ARTICLE

doi:10.1038/nature10172

Genomic island variability facilitates *Prochlorococcus*–virus coexistence

Sarit Avrani[1], Omri Wurtzel[2], Itai Sharon[1]†, Rotem Sorek[2] & Debbie Lindell[1]

Prochlorococcus **cyanobacteria are extremely abundant in the oceans, as are the viruses that infect them. How hosts and viruses coexist in nature remains unclear, although the presence of both susceptible and resistant cells may allow this coexistence. Combined whole-genome sequencing and PCR screening technology now enables us to investigate the effect of resistance on genome evolution and the genomic mechanisms behind the long-term coexistence of *Prochlorococcus* and their viruses. Here we present a genome analysis of 77 substrains selected for resistance to ten viruses, revealing mutations primarily in non-conserved, horizontally transferred genes that localize to a single hypervariable genomic island. Mutations affected viral attachment to the cell surface and imposed a fitness cost to the host, manifested by significantly lower growth rates or a previously unknown mechanism of more rapid infection by other viruses. The mutant genes are generally uncommon in nature yet some carry polymorphisms matching those found experimentally. These data are empirical evidence indicating that viral-attachment genes are preferentially located in genomic islands and that viruses are a selective pressure enhancing the diversity of both island genes and island gene content. This diversity emerges as a genomic mechanism that reduces the effective host population size for infection by a given virus, thus facilitating long-term coexistence between viruses and their hosts in nature.**

Cyanobacteria of the genus *Prochlorococcus* are dominant photosynthetic organisms in the oceans, contributing significantly to global primary production[1]. They are most abundant in oligotrophic waters and have predictable and reproducible distribution patterns over time and space[2,3]. Two different high-light-adapted ecotypes[4], HLI and HLII, span most of the surface oceans and differ from each other in their geographic distribution[5,6].

High-light-adapted *Prochlorococcus* ecotypes are infected by podoviruses with narrow host ranges and myoviruses with comparatively broader host ranges[7]; such viruses, or phages, are common in oceanic waters[8]. It is unclear, however, how the abundant cyanobacteria coexist with their viruses. Theory predicts that viruses should reduce the population size to levels approximately 20-fold lower than actual *Prochlorococcus* concentrations[9], which, for a single ecotype, can be as high as 200,000 cells ml^{-1} (refs 5, 6). A number of hypotheses have been put forward to explain their coexistence[10]. The 'continuous arms race' hypothesis, whereby resistant bacteria emerge followed by viruses with altered host ranges, is often favoured[11,12]. Indeed the presence of predominantly resistant cells, and a small number of susceptible cells that maintain a viable viral population, has been proposed to explain long-term coexistence between cyanobacteria and their viruses in the oceans[13].

Genomes of closely related organisms (belonging to the same species or ecotype) generally contain large syntenic regions, where conserved, core genes—orthologous genes common to all members[14]—are organized in the same order along the genome. These regions are often interrupted at discrete locations by genomic islands—large regions (more than eight kilobases long) of non-conserved, non-core genes that are sporadically distributed among members of the population[15–17]. Genomic islands contain many horizontally transferred genes[17,18], and therefore increase intraspecies variability. They are present in a wide range of bacterial phyla[15–17], including cyanobacteria[18–21]. They often encode cell-surface genes, suggesting that they have a role in grazer and phage avoidance[18–22].

In this study, we set out to understand the mechanisms that allow the long-term coexistence between *Prochlorococcus* and its viruses and the role genomic islands have in this coexistence. Initially we looked for common themes of resistance in a single host and then extended our study to investigate general features of resistance in both HLI and HLII *Prochlorococcus* strains, selected for resistance to ten podoviruses (Supplementary Table 1). In total, we investigated the genetic basis for resistance in 77 mutant substrains, 27 of which were fully sequenced. We further characterized the phenotype of representative mutants with respect to the stage of infection impaired and the adaptive cost associated with resistance.

Genotype of resistance mutants

To characterize the interactions between *Prochlorococcus* MED4 (hereafter MED4) and five podoviruses, we isolated 43 resistant substrains derived from isogenic colonies and 19 from heterogenic populations (Supplementary Methods). To identify resistance-conferring mutations, the genomes of 12 mutants from isogenic colonies were sequenced and compared with the sequence of susceptible controls. Mutations in the remaining substrains were detected by PCR screening.

All substrains resistant to a set of three viruses (P-SSP7, P-TIP1 and P-TIP2) had mutations in a cluster of six non-core genes within genomic island 4 (ISL4) of the MED4 genome (genes PMM1242 to PMM1249; Fig. 1b, Supplementary Table 2 and Supplementary Fig. 1). Most resistant substrains (80%) had a single mutation, but others had another mutation within this same cluster of genes. The mutations caused amino-acid changes, reading frame shifts or premature stop codons (Supplementary Table 3). Three of the genes are exclusive to MED4 among the cyanobacteria and the other three have homologues among only some *Prochlorococcus* strains (Supplementary Table 2). MED4 substrains resistant to all five podoviruses (P-GSP1, P-TIP38, P-SSP7, P-TIP1 and P-TIP2) had mutations in five different genes (Supplementary Tables 2 and 3); three are situated in ISL4 and a fourth

[1]Faculty of Biology, Technion – Israel Institute of Technology, Haifa 32000, Israel. [2]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel. †Present address: Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA.

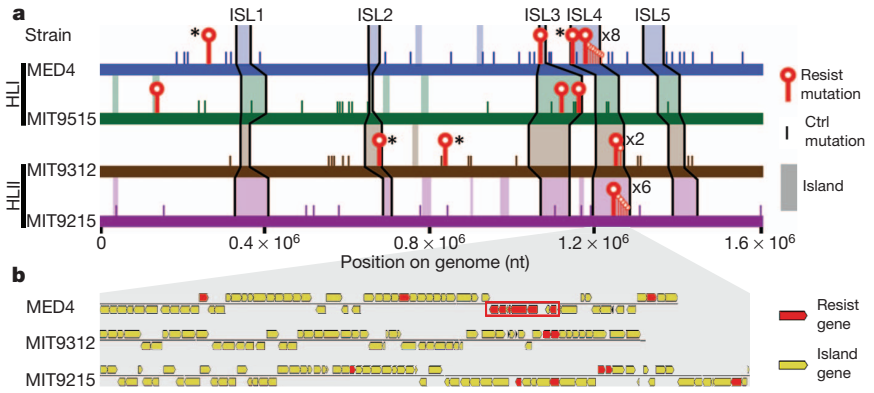

Figure 1 | Distribution of resistance-conferring mutations. **a**, Diagram of high-light-adapted *Prochlorococcus* genomes. Genes with resistance-conferring mutations ('resist') are shown by red lines and circles, and core genes are shown by asterisks. Mutations in controls relative to reference genomes ('ctrl') are shown by short lines. Genomic islands are shaded, and those defined previously[19] are named above the genomes. nt, nucleotides. **b**, Expansion of most of ISL4, the viral susceptibility region. Genes with resistance-conferring mutations are shown in red. The red box surrounds a cluster of six mutant genes between PMM1242 and PMM1249. MIT9515 is not shown as it had no resistance-conferring mutations in this island.

is situated in genomic island 3 (ISL3). Three of these four are non-core genes (PMM1124, PMM1232 and PMM1259), whereas one, in ISL4, is a core gene with homologues in all *Prochlorococcus* strains (PMM1209). By contrast, the fifth mutation is in a core gene (PMM0278) that sits in a syntenic region of the genome (Fig. 1).

Of the 11 MED4 mutant genes identified, four are predicted to be membrane associated and another six are potential cell-wall biosynthesis and modification enzymes (Supplementary Table 2). These include sugar isomerases and methyl-, carbamoyl- and aminotransferases. Therefore, mutations in these genes may cause an alteration to the cell-surface structure of resistant substrains.

Phylogenetic analyses of the non-core genes showed that they do not cluster with *Prochlorococcus* homologues (Fig. 2c and Supplementary Fig. 2), suggesting that they were horizontally acquired from other phyla. Furthermore, the phylogenetic profile of the ISL4 core gene (PMM1209) showed no clear separation between *Prochlorococcus* ecotypes (Fig. 2b), suggesting that there is some degree of gene swapping among such ecotypes. In contrast, the syntenic-region core gene (PMM0278) had a phylogenetic profile congruent with the majority of core genes (Fig. 2a). PMM1209 and the non-core genes had a much higher degree of sequence divergence among *Prochlorococcus* homologues relative to core genes (Supplementary Fig. 3).

To determine whether these findings are a general phenomenon among high-light-adapted *Prochlorococcus* ecotypes, we expanded our study to investigate resistance in 15 additional substrains originating

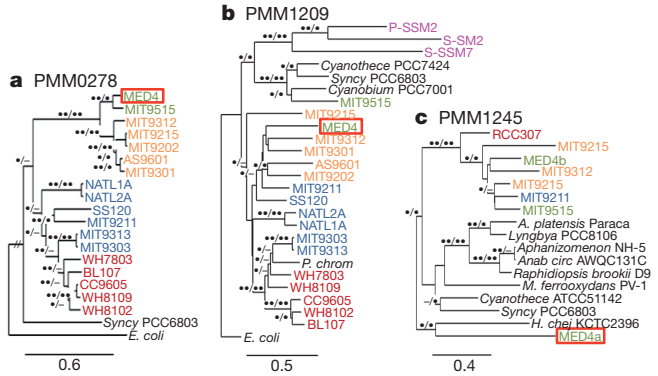

Figure 2 | Phylogeny of representative MED4 mutant genes. Distance trees of PMM0278, a syntenic core gene (**a**); PMM1209, the ISL4 core gene (**b**); and PMM1245, an ISL4 non-core gene (**c**). Mutant genes are enclosed in red boxes. Green, orange and blue indicate HLI, HLII, and low-light-adapted *Prochlorococcus* strains, respectively. Red, black and pink indicate marine *Synechococcus*, other microbes and viruses, respectively. Bootstrap values (single dot, >50%; double dot, >90%) are shown at the nodes (distance/maximum likelihood). *Syncy*, *Synechocystis*; *Anab circ*, *Anabaena circinalis*; *E. coli*, *Escherichia coli*, *P. chrom*, *Paulinella chromatophora*; *H. chej*, *Hahella chejuensis*; *M. ferrooxydans*, *Mariprofundus ferrooxydans*; *A. platensis*, *Arthrospira platensis*. Scale bars show numbers of amino-acid substitutions per site.

from one other HLI strain (MIT9515) or two HLII strains (MIT9312 and MIT9215) that were selected for resistance to five additional podoviruses. Genome sequencing revealed mutations in a total of 13 different genes: 11 non-core and 2 core genes. All but one are in genomic islands; eight in ISL4, two in ISL3 and a core gene (PMT9312_0728) in island 2 (ISL2) (Fig. 1). Phylogenetic analyses of the non-core genes suggest that they have been horizontally transferred from other phyla (Supplementary Fig. 2i–p). Additionally, the ISL2 core gene (PMT9312_0728) had a phylogenetic profile different from those of most core genes (Supplementary Fig. 2g). Many (eight) of the 13 mutant genes are predicted to be cell-surface related and include an integral membrane subunit of the phosphate transporter (*pstA*), lipopolysaccharide and cell-wall biosynthesis genes (Supplementary Table 2).

In this study, myoviruses were not used as a source of selective pressure for the isolation of resistant substrains. However, five of our substrains were resistant to at least one of two myoviruses tested (Supplementary Table 2). These included four with mutations in genomic islands, three of which were in non-core genes. Therefore, mutations in some of the same surface-related island genes confer resistance on both podoviruses and myoviruses. It should be noted that selection for resistant substrains using myoviruses may uncover mutations in additional gene types.

All together, analysis of these 77 resistant strains revealed common features of resistance-conferring mutations among high-light-adapted *Prochlorococcus* ecotypes. The great majority (91%) of the 24 mutant genes are in hypervariable genomic islands. Moreover, 71% localized to a single genomic island, ISL4, indicating that this is a virus susceptibility region. Most mutant genes (83%) are part of the non-core or variable genome and seem to have been horizontally acquired from diverse and distant bacterial phyla, and two of the four core genes seem to have been swapped among members of the *Prochlorococcus* genus. Three of the mutant genes (PMM1209, PMM1246 and PMT9312_1347) also have homologues in the genomes of cyanobacterium phages[23] (Fig. 2b and Supplementary Fig. 2d, j). Most genes are potentially cell-surface related (75%) or neighbour cell-surface-related genes (another 12%). Surprisingly, we did not find a resistance-conferring mutation in the same gene in more than one strain, but we did identify mutations in two homologues from the same strain (Supplementary Fig. 2m). Although we did not observe a direct gene acquisition or loss event, the identification of resistance-conferring mutations in sporadically distributed, horizontally transferred genes strongly suggests that they are dynamically gained by and lost from genomic islands in response to viral selection pressure.

To assess whether non-core island genes are more prone to mutation, we investigated the distribution of mutations in eight sequenced controls relative to published reference genomes of the ancestral strains[18,19,24]. Resistance-conferring mutations were mostly localized to non-core, cell-surface island genes ($P < 10^{-15}$), but no such trend was observed for our control substrains ($P = 0.75$). Control mutations were distributed throughout the genome (Fig. 1) and contained a

mixture of silent, amino-acid-changing and intergenic-region mutations (Supplementary Table 4). These data indicate that our genomic findings are not due to chance but rather are directly associated with resistance.

Impaired attachment

Resistance-conferring mutations in predicted cell-surface genes suggest that they affect virus attachment to the cell surface[10]. To test this hypothesis, we carried out adsorption assays for a subset of resistant substrains. All seven mutants tested showed impaired attachment to the viruses used for selection (Fig. 3). This suggests that the mutant genes encode enzymes involved in the synthesis of viral receptors, co-receptors or molecules that interfere with receptors, or encode the proteins themselves. Preventing attachment and, therefore, entry into the cell is potentially the most effective line of defence, and is a common mode of bacterial resistance to lytic phages[10,25].

Attachment of podoviruses to strain-specific cell-surface components explains their high degree of host specificity; they often infect a single host[7,13]. Mutations in some of these genes also provided resistance to myoviruses with broader host ranges. This suggests that there is differential interference for viral attachment between strains or that myoviruses use different receptors for distinct hosts, a known phenomenon from other systems[10]. Thus, the host range of viruses is determined by the dynamic, exchangeable part of the genome—genomic islands.

Adaptive cost of resistance

Viral infection causes mortality of susceptible cells and selects for populations resistant to infection. Long-term coexistence between host and virus requires, however, a population of susceptible cells supporting viral production[10,26]. This apparent paradox can be reconciled if a cost of resistance exists, whereby susceptible subpopulations out-compete resistant subpopulations under certain environmental conditions[10]. To assess whether resistant substrains had a fitness cost, we compared their growth rates with those of susceptible controls under optimal laboratory conditions. Eleven of the 23 mutants tested grew significantly (up to 50%) more slowly than controls (Fig. 4a and Supplementary Fig. 4). These included all four substrains with core gene mutations and seven out of 19 substrains with mutations in non-core genes. Importantly, these results were consistent for independently isolated substrains with mutations in the same gene (Supplementary Fig. 5). It should be noted that a growth cost of resistance may exist for additional substrains under conditions not tested here. Nonetheless, these data suggest that mutations in core genes are more likely to cause a fitness cost than those in non-core genes, but also highlight the fact that a number of non-core genes have a fundamental role in the physiology of the organism and that their absence would be detrimental.

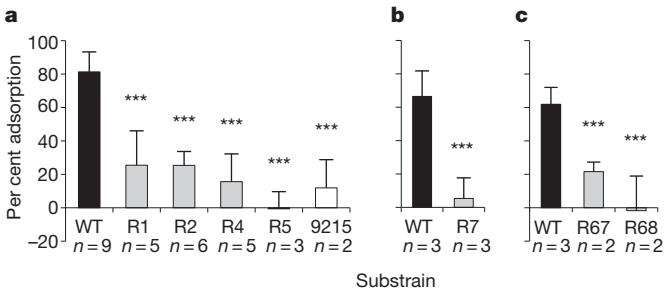

Figure 3 | Attachment of podoviruses to resistant substrains. Adsorption of P-SSP7 (**a**) and P-GSP1 (**b**) to MED4, and P-SSP2 (**c**) to MIT9312. Resistant substrains are indicated by grey bars. Susceptible control substrains (black bars; WT, wild type) served as positive controls for attachment, whereas a non-host strain (white bar) is a negative control for non-specific attachment. Per cent adsorption was determined from the amount of free phage in extracellular medium 4–6 h after phage addition, relative to the amount at 0 h. Data shown are average and s.d. of *n* biological replicates; ***$P < 0.001$, indicating significantly less adsorption than susceptible controls.

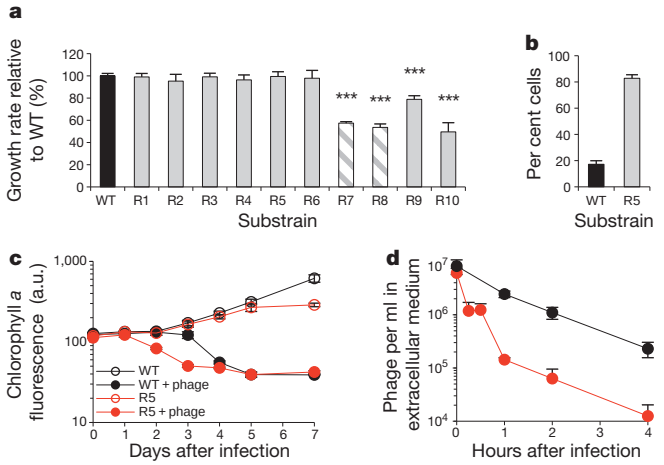

Figure 4 | Cost of resistance. **a**, Growth rates of MED4 resistant substrains with mutations in non-core (grey) and core (hatched) genes, relative to susceptible controls (black). ***$P < 0.001$, indicating significantly slower growth than controls. **b**, Competition between MED4 R5 and susceptible control grown together. **c**, **d**, Infection dynamics (**c**) and adsorption kinetics (**d**) of P-GSP1 infecting MED4 R5 (red) and its control (black). Filled circles, infected cultures; open circles, non-infected cultures. The two substrains responded differently with respect to time ($P < 0.001$). Data shown are average and s.d. of three to six biological replicates. MED4 R5 is resistant to P-SSP7, P-TIP1 and P-TIP2. a.u., arbitrary units.

Growth rate costs are sometimes manifested only under direct competition. However competition experiments between a resistant substrain (R5) and its susceptible control did not reveal a growth cost to resistance either (Fig. 4b).

Another potential, yet undocumented, type of cost of resistance is a greater degree of susceptibility to other viruses. To investigate the possibility of such a susceptibility existing, we challenged resistant substrains with viruses to which they had remained susceptible. A significantly more rapid decimation of mutant populations relative to control populations was observed for five substrains (including R5) infected by a subset of podoviruses and myoviruses (Fig. 4c and Supplementary Fig. 6a). Another three substrains had slower infection dynamics (Supplementary Fig. 6b). Further investigation of R5 revealed a drastic increase in the rate of attachment by two podoviruses still capable of infecting it (Fig. 4d and Supplementary Fig. 6c). These data show that certain mutations confer resistance to some phages yet allow more rapid infection by other phages, representing a novel 'enhanced infection dynamics' fitness cost. This phenomenon may also exist in other host–virus systems where a mutation differentially affects attachment to the host by multiple viruses. Indeed, enhanced attachment of one myovirus was reported for a *Synechococcus* strain resistant to a different myovirus[25], although it is unknown whether this led to enhanced infection dynamics.

Our results indicate that at least 16 out of 23 mutant substrains showed one of two types of adaptive cost of resistance: reduced growth rate or more rapid infection by other viruses.

Coexistence in simple populations

The presence of resistant cells in a population is dependent, in part, on the rate of generation of resistance mutants. Using a modification of the fluctuation test[27], we found a rate of spontaneous mutation leading to resistance in MED4 of between 4.7×10^{-6} and 6.08×10^{-6} per cell per division (Supplementary Table 5), which is approximately 10–1,000-fold higher than for other bacteria[27–29]. Although the mechanism behind this high rate of resistance is unknown, this finding suggests that highly diverse, resistant *Prochlorococcus* populations probably exist in the environment.

The degree to which resistant cells are maintained in a population depends on competition with susceptible cells. We found that heterogenic

populations of MED4 (grown in the absence of viruses for over 600 generations) maintained a majority of susceptible cells. However, a relatively large fraction of resistant cells (0.01–0.2%), both 'normal growing' and slow growing, were also present (Supplementary Fig. 7). This indicates that, contrary to our expectations, slower-growing resistant cells were not competitively excluded. Furthermore, investigation of a resistant substrain (R5) that was maintaining a viable viral population revealed one susceptible colony out of 163, which had reverted to the control genotype. Therefore, susceptible and resistant cells co-occur in simple laboratory populations of *Prochlorococcus*, both in the presence and absence of viruses.

Susceptibility regions in nature

To gain insight into potential viral susceptibility and resistance in field populations of *Prochlorococcus*, we investigated the diversity of homologues of the experimentally identified resistance-related genes in environmental genomic data sets[30,31]. Overall, these genes are highly diverse (Supplementary Fig. 3b), carrying a variety of polymorphisms at resistance-specific positions (Fig. 5a). Some even encode the same polymorphisms as our resistance mutants. Other positions are highly conserved, often carrying the wild-type susceptible polymorphism. Although it is clearly impossible to discern whether environmental populations encoding these sequences are resistant or susceptible to a particular virus, these data are highly suggestive that both resistant and susceptible subpopulations are present in nature.

Another striking feature arising from environmental sequence analyses is the sporadic distribution of resistance-related genes and the scrambling of their order in genomic islands. As is the case for island genes in general[18,19,22], resistance-related island genes were considerably less abundant than syntenic genes (Supplementary Fig. 8). In addition, ISL4 genes do not have a conserved order in environmental populations, as seen from the high number of different genes adjacent to resistance-related genes (Fig. 5b). Furthermore, most genes on these environmental genome segments have homologues in cultured *Prochlorococcus* strains that map to a different region of ISL4 (Fig. 5c). These data are consistent with numerous transfers of these genes within the population at this genomic site[18,19].

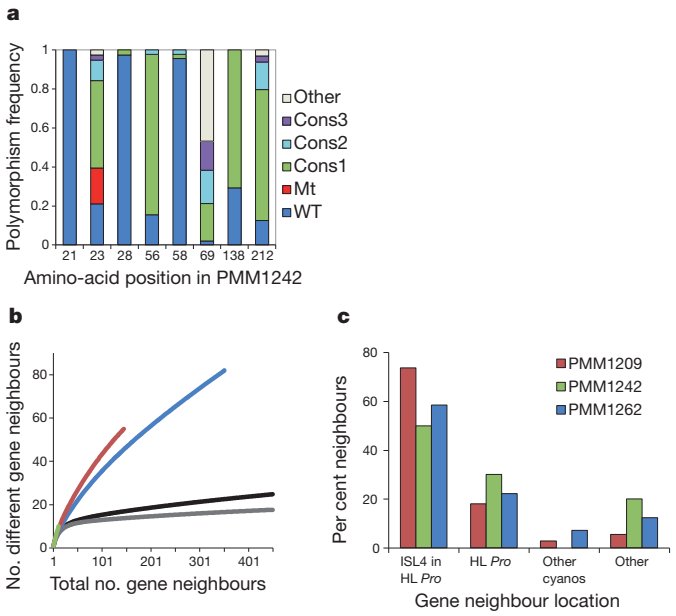

Figure 5 | Resistance-conferring genes in the environment. **a**, Amino-acid frequency at resistance-specific positions in environmental homologues of MED4 gene PMM1242. WT, amino acids identical to controls; Mt, mutant amino acids; Cons1, Cons2 and Cons3, different but conserved amino acids; Other, non-conserved amino acids. Between 37 and 64 sequences were analysed per position. **b**, Number of different neighbouring genes on environmental clones with homologues of target genes. PMM1209 (red) and PMM1242 (green) are within ISL4 and PMM1262 (blue) flanks ISL4. PMM1291 (grey) and PMM1309 (black) are controls for syntenic regions. **c**, Location frequency within sequenced genomes of genes on environmental clones that have homologues of PMM1209, PMM1242 and PMM1262. HL *Pro*, high-light-adapted *Prochlorococcus*; cyanos, cyanobacteria.

Genome evolution and coexistence

The architecture of microbial genomes, with large regions of gene synteny punctuated by hypervariable genomic islands, is a curious phenomenon that seems to have been strongly affected by interactions with viruses. It is well accepted that gene acquisition is facilitated by horizontal gene transfer mediated both by phages and by other mobile genetic elements, and that genomic islands are a repository for such horizontally acquired genes[17,20]. Initial genome locations of islands were probably dictated by the position of transfer RNA genes, which are known integration sites for mobile elements[17], with, as seen from multiple repeat elements[17,19], subsequent and repeated gene acquisition and loss occurring at these sites. Our study suggests that the diversity of genomic islands is at least partly the evolutionary outcome of millions of infection–selection cycles between microbes and viruses, with selection against cells carrying proteins that aid viral attachment. Our results further suggest that gene loss, without parallel gain of genes with similar functionality, would be detrimental to the organism as resistance-conferring mutations often led to a significantly reduced growth rate. This impact of phages on the evolution of genomic islands is likely to extend beyond *Prochlorococcus*, as evidenced by a similar degree of island diversity of cell-surface genes in other cyanobacteria[20,21] and microbes[22,32,33]. We speculate that the physical clustering of horizontally transferred genes in genomic islands is beneficial to microbes, as it facilitates constant gene exchange without disrupting the integrity of the rest of the genome.

Coexistence between *Prochlorococcus* and their phages is probably facilitated by the high diversity of attachment genes in genomic islands. Their sporadic distribution and numerous polymorphisms serve to reduce the effective population size for infection by any particular phage. Therefore, abundant *Prochlorococcus* populations belonging to a single ecotype with common physiological and ecological characteristics are actually an assortment of subpopulations with different susceptibilities to co-occurring phages. We propose that the 'arms race' between bacteria and their viruses leads to the emergence of resistant bacteria in both a sequential and accumulative process, resulting in a continuum of cyanobacteria with different but overlapping ranges of viral susceptibility. Phage pressure then reduces the size of dominant subpopulations, leading to low abundances of a suite of resistance and susceptibility types. Then the small proportion of cells susceptible to a particular phage protects the cyanobacterium from infection owing to reduced probabilities of contact, consistent with the 'numerical refuge' hypothesis[10]. These findings are probably not limited to cyanobacteria, as computer simulations and modelling suggest that microdiversity maintains bacterial population diversity at a subspecies level[22]. Thus, large numbers of taxonomically identical organisms, fulfilling the same ecological role, are probably maintained in the environment as a result of microdiversity in phage susceptibility regions.

The well-known growth cost of resistance is inherent to the physiology of the cell and is probably manifested under a variety of environmental conditions. By contrast, the novel 'enhanced infection dynamics' cost of resistance is local in both time and space, being manifested only in the presence of phages capable of more rapid infection. Thus, the latter resistance cost is probably considerably less detrimental than reduced growth rates while still providing a mechanism for avoidance of competitive exclusion of susceptible cells. Furthermore, this resistance cost is less likely to lead to the extinction of resistant types, resulting in a more diverse set of subpopulations in nature.

An enormous degree of community diversity exists for *Prochlorococcus* and other microbes in the oceans, much of which is manifested at the level of genomic microdiversity among closely

related organisms[18–20,22]. Our experimental findings, together with bioinformatic analyses[18,19,21,22], strongly suggest that this microdiversity is driven to a considerable extent by viruses, through both selection pressure and horizontal gene transfer, leading to an assortment of interchangeable genes in microbial genomes that are involved in viral attachment. This, together with mutation-induced changes in infection dynamics, both accelerations and decelerations, has probably led to a suite of distinct bacterial subpopulations. The resultant community-level complexity has probably produced a complex network of interactions whereby infections constantly occur but only for particular segments of the population at any one time or place, leading to the maintenance of robust and reproducible population and community structure[2,3,34] in the oceans.

METHODS SUMMARY

We used four high-light-adapted *Prochlorococcus* strains to isolate 77 substrains resistant to infection by ten different podoviruses. They were isolated on plates from single colonies in the presence of selecting phages, whereas susceptible controls were isolated in the absence of phages. Genome sequencing of 27 resistant and eight control substrains was performed using Illumina technology to identify resistance-specific mutations. The remaining mutations were identified by sequencing PCR amplicons of genes identified by genome sequencing. Mutant genes were classified as cell-surface related on the basis of predicted membrane domains, functional annotation and gene neighbourhood. They were considered to be core genes if they were present in all 13 sequenced *Prochlorococcus* genomes, and were considered to be part of a genomic island if they localized to regions unique to a single *Prochlorococcus* strain. We identified homologues of mutant genes in environmental data sets using a reciprocal best-BLAST-hit-like approach. The abundance, sequence diversity and gene neighbour diversity of each homologue was estimated. The growth rate of resistant substrains was compared with that of paired controls. The rate of population decline of mutant substrains susceptible to other phages was compared with that of susceptible controls. We determined the degree of attachment to the cell surface from the change in the number of free phages in the extracellular medium with time after phage addition. Experimental details, including genomic island designations (Supplementary Table 6) and primers used (Supplementary Table 7), can be found in Supplementary Methods.

Received 10 February; accepted 4 May 2011.

1. Partensky, F., Hess, W. R. & Vaulot, D. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **63,** 106–127 (1999).
2. Malmstrom, R. R. *et al.* Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceans. *ISME J.* **4,** 1252–1264 (2010).
3. Zwirglmaier, K. *et al.* Basin scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ. Microbiol.* **9,** 1278–1290 (2007).
4. Moore, L. R., Rocap, G. & Chisholm, S. W. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393,** 464–467 (1998).
5. Bouman, H. A. *et al.* Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* **312,** 918–921 (2006).
6. Johnson, Z. I. *et al.* Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311,** 1737–1740 (2006).
7. Sullivan, M. B., Waterbury, J. B. & Chisholm, S. W. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus. Nature* **424,** 1047–1051 (2003).
8. Suttle, C. A. Viruses in the sea. *Nature* **437,** 356–361 (2005).
9. Thingstad, T. F. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic ecosystems. *Limnol. Oceanogr.* **45,** 1320–1328 (2000).
10. Bohannan, B. J. M. & Lenski, R. E. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecol. Lett.* **3,** 362–377 (2000).
11. Comeau, A. M. & Krisch, H. M. War is peace – dispatches from the bacterial and phage killing fields. *Curr. Opin. Microbiol.* **8,** 488–494 (2005).
12. Stern, A. & Sorek, R. The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33,** 43–51 (2011).
13. Waterbury, J. B. & Valois, F. W. Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl. Environ. Microbiol.* **59,** 3393–3399 (1993).
14. Charlebois, R. L. & Doolittle, W. F. Computing prokaryotic gene ubiquity: rescuing the core from extinction. *Genome Res.* **14,** 2469–2477 (2004).
15. Dobrindt, U., Hochhut, B., Hentschel, U. & Hacker, J. Genomic islands in pathogenic and environmental microorganisms. *Nature Rev. Microbiol.* **2,** 414–424 (2004).
16. Hsiao, W. W. *et al.* Evidence of a large novel gene pool associated with prokaryotic genomic islands. *PLoS Genet.* **1,** e62 (2005).
17. Langille, M. G., Hsiao, W. W. & Brinkman, F. S. Detecting genomic islands using bioinformatics approaches. *Nature Rev. Microbiol.* **8,** 373–382 (2010).
18. Kettler, G. C. *et al.* Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus. PLoS Genet.* **3,** e231 (2007).
19. Coleman, M. L. *et al.* Genomic islands and the ecology and evolution of *Prochlorococcus. Science* **311,** 1768–1770 (2006).
20. Dufresne, A. *et al.* Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.* **9,** R90 (2008).
21. Palenik, B. *et al.* The genome of a motile marine *Synechococcus. Nature* **424,** 1037–1042 (2003).
22. Rodriguez-Valera, F. *et al.* Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol.* **7,** 828–836 (2009).
23. Sullivan, M. B. *et al.* Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ. Microbiol.* **12,** 3035–3056 (2010).
24. Rocap, G. *et al.* Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424,** 1042–1047 (2003).
25. Stoddard, L. I., Martiny, J. B. H. & Marston, M. F. Selection and characterization of cyanophage resistance in marine *Synechococcus* strains. *Appl. Environ. Microbiol.* **73,** 5516–5522 (2007).
26. Lenski, R. E. & Levin, B. R. Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *Am. Nat.* **125,** 585–602 (1985).
27. Luria, S. E. & Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28,** 491–511 (1943).
28. Lythgoe, A. & Chao, L. Mechnisms of coexistence of a bacteria and a bacteriophage in a spatially homogeneous environment. *Ecol. Lett.* **6,** 326–334 (2003).
29. Somova, A. G. Frequency of phage-resistant mutations and effect of the bacteriophage on the formation of resistant forms of *Vibrio cholera. Bull. Exp. Biol. Med.* **31,** 99–103 (1966).
30. Rusch, D. B. *et al.* The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **5,** e77 (2007).
31. Yooseph, S. *et al.* The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biol.* **5,** e16 (2007).
32. Parsons, Y. N. *et al.* Suppression-subtractive hybridisation reveals variations in gene distribution amongst the *Burkholderia cepacia complex*, including the presence in some strains of a genomic island containing putative polysaccharide production genes. *Arch. Microbiol.* **179,** 214–223 (2003).
33. Walker, C. B. *et al.* Contribution of mobile genetic elements to *Desulfovibrio vulgaris* genome plasticity. *Environ. Microbiol.* **11,** 2244–2252 (2009).
34. Fuhrman, J. A. *et al.* Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl Acad. Sci. USA* **103,** 13104–13109 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Zinser for the bacterial helper strain; I. Pekarsky for help with adsorption assays; I. Izhaki for advice on statistical analyses; O. Beja, Y. Mandel-Gutfreund, K. Kozek, U. Qimron and Lindell lab members for comments on the manuscript; and T. Dagan for coining the term 'susceptibility region'. Genome sequencing was carried out at the genome sequencing units at the Weizmann Institute of Science and the Technion – Israel Institute of Technology. This work was supported by a European Commission ERC Starting Grant (no. 203406), an ISF Morasha grant (no. 1504/06) and the Technion Russell Berrie Nanotechnology Institute (D.L.); and by an ISF-FIRST grant (no. 1615/09) and an ERC Starting Grant (R.S.). O.W. was supported by an Azrieli fellowship and D.L. is a Shillman fellow.

Author Contributions S.A. and D.L. designed the project and the experiments; S.A. performed and analysed the laboratory experiments, PCR screening and phylogenetic analyses; O.W. performed the bioinformatic analyses and, together with R.S., analysed the genome sequencing data; I.S. analysed the environmental sequences; and D.L. wrote the manuscript with significant contributions from all authors.

Author Information The DNA polymerase and g20 sequences of the phages isolated in this study have been submitted to Genbank under the accession numbers JF837212 to JF837216. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.L. (dlindell@tx.technion.ac.il).