenerl LL, Jansen G, Michiels NK, Schulenburg H. 2013. Sex differences in

host defence interfere with parasite-mediated selection for outcrossing during host–parasite coevolution. *Ecol Lett.* 16:461–468.

McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF, et al. 2013. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci U S A.* 110:3229–3236.

McGraw EA, Merritt DJ, Droller JN, O'Neill SL. 2002. Wolbachia density and virulence attenuation after transfer into a novel host. *Proc Natl Acad Sci U S A.* 99:2918–2923.

Morran LT, Schmidt OG, Gelarden IA, Parrish RC, Lively CM. 2011. Running with the Red Queen: host-parasite coevolution selects for biparental sex. *Science* 333:216–218.

Mueller UG, Gerardo NM, Aanen DK, Six DL, Schultz TR. 2005. The evolution of agriculture in insects. *Annu Rev Ecol Evol Syst.* 36:563–595.

Papp D, Csermely P, Sőti C. 2012. A role for SKN-1/Nrf in pathogen resistance and immunosenescence in *Caenorhabditis elegans*. *PLoS Pathog.* 8:e1002673.

Pukkila-Worley R, Ausubel FM. 2012. Immune defense mechanisms in the *Caenorhabditis elegans* intestinal epithelium. *Curr Opin Immunol.* 24:3–9.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.

Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R. 2004. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118:229–241.

Sachs JL, Mueller UG, Wilcox TP, Bull JJ. 2004. The evolution of cooperation. *Q Rev Biol.* 79:135–160.

Schulenburg H, Hoeppner MP, Weiner J, Bornberg-Bauer E. 2008. Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* 213:237–250.

Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H. 2010. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proc Natl Acad Sci U S A.* 107:7359–7364.

Schultz TR, Brady SG. 2008. Major evolutionary transitions in ant agriculture. *Proc Natl Acad Sci U S A.* 105:5435–5440.

Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, et al. 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A.* 103:8487–8492.

Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* 2:e5.

Tan M-W, Rahme LG, Sternberg JA, Tompkins RG, Ausubel FM. 1999. *Pseudomonas aeruginosa* killing of *Caenorhabditis elegans* used to identify *P. aeruginosa* virulence factors. *Proc Natl Acad Sci U S A.* 96:2408–2413.

Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH. 2006. p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* 2:e183.

Wang L, Spira B, Zhou Z, Feng L, Maharjan RP, Li X, Li F, McKenzie C, Reeves PR, Ferenci T. 2010. Divergence involving global regulatory gene mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome Biol Evol.* 2:478–487.

Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H. 2011. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res.* 39:e132.

Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25:2865–2871.

Zhou YN, Jin DJ. 1998. The rpoB mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 95:2908–2913.