Phase variation and adaptation in bacteria: A 'Red Queen's Race'

R. Jayaraman*

R.H. 35, Palaami Enclave, New Natham Road, Madurai 625 014, India

In nature, bacteria are constantly exposed to many stressful conditions of life. This is particularly true of pathogens. Survival and adaptation under stressful conditions demand multiple strategies, genetic as well as phenotypic. Bacteria have many, pre-programmed, phenotypic stress response systems which can handle a limited number of stresses. Genetically, heritable as well as transient hypermutability mechanisms have been found to facilitate bacterial adaptation to varied and unpredictable stresses; these processes are not reviewed here. Instead, this article will focus on processes which do not increase global mutation rates but cause localized hypermutability in specific loci called contingency genes which have been identified particularly in pathogenic bacteria. These processes are collectively called phase and antigenic variations. Most of the contingency genes are involved in the synthesis or modification of surface-associated structures and enzymes. Phase variation in these genes involves high frequency, reversible, switching of their expression (on to off and off to on). The mechanisms of this switching are reviewed. However, some phase variable genes are not involved in the synthesis or modification of surface structures but are components of type I and type III restriction–modification (RM) systems. The on/off switching of these genes (type III RM genes) leads to regulation of expression of many unlinked genes, impacting several properties of cells. This novel type of control of multiple gene expression by phase variation has been named 'phasevarion'. The adaptive advantages of phase variation in contingency genes and phasevarions in the evasion of host immunity, virulence, niche adaptation and other phenomena are reviewed with some illustrative examples. Phase variation and bacterial adaptation have been likened to the 'Red Queen's Race' in Lewis Carrol's classic *Through the Looking Glass*.

Keywords: Adaptation, contingency genes, phase variation, phasevarions, Red Queen's Race.

THE life of bacteria in nature is plagued by struggle for survival, perpetuation and evolution. Bacteria are constantly subjected to various stresses such as heat, starvation, oxidative stress, DNA damage, exposure to radiation, toxic agents, etc. Phenotypically, bacteria respond to stress by modulation of gene expression through sigma factors, proteolysis and other mechanisms. There are several well-documented mechanisms of stress responses which have been extensively reviewed[1–7]. These pre-programmed response systems enable organisms to sense, respond and adapt to stress. Some of them also result in concomitant mutagenesis which has been categorized as stress-induced mutagenesis[7,8]. However in nature, stresses encountered by bacteria are many in number and variety, and are often unpredictable and unprovoked. In the absence of any pre-determined mechanisms to handle such exigencies, the survival of bacteria is endangered. Obviously, it is not possible to evolve and maintain response mechanisms to each and every conceivable stress. Unforeseen challenges could be best countered by genetic strategies, that is, through mutations and/or horizontal gene transfer. Since spontaneous mutation rates are usually very low, adaptation through spontaneous mutations is difficult to achieve, although not impossible. Elevation of mutation rates (hypermutability) would be a very efficient means to generate genetic diversity and accelerate the speed of adaptation. It is not difficult to understand that when multiple mutations are required for adaptation, hypermutators would be better endowed than normo-mutators since the former can generate many mutations, some of which could be adaptively advantageous. For example, during infection, pathogens face multiple bottlenecks in the form of host defences, antibiotic intervention, inter-host and inter-tissue transmissions, competition by co-infecting organisms and/or commensals, etc. to succeed in establishing successful infection. Survival and growth under such conditions demand multiple mutations and hypermutators could easily meet such demands. It is, therefore, not surprising that clinical isolates of many pathogens such as *Escherichia coli*, *Salmonella* species, *Pseudomonas aeruginosa*, *Haemophilus influenzae*, *Streptococci*, *Staphylococci*, *Neisseriae*, *Helicobacter pylori*, etc. contain high proportions of hypermutators (1–30%)[9–12]. The mechanisms of hypermutability have been reviewed[13]. While hypermutability can be an asset to organisms since it enables easy adaptation, it can also be a liability. A majority of mutations happen to be deleterious and only very few are adaptively beneficial. It has been estimated that in *E. coli*, the mutation rates for deleterious mutations are generally

*e-mail: aarjay007@rediffmail.com

10^{-5}–10^{-4}/cell/generation[14,15], whereas the rates for beneficial mutations are 4–5 orders of magnitude less[16]. Therefore, hypermutability is inevitably associated with an enhanced risk of accumulation of many unwanted and deleterious mutations, especially if the trait persists after adaptation. In spite of such risks, hypermutability is quite common in nature (see above). Apparently, the adaptive benefits of hypermutability outweigh the disadvantages that might accrue due to deleterious mutations. Moreover, when a mutator cell adapts to a stress and grows by acquiring a beneficial (adaptive) mutation, the mutator mutation 'hitch hikes' with the adaptive mutation and is inherited, a phenomenon called indirect or second-order selection. Obviously, hypermutability without risks would be ideal for adaptation. Two processes exist which can accomplish this. One is transient hypermutability which is implicated in the mutagenic processes associated with stress responses. Here, the mutagenic potential does not persist after adaptation and therefore there is no long-term risk. Transient hypermutation (variously called adaptive mutagenesis, stationary phase mutagenesis, stress-induced mutagenesis, post-plating mutagenesis, etc.) will not be discussed here. Several excellent reviews and reports on this topic have appeared in recent years[7,8,17–23]. The other mechanism of risk-free hypermutation is localized hypermutation in which mutagenesis is restricted to certain loci called 'contingency' genes. This review will focus on this process which underlies the phenomena called phase and antigenic variations.

Contingency genes and phase variation

Phase variation was first recognized in strains of *Salmonella* in the early 1920s by the discovery of switching between two flagellar antigenic types, initially detected as variants in colony morphology. The molecular mechanisms of this and similar phenomena began to be unravelled half a century later. The term 'phase variation' is used to denote the high frequency, reversible, on/off switching of the expression of one or more genes, which (usually) encode or modify surface associated structures/components such as fimbriae/pili, lipopolysaccharides, lipoproteins, membrane proteins, iron acquisition proteins, surface enzymes, etc. The frequency of this reversible switching (10^{-4}–10^{-3}/cell/generation) in either direction is several fold higher than spontaneous mutation frquencies and leads to the presence or absence of the surface structure(s), encoded by the respective gene(s) (see below for references). As will be discussed in detail below, phase variation is an efficient means of generating population diversity and helps organisms, especially pathogens, in many ways such as survival in fluctuating environments, niche adaptation, virulence, evasion of host immune response, etc. Phase variable genes are called 'contingency' genes, as opposed to normal, 'house-keeping' genes. They have been found in many bacterial pathogens such as *H. influenzae*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Bordetella pertussis*, species of *Mycoplasma*, *H. pylori*, etc. just to name a few; their biology has been reviewed extensively[24–30]. The most common type of contingency genes present in many bacteria are characterized by the presence of varying numbers of contiguous, simple sequence repeats, or homopolymeric nucleotide tracts of varying lengths (also called microsatellites). These motifs are found within protein-coding sequences or in the promoter region (in between the –35 and –10 sequences) or downstream of the promoter before the translation start codons or upstream of the promoter where the accessory transcription factors act. (In some cases, they are found very much upstream of promoters; see below.) A mechanism called 'slipped strand mispairing (SSM)' has been suggested to explain the reversible on/off switching of the contingency genes[25,31,32]. SSM can occur during processes which involve DNA synthesis (replication, repair, recombination, etc.). It is believed that during DNA synthesis, the template and nascent DNA strands separate transiently and reanneal. During reannealing, the nascent strand can slip forward or backward on the template strand. This will result in the formation of a 'bulge' either in the template strand (due to forward slipping) or in the nascent strand (due to backward slipping), leading to the contraction or expansion, respectively, of the length of the repeat tracts in the next round of replication[25,31]. Depending on the location of the repeat sequences, transcription and/or translation of the genes could be affected in many ways when this event (phase variation) occurs. Some of the possible outcomes of SSM are: reduction/abolition of promoter activity, loss/reduction of accessory factor binding, generation of frameshift mutations leading to the synthesis of altered proteins or, more commonly, generation of premature termination (nonsense) codons.

Since most of the contingency genes encode surface-associated structures or components[32], the progeny cells will lack them when their expression is turned off by phase variation. A clonal population of cells will, therefore, consist of sub-populations in which the contingency genes will be in the on/off states of expression. Since the frequencies of switching are high and cells can have multiple contigency genes, it is possible to generate enormous population diversity by phase variation. (If there are n contingency loci in a cell, each of which could exist in the on/off states of expression, the number of possible sub-populations will be 2^n). As discussed below, such population diversity is extremely useful in adapting to stress, evasion of host immune response, modulation of virulence and many other processes. Although SSM is the most common mechanism of phase variation, it is not the only one. Others include homologous recombination, site specific recombination, transposon insertion/excision and differential DNA methylation. Novel mechanisms of phase variation are also coming to light. Extensive

literature on each of these processes is available. Due to space constraints, only a brief description of the salient features of these processes will be presented below, citing a limited number of examples. An exhaustive review by van der Woude and Baumler[32] is an excellent source for most of the information.

Other mechanisms of phase variation

Homologous recombination

Homologous recombination is associated with a process called 'antigenic variation', which is similar but not identical, to phase variation. Phase variation determines the presence or absence of one or more surface-associated cellular structure(s). In antigenic variation, only one form of a given antigen is present on the cell surface at any given time although the cell has the genetic potential to synthesize a family of that particular antigen (reviewed by van der Woude and Baumler)[32]. The best studied example of homologous recombination-mediated antigenic variation is pilin variation in *N. gonorrhoeae*[33]. This event occurs at a very high frequency (10^{-4} or more/cell/generation) and involves homologous recombination between a transcriptionally active *pilE* gene copy and one of the many (up to 6), transcriptionally silent, *pilS* gene copies, generating variation in *pilE* sequences without any change in the *pilS* sequences. Antigenic variation in the *pilE* gene of *N. gonorrhoeae* requires the products of *recA*, *recF* and *recX* genes. This is actually a process of gene conversion, the molecular details of which have been reviewed[33]. Similar to phase variation, antigenic variation also results in the generation of antigenically diverse subpopulations and helps in adaptation (evasion of host immunity; see below). Antigenic variation has been documented in several bacteria[32].

Site-specific recombination

The process of site-specific recombination in relation to phase variation has also been called conservative site-specific recombination. In this process, the inversion (flipping) of a short DNA segment harbouring the promoter of a phase variable gene results in expression of the gene in one orientation and non-expression in the other. A well-documented example of this process is the phase variation of the *fimA* gene, encoding type 1 fimbriae, in uropathogenic *E. coli* (UPEC). A short (~300 bp) invertible DNA segment intervenes *fimA* and two other genes (*fimB* and *fimE*) which encode two site-specific recombinases. In one orientation of the invertible segment the promoter of *fimA* is correctly positioned for its expression but not in the other. The flipping of this segment, mediated by the site-specific recombinases, results in the on/off switching of *fimA* and consequent phase variation. The flipping requires several other factors besides the recombinases (reviewed by van der Woude and Baumler[32]; also see ref. 35). The adaptive benefits due to *fim* phase variation are discussed in a later section.

Transposition

Phase variation mediated by insertion and excision of transposable elements is known to occur but restricted only to a few IS elements[32]. Generally, transposons have no target specificity in insertions, and excisions are almost always imprecise. However in a few instances insertions have been shown to be target-specific and excisions to be precise. Examples are transposon-mediated phase variation in the *eps* gene of *Pasteurella atlantica* (a marine organism) by the transposon IS492, the *cca* gene of *Staphylococcus epidermidis* by IS2561 (reviewed by van der Woude)[32]. Phase variation in capsule synthesis in *N. meningitidis* by the precise insertion and excision of IS1301 in the *siaA* (sialic acid biosynthesis) gene has been shown (see below).

GATC methylation

Phase variation by mechanisms involving DNA adenine methylase (*Dam*)-catalysed GATC methylation are well documented. A well characterized example is the phase variation of the pyelonephritis associated pili (*pap*) operon of UPEC[34]. In this case, the on/off state of the *papBA* genes is determined by differential binding of two proteins, Pap I and the leucine-responsive regulatory protein (Lrp), at one of the two GATC sites, one proximal to the promoter and the other distal to it. Binding of these proteins at one site blocks methylation at that site and leaves the other site open for methylation. When methylation occurs at the promoter proximal GATC site (Lrp and PapI bound at the promoter distal site), the operon gets into the on state, assisted by cyclic AMP-CAP (catabolite activator protein) complex. When methylation at the promoter proximal site is blocked by the binding of Lrp, the promoter distal GATC gets methylated and the expression is turned off. The above is an over-simplified description of the switching process which has been reviewed extensively[32,34,35]. Another example of GATC methylation-dependent phase variation is that of the *flu* gene of *E. coli* which encodes the multifunctional outer membrane protein Ag43, involved in cell aggregation, biofilm formation and phage adsorption. The regulatory region of *flu* has three GATC sites and the oxidative stress response protein OxyR is a repressor of its expression. When OxyR is bound at its binding site, the GATC sites are masked and hence the Dam methylase cannot access the sites. This turns the expression to the off state. When the GATC sites are methylated, binding of OxyR is blocked and the expression is turned on. Therefore the

on/off transition is determined by a stochastic, concentration-dependent competition between Dam methylase and OxyR for binding to the control site of *flu*[32,36]. It should be noted that the mechanism of phase variation outlined above is epigenetic since no changes in DNA sequences are involved but the expression states (on/off) are heritable[36]. The phase variation of *pap* described above is also epigenetic. These two systems are the only ones documented so far in this class of phase variable genes.

Phase variation driven by a ligand-sensitive repressor

A novel phase variable gene, viz. *nadA*, was reported by Martin *et al.*[37] in *N. meningitidis*. This gene has a tetranucleotide (TAAA) repeat motif (not reported earlier in phase variable genes of this organism), upstream of the –35 region of its promoter. Martin *et al.*[37] showed that the level of expression of *nadA* was dependent on the length of the repeat tract, being high (13, 10 and 8 repeats), or medium (11 and 12 repeats) or low (9 repeats). (In a subsequent report, it was shown that *nadA* encodes an adhesin/invasin, implicated in the pathogenicity of *N. meningitidis*; see below.) Metruccio *et al.*[38] reported the molecular mechanism of phase variation of *nadA*. They identified (i) a cis-acting regulatory region, named growth phase regulatory (GPR) region, located between –108 and –170 bp with respect to the *nadA* transcription start site and (ii) a protein, named NadR (encoded by the gene *nadR*) which binds to GPR and represses the expression of *nadA*. The repressor activity of NadR was observed to be influenced by the length of the tetranucleotide repeat motif. Moreover, they also showed that 4-hydroxyphenyl acetic acid, a product of aromatic aminoacid catabolism, inhibited NadR binding to the *nadA* operators, thereby derepressing (inducing) *nadA* expression. This is a novel mechanism of phase variation. Interestingly, the upstream regions of virulence genes of some other pathogens have been shown to contain repeat sequences, some of which also phase vary (for references see Metruccio *et al.*[38]) but the mechanistic details are not known.

Phase variation in genes encoding restriction–modification (RM) enzymes: Phasevarions

Generally, phase variable genes encode surface-associated cellular components. However, the repeat sequence motifs, characteristic of phase variable genes, have also been found in genes which are not obvious candidates to encode any surface component(s). Well studied examples are the genes encoding restriction–modification (RM) enzymes in many pathogenic bacteria[25,39,40] (also see refs 44 and 64). Although some of these genes were experimentally shown to phase vary (see below), many others were predicted to do so based on their DNA sequences. However, the functional role(s) of phase variation in these genes, remains obscure although several possibilities have been suggested (see below). Among the many-phase variable genes in strains of *H. influenzae* (listed by Moxon *et al.*[25]) the *mod* gene, encoding a DNA methyl transferase, is a component of a type III RM system. It was shown[40,41] that phase variation in this gene is mediated by a tetranucleotide repeat motif (AGTC in some strains and AGCC in some others). Using microarray expression analysis, Srikhanta *et al.*[42] reported that inactivation of the *mod* locus in *H. influenzae* Rd resulted in 2–4 fold-up regulation of seven genes and a similar down regulation of nine others. This is perhaps the first report of a global regulatory role for a phase variable gene. This novel genetic system in which the expression of multiple and unlinked genes is regulated by phase variation in a member of the *mod* gene family was named 'phasevarion' (phase variable regulon)[42]. In a subsequent report, Srikhanta *et al.*[43] reported that approximately 80 genes of *N. meningitidis* and *N. gonorrhoeae* were up/down regulated by null mutations in their respective *modA* genes. Regulation of gene expression controlled by GATC methylation, catalysed by the *dam*-encoded methylase, is a well-known phenomenon. But what distinguishes Dam methylase controlled regulation from the regulation by phasevarions is that, unlike *mod*, *dam* is not a phase variable gene. Therefore, phasevarions seem to be unique and novel modes of control of gene expression. An up-to-date review on phasevarions has appeared recently[44].

Phase variation in type I RM enzymes has also been shown in some strains of *H. influenzae*[45,46]. Wild type strains of *H. influenzae* Rd phase vary into two types which differ in their ability to restrict unmodified phage HPIcI: those which restrict (r^+m^+) and those which do not (r^-m^-). This transition has been shown to be mediated by the length of a pentanucleotide (GACGA) repeat motif, located at the beginning of the *hsdM* gene[46]. The r^+m^+ strains have 4 repeats whereas the r^-m^- variants have either 3 or 5 repeats. Restriction of unmodified phages can only limit the extent of productive infection (reduction in burst size) but cannot confer total resistance which results from alteration of the structure of the receptors located on the cell surface. In *H. influenzae*, Rd30 resistance to phage HPIcI has been shown[46] to occur by phase variation in the *lic2A* gene, resulting in the alteration of the surface lipooligosaccharide (LOS). Here, phase variation has been shown[46] to be mediated by a tetranucleotide (CAAT) repeat motif in the gene, phage-sensitive strains having 22 repeats and resistant ones having 21 repeats. The frequencies of phase variation in the above cases is influenced by mutation in the *dam* gene[46]. In the strain Rd118, switching from the r^+m^+ state to the r^-m^- state (on to off) and vice versa was approximately three-fold less in *dam* mutants. In the strain Rd30, however, phase variation from phage sensitivity to phage resistance (on to off) was approximately 10-fold higher in the *dam* mutant. Phase

variation to phage resistance could be adaptively beneficial since phages exert strong selection pressure for survival. Unlike phase variation in type III RM systems, that in type I RM systems has not been shown to affect the expression of other, unlinked genes.

Adaptive benefits of phase variation

Phase and antigenic variations are usually looked upon, among other things, as a strategy deployed by pathogens to evade immune defence mechanisms mounted by their hosts. Although it is easy to visualize the role of antigenic variation in evasion of immune defence, extrapolation of the same to phase variation is not so easy. The main difficulty is that bacteria lose a surface component when the gene involved in its synthesis phase varies to the off state and thus lack the function ascribed to it. However, this can also be looked upon as a blessing in disguise because cells lacking one or more surface structures can effectively 'hide' themselves from attack by the host's immune mechanisms[36]. Another problem to reconcile with is that phase variable genes have been found in many non-pathogens and non-host associated bacteria also[32]. Evasion of immune defence need not be the only use of phase variation; it could be a general strategy of stress adaptation. Population heterogeneity generated by phase variation could be useful in a more general way since there might be an assortment of cell types which could handle many kinds of stresses. The process could be viewed as an 'insurance' against catastrophe. In a sense, it is similar to bacterial persistence wherein a large population of antibiotic sensitive cells contain a small fraction of antibiotic-tolerant (not resistant) cells (reviewed by Jayaraman[47], and other references cited therein). Extensive literature is available on the adaptive benefits of phase variation. These have been reviewed in depth by several authors[30,32,35,48]. Only a few illustrative examples will be described below.

A recent report by Bayliss *et al.*[49] exemplifies the adaptive value of phase variation in evasion of host immunity. They showed that *N. meningitidis* strain 8047 escaped the lethality due to a specific monoclonal antibody (MAbB5) when the gene lipopolysaccharide glucosyl transferase (*lgtG*) phase varied by losing a single C residue from a tract of 12 Cs in its reading frame, resulting in an out-of-frame to in-frame change (off to on switch). The consequent expression of *lgtG* leads to the incorporation of glucose in the inner core of the surface lipopolysaccharide (LPS) and loss of antibody sensitivity. When *lgtG* is off, phosphoethanolamine gets incorporated in place of glucose in the LPS and the cells are sensitive to killing by MAbB5. A mutation in a mismatch repair gene (Δ *mutS*) was found to greatly enhance the escape from antibody-mediated killing, probably by increasing the phase variation rate, thereby increasing the number of cells in the population permissive for translation of *lgtG*, prior to exposure to the antibody[49]. Phase variation in the *lgtG* gene has been shown to be enhanced in *mutS* mutants[50]. This is an elegant demonstration of how phase variation could impact the fitness of a pathogen by countering the immune defences of the host. Antigenic variation in *pilE* by homologous recombination[33], described in an earlier section, is a very effective means of evading host immunity in *N. gonorrhoeae*.

Organisms belonging to the genus *Mycoplasma*, comprising many species, are among the smallest of free-living microbes. Many of them are pathogenic for humans as well as non-human animals, producing persistent infections, even in the face of adequate host immunity. This has been shown to be due to their remarkable ability to undergo phase and antigenic variations which alter their surface architecture, resulting in successful escape from host immune responses[51]. What is remarkable is that even with their very small genome size they are as proficient as bacteria with larger genomes in evading host immunity and maintaining infection.

In recent years, there has been a resurgence of pertussis (whooping cough), even in vaccinated populations. This indicates that the causative organism (*Bordetella pertussis*) has developed mechanisms to evade host immune defences, adapted to survive and cause disease. It is believed[52] that phase variability could be an important mechanism in this phenomenon.

Phase variation has also been shown to influence bacterial virulence and ability to initiate and maintain colonization of hosts during infection. A good example of this is provided by the study of niche adaptation in *H. pylori*. This organism has 31 phase variable genes (reviewed by Salaun *et al.*[53]). Phase variation in all the genes was monitored over a period of 150–360 days post-infection in a murine infection model[54]. It was observed that 10 of them phase varied at different times during the period, some early and some late. This suggested that phase variation determined not only the initial fitness for colonization but also the subsequent niche adaptation. To cite another example, turning on the expression of the *fimA* gene by phase variation has been shown to be essential for establishment of urinary tract infections by uropathogenic *E. coli* (UPEC)[55]. Snyder *et al.*[56] constructed mutants of UPEC in which phase variation of the *fim* operon was 'locked' in the on/off states and examined their virulence in a mouse infection model. The off-locked mutants were shown to be severely impaired in their ability to colonize the kidneys and bladder of mice whereas the on-locked mutants were as virulent as the wild type strain, showing the importance of the on state of *fim* expression in pathogenicity. Since the on state is affected by phase variation it follows that phase variation is needed for pathogenicity. *Klebsiella pneumoniae* is another uropathogen but less pathogenic and less prevalent than UPEC. In a recent report[57], it was shown that the reduced

pathogenicity of *Klebsiella pneumoniae* could be correlated with reduced ability to turn on the expression of the *fim* operon by phase variation. A gene, *fimK*, was shown to reduce piliation, and the absence of a gene (*fimX*), encoding a site-specific recombinase, were also shown to be contributing factors[57].

Yet another mechanism linking phase variation to bacterial virulence has been reported in *N. meningitidis*. Usually, *N. meningitidis* resides as a commensal organism in the human nasopharynx. Occasionally, it turns virulent and enters the blood stream causing sepsis. From the blood stream it could cross the blood–brain barrier, infect the brain and lead to fatal meningitis. What triggers the transition to virulence is largely unknown but it occurs more frequently in some strains of *N. meningitidis*, called hyper-virulent strains, than others. A gene called *N*eisseria *A*dhesin *A* (*nadA*) has been identified in some hypervirulent strains[58]. The NadA protein is a surface antigen which promotes adhesion to and entry into (invasion) epithelial cells. These properties were shown by *E. coli* also when *nadA* was expressed in them[58]. Inactivation of *nadA* reduced the ability to adhere and invade. These pathogenic properties could be correlated with phase viable expression of *nadA*. The novel mechanism of phase variation of *nadA* was described in an earlier section[38].

N. meningitidis and *N. gonorrhoeae* have a family of outer membrane proteins called Opa (opacity-associated) proteins, encoded by a group of phase variable, independently expressed, *opa* genes, numbering 3–4 in the former and 10–11 in the latter[59,60]. The Opa proteins are necessary for adhesion of the pathogen to host cells and also its internalization. These events involve the interaction between host cell receptors and the Opa proteins[59,60]. Phase variation in the *opa* genes is very frequent ($\sim 10^{-3}$/cell/generation) and involves pentameric (CTCTT) repeats in the coding sequences of the genes. Depending on the pattern of their expression, different cells in a population will have different complements of Opa proteins resulting in population heterogeneity. Whereas phase variation in the pilin genes in Neisseriae could be linked to evasion of host immunity[32,33], that *opa* genes could be linked to their invasiveness. A recent analysis of whole genome sequences of commensal and pathogenic strains of Neisseriae has revealed a high degree of similarity in the repertoire of virulence-associated genes between the two[61]. Interestingly, most of the commensal strains, with a few exceptions, lack the *opa* genes[61]. What are the critical determinants of the virulence/avirulence of Neisseriae is an area in which considerable research effort is focussed currently[61,62]. There are other examples to show the link between phase variation and pathogenicity of Neisseriae (reviewed by Hill *et al.*[63]). For instance, turning off the production of capsule in *N. meningitidis* by SSM in the polyC tract of the *siaD* gene and/or IS1301 insertion in the *siaA* gene enhances the adhesion of the pathogen to target cells by fully exposing sub-capsular adhesins[63]. A good deal of information on meningococcal pathogenesis, not discussed here due to space constraints, can be found in the review by Hill *et al.*[63].

As summed up by van der Woude[48], 'an encompassing view on the role of phase variation is that the generation of diverse sub-populations enhances the chance that at least one can overcome a stressful challenge, in essence a "hedge betting" strategy'.

Hypotheses on the adaptive advantages of phase variation in the *mod* genes of type III RM enzymes have centred around one of two possibilities, namely, DNA restriction activity or control of gene expression[39–44,64]. Many organisms which possess type III RM systems are naturally competent for DNA uptake and transformation. The presence of an active RM system will restrict incoming heterologous DNA (via transformation, phage infection, horizontal gene transfer, etc.) since its function is considered to be protection of cells from foreign DNA. One of the possible roles suggested for phase variation in the RM systems is that their off state of expression might temporarily relax this barrier and permit the uptake and integration of foreign, but useful, DNA. Recently, a prophage has been implicated in the invasiveness of *N. meningitidis*[65,66]. It has been suggested[39] that phase variation in the RM system of this organism might facilitate the horizontal spread of this phage among populations of Neisseriae, thereby enhancing their fitness (invasiveness).

A different role, namely, regulation of gene expression, has been proposed for phase variation in type III RM systems (see below). During a detailed analysis of the sequences of *mod* genes in various strains of *H. influenzae*, it was observed that the *mod* alleles could be grouped into 15 groups[40]. Strains belonging to three out of these 15 groups were found to be non-phase variable since they lacked the characteristic tetranucleotide repeats. Therefore, they cannot be considered to be part of any phasevarion and are probably dedicated solely to RM function. Among the strains possessing phase variable *mod* genes, 17% (7/41) had inactivating mutations in their *res* genes (encoding the restriction endonuclease)[40]. Fox *et al.*[40] proposed that the RM system in strains possessing phase variable *mod* genes and nonfunctional *res* genes have evolved to function exclusively as regulators of gene expression, losing their original RM function. Although obvious *res* mutations have been detected only in a few strains with phase variable *mod* genes, the possibility of yet undetected mutations, silencing the *res* in other strains cannot be ruled out[40]. Interestingly, none of the strains with non-phase variable *mod* genes had mutations in their *res* genes, supporting the notion that they are solely dedicated to RM function. In strains of *N. meningitidis* and *N. gonorrhoeae* possessing phase variable *mod* genes, a majority (70%) had frameshift or deletion lesions in their *res* genes[43]. Wherever the genes of the RM system has evolved to function as regulators of gene

expression (phasevarions), phase variation in those genes can influence several properties such as virulence, antibiotic sensitivity, biofilm formation, nutrient uptake, response to stress, antigenic composition, etc.[43].

If the *mod* gene of a strain is phase variable and if there is no inactivating mutation in the *res* gene (functional RM system), turning *mod* expression off will be inconsequential since Res will be functional only in association with Mod[67] (also see figure 1 of ref. 44). However, when *mod* expression is turned on after several generations of growth in the off state (restoration of a functional RM system), a problem of suicidal, self restriction of the genome will arise. When such an event occurs, the genome will be initially unprotected (unmethylated) and will be susceptible to restriction when the active RM system is restored. Surprisingly, it has been possible to isolate an variants from off state cultures of strains which have no obvious *res* inactivating mutations (M. P. Jennings, pers. commun.). 'Therefore Res is either inactive (via a non-obvious silencing mutation) or has no lethal phenotype for some other reason' (M. P. Jennings, pers. commun.). Another possibility is that since the contribution of type III RM system towards defence against foreign (unmodified) DNA is only marginal[40], a good fraction of the population could survive. Nevertheless, this intriguing question awaits a definitive answer.

Concluding remarks

The success of adaptive evolution by natural selection depends upon how well organisms adjust and adapt to their environments. Bacteria, especially pathogens, encounter several unpredictable threats to their survival and are constrained to keep evolving adaptive strategies to tide over such threats. The threats also change and evolve. For example, pathogens would try many tricks to fool their hosts to establish themselves during infection. The hosts, in their turn, would try to come up with more, 'new and improved', hurdles to contain and eliminate the invader. Thus, there is a constant coevolution of the threat and the response. This is likened[68] to the Red Queen's Race in Lewis Carrol's classic, *Through the Looking Glass* in which the Red Queen tells Alice 'It takes all the running you can do to keep in the same place. If you want to get somewhere else, you must at least run twice as fast'. In simple, layman's language, stress and response could be compared to a predator and its prey, respectively. If the predator runs fast to catch the prey, the prey has to run faster to escape the predator. Evolutionarily speaking, phase variation (as well as heritable and transient hypermutability mechanisms, not discussed presently) empower bacteria to 'run faster' to deal with unforeseen and life-threatening stresses. The literature reviewed in the foregoing pages, though not exhaustive, illustrates how bacteria exploit the power of phase variation to evolve counter offensive strategies for survival and perpetuation by generating genetic/population diversity. When a large population diversity is available it is possible that some among them could be adaptively useful and they could be taken advantage of to counter the threat. As pointed earlier, phase variation insures a cell population against challenges to survival.

1. Hengge-Aronis, R., The general stress response in *Escherichia coli*. In *Bacterial Stress Response* (eds Storz, G. and Hengge-Aronis, R.), ASM Press, Washington DC, USA, 2000, pp. 161–178.
2. Hengge-Aronis, R., Signal transductions and regulatory mechanisms involved in the control of sigma(S) (RpoS) subunit of RNA polymerase. *Microbiol. Mol. Biol. Rev*., 2002, **66**, 373–395.
3. Weber, H., Polen, T., Hueveling, J., Wendisch, V. F. and Hengge, R., Genome-wide analysis of the general stress-response network in *Escherichia coli*: σ S-dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.*, 2005, **187**, 1591–1603.
4. Hengge, R., The two-component network and the general stress sigma factor RpoS in *Escherichia coli*. *Adv. Exp. Biol. Med.*, 2008, **631**, 41–63.
5. Hengge, R., Proteolysis of RpoS and the general stress response in *Escherichia coli*. *Res. Microbiol.*, 2009, **160**, 667–676.
6. Ron, E. Z., Editorial: An update on bacterial stress response. *Res. Microbiol.*, 2009, **160**, 243–244.
7. Galhardo, R. S., Hastings, P. J. and Rosenberg, S. M., Mutation as a stress response and regulation of evolvability. *Crit. Rev. Biochem. Mol. Biol.*, 2007, **42**, 399–435.
8. Foster, P. L., Stress induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.*, 2007, **42**, 373–397.
9. de Visser, J. A. G. M., The fate of microbial mutators. *Microbiology*, 2002, **148**, 1247–1252.
10. Oliver, A., Hypermutation in natural bacterial populations: consequences for medical microbiology. *Rev. Med. Microbiol.*, 2005, **16**, 1–25.
11. Labat, F., Pradillon, O., Gary, L., Peuchmaur, M., Fantin, B. and Denamur, E., Mutator phenotype confers advantage in *Escherichia coli* chronic urinary tract infection. *FEMS Immunol. Med. Microbiol.*, 2005, **44**, 317.
12. Denamur, E. and Matic, I., Evolution of mutation rates in bacteria. *Mol. Microbiol.*, 2006, **60**, 820–827.
13. Jayaraman, R., Mutators and hypermutability in bacteria: the *Escherichia coli* paradigm. *J. Genet.*, 2009, **88**, 379–391.
14. Kibota, T. T. and Lynch, M., Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature*, 1996, **331**, 694–696.
15. Boe, L., Danielsen, M., Knudsen, S., Petersen, J. B., Maymann, J. B. and Jensen, P. R., The frequency of mutations in populations of *Escherichia coli*. *Mutat. Res*., 2000, **448**, 47–55.
16. Imhoff, M. and Schlotterer, C., Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Nature*, 2001, **381**, 694–696.
17. Wright, B. E., Stress directed mutations and evolution. *Mol. Microbiol*., 2004, **52**, 643–650.
18. Roth, J. R., Kugelberg, E., Reams, A. B., Kofoid, E. and Andersson, D. I., Origin of mutations under selection: the adaptive mutation controversy. *Annu. Rev. Microbiol.*, 2006, **60**, 477–501.
19. Galhardo, R. S., Do, R., Yamada, M., Friedberg, E. C., Hastings, P. J., Nohmi, T. and Rosenberg, S. M., DinB up regulation is the sole role of the SOS response in stress-induced mutagenesis in *Escherichia coli*. *Genetics*, 2009, **182**, 55–68.
20. Hastings, P. J., Adaptive amplification. *Crit. Rev. Biochem. Mol. Biol.*, 2007, **42**, 271–283.

21. Gonzalez, C., Hadany, L., Ponder, R. C., Price, M., Hastings, P. J. and Rosenberg, S. M., Mutabilty and importance of a hypermutable cell subpopulation that produces stress induced mutations in *Escherichia coli*. *PLoS Genet.*, 2007, **4**, e1000208.
22. Andersson, D., Hughes, D. and Roth, J., The origin of mutants under selection: interaction of mutation, growth and selection. In *Eco Sal: Escherichia coli and Salmonella, Cellular and Molecular Biology* (eds Bock, A. *et al.*), ASM Press, Washington DC, USA, 2009.
23. Roth, J., Genetic adaptation: a new piece for a very old puzzle. *Curr. Biol.*, 2010, **20**, R15–R17.
24. Moxon, R. E., Rainey, P. B., Nowak, M. A. and Lenski, R. E., Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.*, 1994, **4**, 24–33.
25. Moxon, R., Bayliss, C. and Hood, D., Bacterial contingency loci: the role of simple sequence repeats in bacterial adaptation. *Annu. Rev. Genet.*, 2006, **40**, 307–333.
26. van Belkum, A., Scherer, S., van Alphen, L. and Verbrugh, H., Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. *Infect. Immunol.*, 1998, **65**, 5017–5027.
27. Moxon, R. E. and Wills, C., DNA microsatellites: agents of evolution? *Sci. Am.*, 1999, **280**, 94–99.
28. Metzgar, D. and Willis, C., Evidence for the adaptive evolution of mutation rates. *Cell*, 2000, **101**, 581–584.
29. Bayliss, C. D., Field, D. and Moxon, R. E., The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningiditis*. *J. Clin. Invest.*, 2001, **107**, 657–662.
30. Salaun, L., Snyder, L. A. and Saunders, N. J., Adaptation by phase variation in pathogenic bacteria. *Adv. Appl. Microbiol.*, 2003, **52**, 263–301.
31. Hallet, B., Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. *Curr. Opin. Microbiol.*, 2001, **4**, 570–581.
32. van der Woude, M. V. and Baumler, A. J., Phase and antigenic variation in bacteria. *Clin. Microbiol. Rev.*, 2004, **17**, 581–611.
33. Kline, K., Hechman, E. V., Scaar, E. P. and Seifert, H. S., Recombination, repair and replication in the pathogenic *Neisseriae*: the 3Rs of molecular genetics of two human-specific bacterial pathogens. *Mol. Microbiol.*, 2003, **50**, 3–13.
34. Hernday, A., Krabbe, M., Braatten, B. and Low, D., Self-perpetuating epigenetic pili switches in bacteria. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 1640–1647.
35. Bayliss, C. D., Determinants of phase variation rates and fitness implications of differing rates for pathogens and commensals. *FEMS Microbiol. Rev.*, 2009, **33**, 504–520.
36. van der Woude, M. V., Some types of bacterial phase variations are epigenetic. *Microbe*, 2008, **3**, 21–26.
37. Martin, P. *et al.*, Experimentally revised reportoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. *Mol. Microbiol.*, 2003, **50**, 245–257.
38. Metruccio, M. M. E. *et al.*, A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. *PLoS Patho.*, 2009, **5**(12), e1000710.
39. Bayliss, C. D., Callaghan, M. J. and Moxon, R. E., High allelic diversity in methyltransferase gene of a phase variable type III restriction modification system has implications for the fitness of *Haemophilus influenzae*. *Nucleic Acids Res.*, 2006, **34**, 4046–4059.
40. Fox, K. L., Dowideit, S. J., Erwin, A. L., Srikhanta, Y. N., Smith, A. L. and Jennings, M. P., *Haemophilus influenzae* phasevarions have evolved from type III DNA restriction systems into epigenetic regulators of gene expression. *Nucleic Acids Res.*, 2007, **35**, 5242–5252.
41. De Bolle, X. *et al.*, The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase variation rate of a gene with homology type III DNA methyltransferases. *Mol. Microbiol.*, 2000, **35**, 211–222; Erratum. *Mol. Microbiol.*, 2002, **46**, 293.
42. Srikhanta, Y. N., Maquire, T. L., Stacey, K. J., Grimmond, S. M., and Jennings, M. P., The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 5547–5551.
43. Srikhanta, Y. N. *et al.*, Phasevarions mediate random switching of gene expression in pathogenic *Neisseriae*. *PLoS Path.*., 2009, **5**(5), e1000400.
44. Srikhanta, Y. N., Fox, K. L. and Jennings, M. P., The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.*, 2010, **8**, 196–206.
45. Glover, S. W. and Piekarowicz, A., Host specificity of DNA in *Haemophilus influenzae*; restriction and modificatioin in strain Rd. *Biochem. Biophys. Res. Commun.*, 1972, **46**, 1610–1617.
46. Zaleski, P., Wojciechow, M. and Piekarowicz, A., The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection. *Microbiology*, 2005, **151**, 3361–3369.
47. Jayaraman, R., Bacterial persistence: some new insights into an old phenomenon. *J. Biosci.*, 2008, **33**, 795–805.
48. van der Woude, M. V., Re-examining the role and random nature of phase variation. *FEMS Microbiol. Lett.*, 2006, **254**, 190–197.
49. Bayliss, C. D., Hoe, J. C., Makepeace, K., Martin, P., Hood, D. W. and Moxon, R. E., *Neisseria meningiditis* escape from the bactericidal activity of a monoclonal antibody is mediated by phase variation of *lgtG* and enhanced by a mutator phenotype. *Infect. Immunol.*, 2008, **76**, 5038–5048.
50. Martin, P., Sun, L., Hood, D. W. and Moxon, R. E., Involvement of genes of genome maintenance in the regulation of phase variation frequencies in *Neisseria meningitidis*. *Microbiology*, 2004, **150**, 3001–3012.
51. Citti, C., Nouvel, L. and Baronowski, E., Phage and antigenic variation in mycoplasma. *Future Microbiol.*, 2010, **5**, 1073–1085.
52. Gogol, E. B., Cummings, C. A., Burns, R. C. and Relman, D. A., Phase variation and microevolution at homopolymeric tracts in *Bordetella pertussis*. *BMC Genomics*, 2007, **8**, 122.
53. Salaun, L., Linz, B., Suerbaum, S. and Saunders, N. J., The diversity within an expanded and redefined repertoire of phase variable genes in *Helicobacter pylori*. *Microbiology*, 2004, **150**, 817–830.
54. Salaun, L., Ayraud, S. and Saunders, N. J., Phase variation mediated niche adaptation during prolonged experimental murine infection with *Helicobacter pylori*. *Microbiology*, 2005, **161**, 917–923.
55. Snyder, A. J. *et al*., Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection. *Infect. Immun.*, 2004, **72**, 6373–6381.
56. Snyder, J. A., Lloyd, A. L., Lockatell, C. V., Johnson, D. E. and Mobley, H. L. T., Role of phase variation of type 1 fimbriae in a uropathogenic *Escherichia coli* cystitis isolate during urinary tract infection. *Infect. Immun.*, 2006, **74**, 1387–1393.
57. Rosen, D. A., Pinker, J. S., Jones, J. M., Walker, J. N., Cleg, S., and Hultgren, S. J., Utilization of an intracellular bacterial community pathway in *Klebsiella pneumoniae* urinary tract infection and the effects of *fimK* on type I pilus expression. *Infect. Immun.*, 2008, **76**, 3337–3345.
58. Capeechi, B. *et al.*, *Neisseria meningitidis* NadA is a new invasion which promotes bacterial adhesion and penetration into human epithelial cells. *Mol. Microbiol.*, 2005, **55**, 687–698.
59. Dehio, C., Grey-Owen, S. D. and Meyer, T. F., The role of *Neisseria* Opa proteins in interaction with host cells. *Trends Microbiol.*, 1998, **6**, 489–495.

60. Hauck, C. R. and Meyer, T. F., Smal talk: Opa proteins as mediators of *Neisseria* – host cell communication. *Curr. Opin. Microbiol.*, 2003, **6**, 43–49.
61. Marri, P. R. *et al.*, Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. *PLoS ONE*, 2010, **5**, e11835.
62. Schoen, C. *et al.*, Whole-genome comparison of disease and carriage strains provides insights into virulence evolution in *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 3475–3478.
63. Hill, D. J., Griffiths, N. J., Borodina, E. and Virji, M., Cellular and molecular biology of *Neisseria meningitidis* colonization and invasive disease. *Clin. Sci.*, 2010, **118**, 547–564.
64. Fox, K. L., Srikhanta, Y. N. and Jennings, M. P., Phase variable type III restriction–modification systems of host-adapted bacterial pathogens. *Mol. Microbiol.*, 2007, **65**, 1375–1379.
65. Bille, E. *et al*., A chromosomally integrated bacteriophage in invasive *meningococci. J. Exp. Med*., 2005, **201**, 1905–1913.
66. Moxon, R. E. and Jansen, V. A., Phage variation: understanding the behaviour of an accidental pathogen. *Trends Microbiol.*, 2005, **13**, 563–565.
67. Dryden, D. T., Murray, N. E. and Rao, D. N., Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res*., 2001, **29**, 3728–3741.
68. Van Valen, A new evolutionary law. *Evol. Theory.*, 1973, **1**, 1–30.

ACKNOWLEDGEMENTS. I thank Arul Jayaraman and Sachin Jayaraman for their valuable help in literature search. I also thank the anonymous reviewers for their critical comments.

Received 30 August 2010; revised accepted 14 January 2011