Abstract

In nature, bacteria are exposed to multiple stress conditions posing threat to their life. However, bacteria evolve constantly and employ myriad of genetic and phenotypic strategies for successful survival. One such adaptive process is phase variation - a random, frequent and reversible ON/OFF switching of contingency genes generating phenotypical variations within the bacterial population. Pathogenic and commensal bacteria make use of this heritable process as a tool for adaptation, generating intra-strain diversity and immune evasion. Unveiling molecular basis of this aspect of bacterial pathogenesis is vital in yielding significant implications to understand health and diseases. The current review was planned to overview this hot topic of research, its molecular basis and biological significance.

Keywords: Phase variation, Phase variation, Contingency gene.

Introduction

Survival of a single bacterial species might be tested in diverse environmental conditions.¹ For effective colonisation and fruitful survival, pathogenic bacteria must adapt their physiology to dynamic and unpredictable changes occurring in the host environment or during transmission from one host to the other. Bacterial adaptation strategies encompass either selection of fittest variant form as a result of mutation or by sensing and responding to stringent and changing host conditions.² One such adaptive process is Phase variation (PV) that mediates phenotypic heterogeneity in bacterial population through molecular mechanisms.³

PV is a heritable and reversible process and its important feature is interchangeable ON/OFF expression switching of genes. This means that cells with gene expression in OFF mode still retain the ability to switch it ON. Such ON and OFF switching of gene(s) results in two bacterial sub-population: one has the ability to express the gene in full swing, while in the other the expression of phase variable genes are either reduced or completely OFF. The rate of switching varies far and wide: it might be as common as one in hundred per generation to as rare as one in thousand per generation. PV frequency of certain genes can be varying between strains and can also be affected by environmental conditions. This process is a random event which make it difficult to guess which bacterial cell within bacterial population will be subject to change.²,³ In addition, PV is also referred as “programmed event” because bacterial genome that exhibit such events is organised so that may lead to certain events to occur.⁴,⁵ Gram-negative bacteria appeared to be the best example of PV as they give rise to visible phenotypic changes such as colony morphology and aggregation. In addition, this process is also associated with genes that are not involved in host-pathogen interaction and act as a component of bacterial defence system.⁵

PV is a product of hyper mutable simple sequence repeats (SSRs) composed of one to several nucleotides and are located either within reading frames or in the promoter region of the gene, known as contingency loci. Such contingency loci are directly involved in interaction with host components and are frequently subjected to polymerase slippage to generate phenotypic variation for survival in adverse environmental conditions.⁶ These SSRs are reported to be associated with pathogenic bacteria adaptation and virulence such as Bordetella pertussis, Campylobacter jejuni, Helicobacter pylori, Haemophilus influenzae and Neisseria meningitides.²,⁷,⁸ SSRs present within the coding region of a gene known as Intragenic SSRs can cause frame-shift mutation within coding sequences and thus can affect its translation and lead to premature termination of translation or altered protein C terminal. On the other hand, intergenic SSRs interfere with transcription process by altering the distance between transcriptional factor binding sites or promoter elements.⁹

PV is a vital tool for bacterial success in term of colonisation, survival and virulence. In case of host-pathogen interaction, this process help pathogens in escaping immune system of host and modulation, a consideration that affects colonization, role in biofilm...
formation and detachment as well as in vaccine development. Moreover, PV is considered as an important defence tool by protecting cell from foreign deoxyribonucleic acid (DNA), phages and plays an indirect role in strain evolution. Research areas of interest in PV ranges from its occurrences at molecular level that mediate heterogeneity in bacterial population and the impact of PV on bacterial population. Advances have been made in understanding PV associated with pathogenic and commensally bacteria that yield important implications in understanding health and disease.\textsuperscript{1,2-9}

**Mechanisms of Phase Variation**

Diphasic Salmonellas by Andrews was the first published report of PV in 1992 and was distinguished sharply by agglutination behaviour. Subsequent studies revealed that Salmonella’s diphasic agglutination stage was due to promoter sequences inversion containing DNA sequences that control expression of flagellar biosynthesis genes. Since then a number of bacterial systems have been reported that undergo genomic rearrangement.\textsuperscript{5}

1. **Site-Specific Recombination**

Site-specific recombining sequences are short and recombination requires single site of reaction within recombining sequences.\textsuperscript{10} In a number of instances, site-specific inversion is a simple ON to OFF switching of genes positioned within or neighbouring invertible region. This process of switching can be accomplished by varying the spatial arrangement of regulatory region elements or promoters within the region of the affected gene. The well-characterised example of site-specific DNA inversion mechanism are Escherichia coli type 1 fimbriae and Salmonella typhimurium flagella.\textsuperscript{11}

1.1. **Site-specific Inversion**

Site-specific inversion require the presence of specific recombinase enzymes which invert particular DNA fragment resulting in ON to OFF or OFF to ON switching of gene placed within the invertable segment of DNA or next to the switch. This strategy of inversion is adopted by different bacterial species during infection process by changing the expression of selected genes. The inversion event may be simple by changing the expression of a single gene like in case of E.coli pillin gene or it may be complex by affecting multiple genes expression as in case of S.Typhimurium multiple types of flagglin.\textsuperscript{12} Variation in Type 1 fimbriae is achieved by inversion of 314 base pair (bp) chromosomal sequences positioned upstream of structural subunit encoding gene, fimA. The promoter of fimA gene is present within the invertible sequences: correct positioning of promoter will promote the transcription of fimA gene: ON orientation while in case of incorrect positioning of fimA gene promoter will inhibit its transcription: OFF orientation. In addition to regulatory proteins, histone like protein (H-NS), integrated host factor (IHF), leucine response protein (Lrp), product of fimB and fimE gene controlled the inversion of DNA segment. Recently it has been found that ribonucleic acid (RNA) polymerase, sigma S(RpoS), a sigma factor, influences type 1 fimbriae gene expression.\textsuperscript{13}

The product of both fimB and fimE can mediate the process DNA inversion and show sequence homology with lambda integrases group of site-specific recombinases. The product of fimE results in ON to OFF action, while fimB product action will promote both ON/OFF and OFF/ON activity. Fimbriate to afimbriate switch can be mediated by FimE gene product at twice magnitude frequency than FimB gene product, providing an efficient mechanism of rapidly ending fimbriae synthesis in reply to suitable environmental stimuli.\textsuperscript{14}

Methylation blocking factor (Mbf) or presently known as Lrp is required for FimB and FimE mediated inversion. Type 1 fimbriae inversion is mediated by Lrp and is potentiated in leucine presence. There are three binding sites for Lrp gene product within invertible sequences. DNA confirmation produced by Lrp binding to site 1 and 2 in comparison to binding of Lrp to all three sites is more likely to undergo inversion. The action of leucine is more selective on site 3 bounded leucine as it encourages dissociation from this site and mediates the inversion of fim.\textsuperscript{14,15}

Lrp and IHF may act in concert and may promote DNA bending. It has been proposed that IHF binds to two sites: site I and site II. Site I is present downstream of FimE gene while site II is present within invertible element. IHF can introduce DNA bending and before recombination event is responsible for aligning inverted repeats present at the terminal sites of fim element. Binding of IHF to site II occur in concert with Lrp by binding to DNA that result in two loops, thus aligning recombination sequences at juxtaposition to ensure strand exchange phenomena. The binding of IHF to site I is not very clear but binding of IHF to site II perform structural role in recombination process.\textsuperscript{15}

In E. coli and S. typhimurium, an important component of nucleoids is the H-NS protein. This protein is present in three iso-forms and has the capability of compacting DNA into nucleosome-like structures. Pleiotropic changes can be resulted when mutation occurs in H-NS gene of E. coli. These pleiotropic effects include deregulation of type 1
site-specific recombination. It has been indicated recently that H-NS protein binds to upstream position of fimA promoter and accelerates its activity in ON orientation switch.14,15 The binding region of H-NS protein located immediately downstream of fimE, encompasses site I binding of IHF and extending into the switch region. These findings raise the possibility that IHF interfere with the binding of H-NS in such a way that H-NS enhances transcription from fimA promoter only in IHF presence.15

In comparison to type 1 fimbrial system, which exhibits a simple ON/OFF phenomena of PV, site-specific recombination phenomena in Salmonella flagella results in a more complex ON/OFF, ON/OFF PV. In Salmonella typhimurium, this inversion event can be achieved by reversible inversion of 966 bp DNA, bounded by two homologous chromosomal sites such as hiXL and hiXR. These homologous chromosomal hix sites contain two symmetrical half-sites of 12 base pair separated by a bridge of 2bp. In such ON/OFF to ON/OFF switching, a switch occurs between the expression of two antigenically different types of flagella i.e. H1 and H2. The promoter for transcription of H2 flagellin (h2) gene and rh1 repressor gene, located adjacent to H2 flagellin gene, is present within the invertible region. When the configuration of H2 is in ON condition, H2 type of flagella and rh1 gene product (rH1) are produced, therefore transcription of h1 gene is repressed. As a result of inversion, repressor expression is stopped and h1 flagellin is expressed.16

Site-specific recombinase (Hin) enzyme is required for DNA segment inversion. This enzyme is present within invertible DNA segment and has the ability to bind and catalyse the phenomena of site-specific recombination between the hix sites at either end of invertible DNA segment. An accessory protein known as Fis is required to stimulate Hin activity by binding to recombination enhancer site.17

1.2. Transposition Mechanism

Transposition mechanism leads to either insertion or excision of transposable element and is restricted to few insertion sequence (IS) elements. Transposition event provides a reversible mechanism of restoring original sequences of recipient DNA if the excision event is precise, but in most cases the excision is not precise as a result of which original sequences of recipient DNA is not restored. Transposition mechanism seems to less target-specific in terms of insertion, while excision events are almost imprecise. In few cases, however, insertion events showed target specificity while excision events to be precise.1

Transposition-mediated PV examples includes eps gene of marine organism, Pasteurella atlantica by IS492 transposons. Certain isolates of this marine bacterium are known to exhibit variation in an eps locus that codes for extracellular polysaccharide and the expression of this gene is regulated by the presence or absence of IS942. A pair of recombinase enzymes is required for this event encoded by two genes, MooV and piv that mediate precise exision and insertion event respectively.19 Staphylococcus epidermis, a gram-negative bacterium, is also known to vary the expression of ICA operon by transposon-mediated phase expression mechanism. The expression of this operon results in the formation of adhesin (polysaccharide), play essential role in biofilm formation and cell-cell interaction.19 PV is also found to be mediated via insertion and excision of plasmid sequences. This mechanism of PV has been found in some isolates of Legionella pneumophila, a gram-negative bacterium. PV in these isolates is connected with reversible insertion and excision of a plasmid that is about 30-kb in size into the chromosome. Although the molecular basis is not known, but it seems to be RecA-dependent mechanism.19

2. General Recombination

A prominent mechanism for genomic diversity and rearrangement of genetic material in prokaryotes is general homologous recombination. Homologous recombination generally requires long stretch of homologous DNA that is greater than 50 base pair in length and relies on number of DNA repair and cell maintenance machinery proteins. Unidirectional exchange of DNA between two allele of a gene results in gene conversion. Combinatorial variation results when alleles undergo constant changes as a result of recombination. Gene conversion mechanism varies among species and requires general homologous recombination machinery. Nevertheless, this mechanism is different from general recombination in several features i.e., the frequency of recombination is much higher compared to other Rec-dependent recombination plus recombination can occur between less-stretched homologous regions.20 Type IV pilin in Neisseria gonorrhoeae appeared to be the best source of understanding gene conversion mechanism that further leads to antigenic variation of the mentioned pilli. Proteins that are responsible for antigenic variant of type IV pilli are conserved for two-third of N-terminus and vary at the remaining C-terminus. One-way variation causes expression of pilE locus using one of the six silent pilS loci present in the genome. These six pilS loci are placed from pilE locus at a distance of about 900kb. The gene present at pilE locus has both conserved and variable regions while that present at pilS locus mainly contains variable...
region. Rec-F like pathway, in which Rec-A is a vital player, plays an important role in unidirectional exchange plus Rec-A is required for antigenic variation. The "hybrid intermediate" model is the most recent proposed model that describes how this genetic exchange occurs and several critical aspects of this model have been verified experimentally. According to this model, the first event is the formation of pilE-pilS, an extra chromosomal circular hybrid molecule, formed as a result of RecA independent recombination between short homologous region of pilS locus and pilE sequences of the donor chromosome. In the second step the hybrid molecule donates sequence of pilS to pilE on recipient chromosome in the presence of RecA and requires higher region of homology as well as some homologous region within the gene.\(^{20,21}\) An important feature of this recombination is that despite one-way genetic exchange, the fidelity of chromosome is maintained and sequences at pilS locus remain unchanged. It is also possible that recombination event leads to deletion of pil gene region, forming irreversible, non-expressing phenotypes. Mechanisms and molecular pathways underlying antigenic variation of Borrelia hermsii variable major lipoprotein (Vmp) and B. burgdorfer surface lipoprotein (VlsE) is less understood. Four types of mechanisms have been identified via indepth analysis of genetic exchanges that lead to seroconversion events in B. hermsii. The first mechanism undergoes nonreciprocal recombination of genes present at silent loci (archival) from linear plasmid to Vlp7 expressed loci present near telomeric regions of plasmid Ip28-1. The second mechanism only occurs at expression sites and results in loss of DNA fragment and involves less occurrences of intraplasmidic recombination at duplicate sequences containing region. The third mechanism introduces point mutation at expression locus that may also result from archival loci. The last and final mechanism is not much clear, but lead, to silencing of gene transcription at expression site on Ip28-1 plasmid and expression of Vsp33 from internally located site of 53kb plasmid. In campylobacter, antigenic variation of Surface layer proteins (SLPs) is the result of homologous recombination that leads to its antigenic variation. This involves re-assortment of eight Surface associated proteins (SAP) genes each encoding distinct antigenic SAP and a single SAP promoter. The inversion of DNA involves 6.2kb fragment with single promoter or sometime in addition require flanking region with either one or more variable sap cassettes. The DNA is arranged in such a position that only one promoter is used to transcribe only one of the eight SAP gene. In case of H. Influenza type B, gene duplication of complete capsule (CAP) gene causes heritable variation in production of capsule production level which can also be enhanced by IS-like sequences.\(^{20,21}\)

System of complex recombination induces variation based on deletion incidents which caused only one-way ON/OFF. In order to inactivate phase variants so that they do not accumulate in a specific population, gene duplication or gene transfer events occur in most cases. Neisseria gonorrhoeae revealed RecA-dependent high-frequency phenotype switching of type IV pilin. Structural subunit of type IV pilus is transcribed from pilE (E for expressed) locus while transcriptionally inactive alleles known as pilS (S for silent) is present in a different place on chromosome. The 5’ end of pilE gene may be the target of recombinational deletion mechanism that result in ON to OFF variant.\(^{20-22}\)

### 1.3. Slipped Strand Mechanism (SSM)

SSM is thought to be the most common mechanism of ON/OFF switching of contingency genes. This process of SSM occurs during the process of DNA synthesis i.e., DNA replication, repair and recombination. During DNA synthesis template and nascent strand transiently separates from each other and then reanneal. In reannealing step, nascent strand on the template strand can be slipped either in forward direction or in backward direction, resulting in "bulge" formation. The bulge is formed either on the template strand due to forward slippage or in the nascent strand due to backward slippage, leading to either contraction or expansion of the repeat tracts which in turn could affect in many ways transcription or translation of the affected contingency gene depending on the position of the repeat tract. Some of the possible outcomes of SSM include premature termination due to nonsense codon generation or altered protein due to frame-shifted mutation, loss or reduction of accessory proteins binding, loss or reduction of promoter’s activity.\(^{3,24}\)

### 1.4. Phase Variation via Differential Methylation

Variation mediated by differential methylation is described as epigenetic event as phenotypes are altered but not genotype, therefore maintaining the integrity of genome. This process undergoes deoxyadenosine methylase (Dam) dependent methylation of GATC sequence in E. Coli and then methylate adenosine at position N.\(^6\) The well-documented example of this process is pyelonephritis-associated pili (pap) operon of Uropathogenic E. coli (UPEC).\(^{21,24}\) Two sites of GATC (GATC-I and GATC-II), each of which is present in pap locus, are differentially methylated. Both these GATC sites affect divergently transcribed operons such as papI and papBA.
Methylation of both these sites are regulated by competitive action of two elements i.e. Dam and Lrp. Two possibilities may result from this: one, methylation of GATC-I and not of that GATC-II results in phase ON cells; two, in conditions where GATC-I is not methylated and GATC-II is methylated, phase OFF cells are generated.

Promoter of pap gene is present upstream of papBA, overlapping GATC-II site. Lrp molecule binding to this region leads to inhibition of papBA transcription. On the other hand, binding of Lrp molecule to alternative sites results in lifting inhibition. Lrp molecule binding to site 4, 5, and 6 is facilitated by papI protein binding to Lrp/DNA complex at GATC-I site and by GATC-II site methylation. Other proteins which play an important role in the pathogenicity of Neisseria Adhesion A (nadA) and other phase variable genes include PapB, cAMP receptor protein (CRP) and H-NS. DNA binding by PapB is in oligomeric fashion and leads to auto-regulation of its own gene expression. Upstream of papBA promoter CRP proteins binds and is needed for papBA and papI transcription.

1.5. Ligated-sensitive Repressor Mediated Phase Variation

A novel phase variable gene in Neisseriameningitidis entitled as Neisseria Adhesion A (nadA) undergo phase variation because of tetra nucleotide repeat located -35 bp upstream of the nadA gene promoter which controls the expression of this gene. The repeats vary in length; it may be as high as 13, 10 and 8, or it may be in medium length as 11, 12, or it may be as low as 9. Subsequent studies revealed that this gene encode adhesion protein which play an important role in the pathogenicity of N.meningitides. The growth phase regulatory (GPR) region which lies between -108 and -170 bp with respect to nadA gene transcription start point and NadR protein, encoded by nadR gene, binds to GPR and repress nadA gene expression. NadR protein repressor activity was found to be affected by the length of tetra nucleotide repeat motif. In addition, aromatic amino acid catabolism which produce 4-hydroxyphenyl acetic acid inhibiting NadR binding to nadAthus resumes the expression of nadA gene. Further studies revealed that upstream region of many virulence gene in pathogenic bacteria possess repeat tracts that exhibit PV phenomena but detail mechanism is still unknown.

Phase Variable Genes of Restriction Modification (R-M) System (Phasevarion)

The genes of R-M system in pathogenic bacteria represent the well-studied example of phase variable genes. Only few genes are experimentally showed to vary but a number of others are predicted on the basis of DNA sequence. In Haemophilus influenzae, type III RM system DNA methyltransferase enzyme encoding gene, mod gene, mediate PV through tetranucleotide repeat tract motifs. Microarray expression systems revealed that inactivation of H. influenzae mod gene results in 2-4 fold up-regulation of seven genes while down-regulate nine others. Such a system in which multiple genes expression are regulated is termed as phasevarion (phase variable regulon). Approximately 80 genes of N. gonorrhoeae and N. Meningitides were found to be up and down-regulated due to null mutation in their respective mod gene.

Genes of type I R-M system of H. influenzae has also been found to show PV phenomena. H. influenzae strain uses two ways to phase vary and differ from one another by their ability to restrict unmodified phage; those which can restrict are entitled as r+m+ and those which are not are designated as r-m-. This transition is due to varying length of pentanucleotide repeat motif (GACGA) present in the beginning of hsdM gene. The r+m+ strains possess 4 such repeat while r-m- posses 3 to 5 repeats. In H. influenza, PV in lic2A gene is mediated by tetranucleotide (CAAT) repeat motifs that alter lipooligosaccharide (LOS) surface and generate Rd30 resistance to phage HPic.

Survival Benefits of Phase Variation

PV is not only a way of evading host immune responses but is a general strategy of adaptation to multiple stress conditions in a given set of environmental conditions. Survival benefits of PV in aspect of evading host immune system can be explained by a recent report. According to this study, lethality of a specific monoclonal (MabB5) antibody can be neutralised by 8047 strain of N. Meningitides in such a way that the gene encoding lipopolysaccharide glucose transerase enzyme loses a single cystidine residue from homopolymeric tract of 12 Cs present in coding region of the gene that inturn results in ON switch of the mentioned gene by converting the open reading frame from out-of-frame to in-frame. The above-mentioned event then lead to loss of antibody sensitivity by incorporating glucose molecule in the inner core of LPS surface. In case when the expression of IgtG gene is in ON condition, phosphoethanolamine gets incorporated in place of glucose in LPS and the cells become vulnerable to MabB5 monoclonal antibody-mediated killing. Escape from killing can also be facilitated when mutation occurs in mismatch repair gene (mut S) that increases PV rate by enhancing number of cells in population permissive for IgtG gene translation before exposure to the monoclonal antibody. Majority species of Mycoplasma, free living microbes, are pathogenic for both humans and non-human animals by causing persistent infections even in the presence of strong immune responses. These smallest organisms have...
an extraordinary ability of undergoing phase and antigenic variation that alters the cell architecture and hence helps in successful survival of these small genome bearing pathogens.\textsuperscript{29,30}

PV is also responsible for enhancing bacterial virulence and not only determines initial fitness, but also helps in subsequent niche bacterial adaptation. An excellent example of this is variation in 10 out of 31 genes that vary during different stages of infection in murine model and play an important role in H. pylori niche adaptation.\textsuperscript{30} Similarly, PV-mediated expression of fimA gene is another example that is essential for Uropathogenic E. coli (UPEC) to cause urinary tract infection (UTI).\textsuperscript{18} A mutant of UPEC was constructed in which fimA operon expression was locked in ON/OFF state and its virulence was checked in a mouse model. It was found that the locked ON mutants were able to colonise mice kidney and bladder, exhibiting virulence like that of wild strains, while the OFF locked model were not able to colonise kidney and bladder of mice. So it was concluded that PV-mediated expression of fimA is necessary for UPEC pathogenicity.\textsuperscript{31} Citing another example of less pathogenic and less prevalent uropathogen Klebsiella pneumonia, it has been observed that pathogenicity of K. pneumoniae is correlated with reduced ability of turning ON of fim operon expression via PV.\textsuperscript{32} The virulence of N. meningitides, a commensal bacteria present in human nasopharynx, is linked with PV. Sometimes this organism become virulent and enters into the blood stream where it causes sepsis and from blood stream it could pass the blood-brain barrier where it infect brain and causes fatal meningitides. The transition to virulent strains, Hyper virulent strains, is not known much, but nadA protein is found in these strains that help in adhesion and entry of virulent strain into epithelial cells.\textsuperscript{25}

Phase varied expression of modification (mod) gene of Restriction-Modification (R-M) system especially of type III R-M system are naturally competent for foreign DNA uptake and transformation in many bacteria. The active form of R-M systems genes otherwise restrict incoming foreign DNA as these R-M system are used as defence tool. The OFF expression of these genes facilitates the uptake of useful DNA from environment, thus increasing the fitness of host.\textsuperscript{33} Similarly, as stated earlier, PV of R-M system genes also results in regulation of different gene expression that is why it gains the term of “phase varion.”

**Conclusion**

The success of an organism in a given environment depends on how well that organism is adapted and adjusted to that environment. Pathogenic bacteria faces number of threats in its host environment and keeps on evolving different strategies to overcome such threats. However, the host environments in which these bacteria are living also changes and are in constant way of evolution to tackle the foreign invader with improved attack. Simply we can say that there is constant co-evolution between stress and response. PV empowers bacteria to adapt faster in order to deal life-threatening stresses and is a way of producing intra-strain diversity, evading host of immune responses and pathogenic phenotypes.

**Disclaimer:** None.

**Conflict of Interest:** None.

**Source of Funding:** None.

**References**


