Biochemical Society Transactions (2019)
https://doi.org/10.1042/BST20180633

Review Article

Phase-variable bacterial loci: how bacteria gamble to maximise fitness in changing environments

Zachary N. Phillips, Greg Tram, Kate L. Seib and John M. Atack

Institute for Glycomics, Griffith University, Gold Coast, Queensland 4215, Australia

Correspondence: John M. Atack (j.atack@griffith.edu.au)

Phase-variation of genes is defined as the rapid and reversible switching of expression — either ON-OFF switching or the expression of multiple allelic variants. Switching of expression can be achieved by a number of different mechanisms. Phase-variable genes typically encode bacterial surface structures, such as adhesins, pili, and lipooligosaccharide, and provide an extra contingency strategy in small-genome pathogens that may lack the plethora of 'sense-and-respond' gene regulation systems found in other organisms. Many bacterial pathogens also encode phase-variable DNA methyltransferases that control the expression of multiple genes in systems called phasevarions (phase-variable regulons). The presence of phase-variable genes allows a population of bacteria to generate a number of phenotypic variants, some of which may be better suited to either colonising certain host niches, surviving a particular environmental condition and/or evading an immune response. The presence of phase-variable genes complicates the determination of an organism's stably expressed antigenic repertoire; many phase-variable genes are highly immunogenic, and so would be ideal vaccine candidates, but unstable expression due to phase-variation may allow vaccine escape. This review will summarise our current understanding of phase-variable genes that switch expression by a variety of mechanisms, and describe their role in disease and pathobiology.

Introduction

Phase-variable bacterial loci rapidly and reversibly switch their expression. In many small-genome, host-adapted bacterial pathogens, phase-variation serves as an extra contingency strategy to allow adaptation to changing conditions [1], and form part of the 'tinkerer's evolving toolbox' [2]. The mechanisms and processes behind the evolution of these loci have been discussed in excellent detail previously [2,3], and as such this review will focus on the role, and the advantages, of phase-variable gene expression during pathobiology in a number of important human pathogens. Phase-variable genes often encode bacterial cell-surface features, such as adhesins, iron acquisition proteins, pili and lipooligosaccharide (LOS) biosynthetic enzymes [4–9]. The suite of phase-variable genes in a particular species is referred to as the 'phasome' [10]. Variable expression of the phasome within a bacterial population results in a variety of phenotypically distinct individuals which may be better equipped to colonise certain host niches or better able to evade a pre-primed host immune response. For example, expression of an adhesin may be required for initial colonisation of the host, and variants where the adhesin is expressed may be selected for during this stage of infection. However, this adhesin may be highly immunogenic, and its expression may be selected against if the host has a pre-primed immune response, as those variants expressing high levels are killed by the immune system. Modifications to LOS may impart serum resistance and resistance to neutrophil-mediated killing [11,12], but may be selected against in different host niches [8]. The polysaccharide capsule of *Neisseria meningitidis* is absolutely required for resistance to serum and always present in invasive isolates [13], but variants expressing high levels of capsule show decreased adherence to host cells [14], meaning they may be

Received: 14 June 2019
Revised: 2 July 2019
Accepted: 5 July 2019

Version of Record published:
24 July 2019

Biochemical Society Transactions (2019)
https://doi.org/10.1042/BST20180633

less equipped to initially colonise the host. As such, a 'back-and-forth' selection and counter-selection for the different phenotypes resulting from phase-variable gene expression occur as bacteria colonise different niches or encounter different pressures.

The rapid and reversible switching of gene expression means that many antigens encoded by phase-variable genes are often discounted as vaccine candidates. However, under certain circumstances, phase-variable genes can be used as vaccine antigens if they are highly immunogenic or their expression is high during colonisation and/or disease. For example, the NadA protein forms part of the 4CMenB vaccine against *N. meningitidis* serogroup B (licensed as Bexsero) [15], and although it undergoes phase-variation due to DNA repeats in the region upstream of its promoter, it is highly expressed during infection [16], and thus required for a key stage in disease. Therefore, targeting this protein with a vaccine will result in protection against a key stage of disease. Similarly, the adhesin Hia switches between high and low expression states by variation in length of a DNA repeat tract in its promoter region, but is being investigated as a vaccine candidate for non-typeable *Haemophilus influenzae* (NTHi) as it is able to induce high levels of serum anti-Hia antibodies in a Chinchilla model of NTHi disease [17]. Targeting Hia with a vaccine would in theory prevent the initial colonisation of the host as *hia* expression is selected for during a Chinchilla model of host colonisation [4], meaning protection could be achieved by preventing colonisation.

Mechanisms of phase-variation

Genes can phase-vary through a number of genetic mechanisms [1] (Figure 1), including variation in the length of hypermutable simple DNA sequence repeat (SSR) tracts, recombination-mediated shuffling between expressed and silent loci, promoter inversions, and by epigenetic mechanisms [1,18]. SSR tracts are unstable and vary in length through polymerase slippage during replication. Longer SSR tracts exhibit higher rates of phase-variation [19–21]. If an SSR tract is located in the open reading frame (ORF) of a gene, variation in tract length can result in the expression of the gene (ON), or due to a frame-shift mutation downstream of the SSR resulting in a premature stop codon, the gene is not expressed (OFF), or in some cases a truncated protein is expressed. SSR tracts can also be located in the promoter of a gene, where they result in a gradient of protein expression. Alternatively, recombination, or shuffling, between expressed and silent variants of particular loci results in the switching of expression between multiple allelic variants of a single protein. This often occurs via recombination between inverted repeats (IRs) that are present within these loci. This type of phase-variable gene expression is also referred to as antigenic-variation. Furthermore, many bacterial promoters are invertible, which results in ON-OFF switching of their respective genes [22], with promoter inversions often catalysed by an associated recombinase. In addition, differential methylation of DNA at particular target sequences in promoters can lead to up- and down-gene regulation by epigenetic mechanisms, and is dependent on the interaction of the methylated site or the methyltransferase with a regulatory protein at the same site [23]. Phase-variable regulons — phasevarions — control differential regulation of multiple genes through phase-variation of a single gene encoding a methyltransferase [24], adding a further level of complexity to understand gene expression in a number of pathogens. This review will highlight a variety of interesting and well-studied examples of phase-variable genes in bacteria, and the role that phase-variable gene expression contributes to the biology of the organisms containing them. It will also highlight the implications of phase-variation on vaccine development.

Phase-variable expression of genes via genetic mechanisms

Adhesins

Many adhesins found in bacteria are phase-variable [1]. Whilst these are required for colonisation, as their expression is required to adhere to host cells, they are also often highly immunogenic; therefore individuals that have lower levels or no expression of the adhesins are better equipped to survive an immune response. For example, the HMW1 and HMW2 adhesins are found in ~75% of NTHi isolates [25]. Genes encoding both adhesins contain heptanucleotide $TCTTTCA_{(n)}$ repeats in their promoter regions. As the number of $TCTTTCA_{(n)}$ repeats in this SSR increases, the level of expression of HMW1/2 decreases [26]. HMW1/2 are required for binding to related host cell receptors in the human airway [27,28], but they are also highly immunogenic, and are currently being investigated as a candidate for a vaccine against NTHi [29,30].

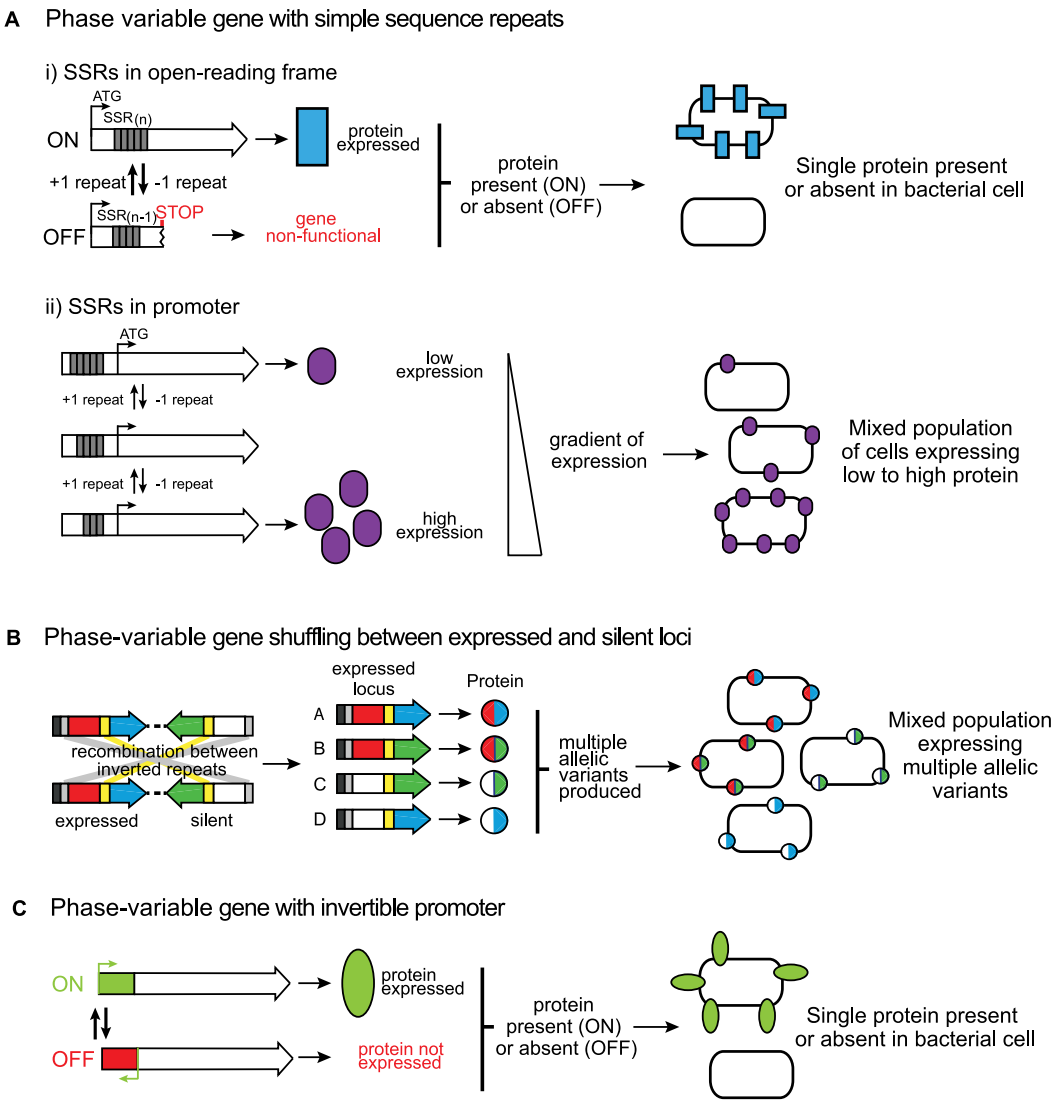

Figure 1. Illustration of the major ways bacterial loci phase-vary.

(**A**) Variation in the number of simple sequence repeats (SSRs; grey boxes) associated with a gene: (i) if the SSR tract is located in the ORF of a gene, this results in ON-OFF switching of expression of the encoded protein. If the number of SSRs leads to the gene remaining in-frame, the gene is expressed (ON). However, gain or loss of repeat units in the SSR tract results in a frame-shift downstream of the SSR tract, and a premature stop codon means the gene is non-functional (OFF). Therefore, the bacterial population contains variants that express the protein or that do not; (ii) alternatively, if the SSR tract is located in the promoter of gene, this can result in graded protein expression, from high expression to low/no expression, with the resulting bacterial population containing a mixture of variants. (**B**) Phase-variable loci can shuffle between expressed (red + blue) and silent (white + green) loci, often by recombination between inverted repeats (IRs; yellow and grey boxes). In the example shown, this results in four different gene variants possible in the expressed locus, meaning four protein allelic variants are expressed, with the resulting bacterial population containing all four protein variants. (**C**) Inversion of a promoter means that the gene it controls is either expressed (promoter in the correct orientation to allow expression; green box), or not expressed, as the promoter is pointing in the opposite orientation (red box). Similar to the ON-OFF switching seen with SSR tracts in an ORF, this results in the bacterial population containing variants expressing the protein, or not.

The adhesin SabA in the human gastric pathogen *Helicobacter pylori* is capable of phase-variation through changes in the length of two SSR tracts associated with the *sabA* gene. The *sabA* gene contains a $CT_{(n)}$ tract in its ORF, and a $T_{(n)}$ tract in its promoter [31]. ON-OFF switching occurs due to variation in the length of the

Biochemical Society Transactions (2019)
https://doi.org/10.1042/BST20180633

$CT_{(n)}$ tract [31], with fine tuning of expression occurring via $T_{(n)}$ tract variation [32]. This fine tuning of *sabA* expression is a result of changes in the DNA structure of the promoter region through T tract length variation, leading to differences in RNA polymerase binding and therefore levels of transcription.

The opacity proteins of *Neisseria gonorrhoeae* phase-vary through both ON-OFF switching and by expression of multiple allelic variants [33]. Expression of allelic variants is the result of the number of $CTCTT_{(n)}$ repeats in the ORF of multiple variable copies of the *opa* genes [33,34]. The selective pressure of the host immune system drives variation of Opa protein expression [35], allowing evasion of pre-primed immune responses against this organism.

Autotransporter proteins

Autotransporters, also called Type V secretion systems, are characterised by an N-terminal passenger domain and a C-terminal transmembrane domain, through which the functional passenger domain is translocated to the bacterial surface [36,37]. Many autotransporters function as adhesins, and a selection for and against their expression during pathobiology would similarly provide advantages to the organisms containing them. The adhesin Hia in NTHi contains a $T_{(n)}$ SSR tract in its promoter, and switches between low and high expression levels due to variation in the length of this tract [4]. Selection for Hia expression occurs during colonisation of the host nasopharynx, and was commensurate with 34T residues present in the $T_{(n)}$ tract. However, Hia is immunogenic, and it was demonstrated that $T_{(n)}$ tract lengths that result in low Hia protein expression levels (30T residues) are selected for during *in vitro* opsonophagocytic killing assays [4]. Thus, selection and counter-selection for Hia protein expression levels, mediated by phase-variation, occur during NTHi colonisation and pathogenesis.

A number of autotransporters in *N. meningitidis* are phase-variable. For example, NadA is a major adhesin in *N. meningitidis*, and forms part of the 4CMenB (Bexsero) vaccine against *N. meningitidis* serogroup B [15]. The *nadA* promoter contains a $TAAA_{(n)}$ repeat, with the number of $TAAA_{(n)}$ repeats affecting the spacing of key regulatory elements [38], and consequently differential expression. NadA has been shown to be highly expressed during infection [16], meaning that NadA could be targeted by a vaccine during disease even though it is phase-variable. MspA is an outer-membrane serine protease in *N. meningitidis* that can also be found in the extracellular medium due to auto-catalytic cleavage that releases the passenger domain. ON-OFF switching of expression of MspA occurs due to variation in the length of a $C_{(n)}$ SSR tract located in the ORF of the *mspA* gene [39]. MspA is expressed at high levels during invasive disease, [39,40] and is required for adhesion to human epithelial and endothelial cells [41]. The AutB adhesin switches ON-OFF due to changes in the length of an $AAGC_{(n)}$ SSR tract in its ORF [42]. Expression of AutB results in increased biofilm formation, but when phase-varied OFF, *N. meningitidis* is able to cross epithelial layers at a higher rate [42]. AutB is highly immunogenic, so switching OFF of expression would also allow evasion of an immune response. Thus, like Hia in NTHi, selection and counter-selection for AutB expression likely occur during meningococcal colonisation and disease progression.

Pili, fimbriae, and flagella

Type IV pili in the pathogenic *Neisseria* can phase-vary by multiple mechanisms. For example, the expression of *pilC*, encoding the putative tip adhesin of the Type IV pili [43], switches ON-OFF by variation in the length of a $G_{(n)}$ tract [44,45]. The major pili protein subunit, encoded by *pilE,* shuffles between multiple allelic variants by recombining with silent variable *pilS* loci [46,47]. *N. gonorrhoeae* has one expressed *pilE* gene and up to 19 silent variable *pilS* genes, distributed in four or five locations across the genome [48]. *N. meningitidis* typically encodes four to eight variable *pilS* sequences, contained in a single locus on the chromosome [49]. Shuffling between *pilE* and *pilS* loci can also lead to non-functional sequences in the *pilE* locus, meaning cells are non-piliated [50]. RecA, Rep, and RecJ proteins, involved in DNA recombination, are all essential for pilin antigenic-variation in *N. meningitidis* [51]. RecA is amongst the seven proteins that are absolutely essential for pilin variation in *N. gonorrhoeae* [48,52], in addition to a guanine quadruplex motif upstream of the *pilE* gene [53]. Interaction between cells expressing different pilin variants within *N. gonorrhoeae* populations results in variable colony formation [54], which could, for example, influence biofilm formation and therefore gonococcal pathogenesis and treatment.

The best-studied example of fimbrial phase-variation occurs in Type I fimbriae of *Escherichia coli* [55]. An invertible DNA element, *fimS*, which contains the *fimA* promoter, is encoded upstream of the major fimbrial

subunit gene, *fimA*. Inversions in *fimS* result in ON-OFF switching of the *fimA* gene [55] and consequently production, or not, of Type I fimbriae.

Clostridium difficile is a major nosocomial pathogen and the cause of potentially fatal colitis. It is able to switch expression of its flagellar ON-OFF via inversions in the DNA associated with the *flaB* gene [56], termed the 'flagellar switch'. Inversions in this promoter also alter production of the toxins TcdA and TcdB by *C. difficile* [57], which could have implications in *C. difficile* pathobiology.

Iron acquisition proteins

N. meningitidis encodes two separate haemoglobin receptors that are phase-variably expressed: HpuAB and HmbR [58]. ON-OFF switching occurs commensurate with the length of a $G_{(n)}$ tract located within the ORF of the *hpuA* and *hmbR* genes [6,58]. HmbR also shows allelic variation between different meningococcal strains, implying that the selection for variants occurs *in vivo* [59]. A strain lacking both HpuAB and HmbR was less virulent in a rat model of infection [60], but not impaired in its growth in human blood [61]. Examination of *in vivo* isolates showed that the majority of strains were phase-ON for *hpuA* or *hmbR*, implying that haemoglobin acquisition is key for systemic disease [62,63].

Haemoglobin and haptoglobin binding proteins are also phase-variable in *H. influenzae*. Genes for the related proteins HgpA, HgpB, and HgpC contain $CCAA_{(n)}$ SSR tracts in their ORFs, and show ON-OFF switching of expression [7]. HgpA is required for full virulence in an infant rat model of invasive disease [64], implying that the selection for HgpA ON would likely occur in invasive *H. influenzae*/NTHi isolates, although this remains to be investigated.

A novel iron acquisition protein, Irp, expressed by the bovine pathogen *Mannheimia haemolytica* has been reported to undergo phase-variation [65]. The *irp* gene contains an $A_{(n)}$ SSR tract, with ON-OFF switching of *irp* expression resulting from changes in the number of adenine residues present [65]. In addition, a stem-loop structure formed by a short IR upstream of this $A_{(n)}$ tract contributes to rates of phase-variation [65], demonstrating a complex mode of phase-variable expression of this iron acquisition protein in *M. haemolytica*.

Lipooligosaccharide biosynthetic enzymes

Lipooligosaccharide (LOS) is a major virulence factor in a number of bacterial pathogens, such as NTHi and the pathogenic *Neisseria*, and has been proposed as a vaccine candidate for *N. gonorrhoeae* [66]. LOS contributes to NTHi survival *in vivo* [8,67]. Many LOS biosynthetic genes contain SSR tracts and are phase-variably expressed [68,69]. Variation of the expression of enzymes required for LOS biosynthesis results in a heterogenous LOS. In NTHi, at least seven LOS biosynthetic genes are phase-variable: *lic1A*, encoding a phosphorylcholine transferase [70], *lic2A* encoding a galactosyltransferase [71], *lic3A* and *lic3B* encoding related sialyltransferases [67,72], l*ex2A* encoding a glucosyltransferase [73], *lgtC* encoding a galactosyltransferase [74], and *oafA* encoding an O-acetyltransferase [12]. Variable expression of these genes is selected for when NTHi colonise or cause disease in humans. For example, the addition of galactose by Lic2A is required for resistance to human serum [75], and protects cells from neutrophil-mediated killing [11], but *lic2A* expression is switched OFF in the majority of invasive NTHi isolates [76]. This indicates a complex role for LOS modified by Lic2A, and demonstrates the fine tuning of phenotype afforded by phase-variable loci. Expression of the O-acetyltransferase OafA is turned OFF in NTHi during middle ear infection [8], but turned ON during invasive disease [76]. During experimental infection of human volunteers with NTHi, both *lex2A* and *lic1A* were shown to switch from OFF to ON during nasopharyngeal colonisation [9]. These findings together show the rapid adaptability to be gained by LOS phase-variation during host colonisation and disease.

C. jejuni, a major human gastric pathogen, also switches the expression of many LOS biosynthetic enzymes due to the presence of SSRs in these genes [77]. The addition of terminal GM1 or GM2 gangliosides is dependent on variation in the length of a $G_{(n)}$ tract in the *wlaN* gene, encoding a beta-1,3 galactosyltransferase [78]. These structures mimic host glycans, allowing *C. jejuni* to evade host immune responses [79], and is also the basis for the auto-immune disease Guillan–Barre syndrome [80]. LOS phase-variation can also result from shuffling between variable biosynthetic loci, leading to 'mosaic' LOS structures, caused by different specificities of the encoded enzymes [81]. *C. jejuni* isolates do not necessarily encode all LOS biosynthetic genes or contain point mutations in one or more biosynthetic loci [82,83], so in addition to switching expression of genes ON-OFF, LOS heterogeneity can occur through lack of functional biosynthetic enzymes.

Biochemical Society Transactions (2019)
https://doi.org/10.1042/BST20180633

Phase-variation of gene expression through epigenetic regulation

Epigenetic regulation by DNA methyltransferases

Epigenetics is the study of heritable gene expression changes that occur without a change in the DNA sequence [84]. DNA methylation at adenine residues is the most common form of epigenetic regulation in bacteria [85]. The DNA methyltransferase Dam (DNA adenine methyltransferase) is a well-studied example of epigenetic regulation in bacteria. Dam regulates genes by methylating DNA or binding and competing with regulatory proteins at specific GATC target sites [23,86]. For example, variable expression of the Pap pilus and antigen 43 (Ag43) in *E. coli* is mediated by Dam methylation at their respective promoters. Methylation of the *pap* promoter by Dam alters the affinity of the LRP regulatory protein for DNA, and results in ON-OFF phase-variable switching of Pap pilus expression [5,87]. Mutants lacking OxyR are locked ON for Ag43 expression, whereas strains lacking Dam are locked OFF for Ag43 expression, implying OxyR competes with Dam for unmethylated GATC sites in the promoter region of the *ag43* gene [88].

Phase-variable DNA methyltransferases

In addition to phase-variable genes encoding surface features, many bacterial pathogens encode cytoplasmic methyltransferases, associated with restriction-modification (R-M) systems, that are subject to phase-variation. Phase-variation of methyltransferase expression results in differential methylation throughout the genome, leading to variable expression of *multiple* genes through epigenetic mechanisms. These systems are called phasevarions (*phase*-*vari*able regul*ons*), and have been described in a number of human-adapted pathogens [89–95]. All phasevarions described to date regulate expression of multiple genes, and many include proteins involved in host colonisation, survival, and pathogenesis, and many regulate putative vaccine candidates. The genes regulated by phasevarions do not contain any identifiable features, and so complicate the identification of an organism's stably expressed protein repertoire. Phasevarions and their role in pathobiology and vaccine development have been described in detail in a number of recent reviews [18,24,96], and as such this review will only briefly describe them.

Many of the phasevarions described to date are controlled by phase-variation of Type III *mod* genes [97]. In these systems, the methyltransferase (Mod) phase-varies between two states (ON or OFF) by variation in the number of SSRs in the encoding *mod* gene [98]. A recent survey of all Type III methyltransferases in REBASE showed that nearly 20% of Type III *mod* genes contain SSRs, are therefore able to phase-vary, and potentially able to control a phasevarion [99]. *mod* genes are highly conserved (>95% DNA sequence identity) in their 5′ and 3′ regions, but contain a highly divergent central domain, the TRD (for Target Recognition Domain). The TRD determines the DNA sequence that is methylated by the Mod methyltransferase. Thus, *mod* genes can exist as multiple allelic variants, due to TRD variation, and which therefore encode enzymes that methylate a *different* DNA target sequence. Methylation of a different DNA target sequence means different Mod proteins regulate the expression of a different suite of genes; i.e. they control different phasevarions. For example, *H. influenzae* contains 21 different *modA* alleles [95]; *Neisseria* species contain 7 *modB* alleles [100]; and *H. pylori* contains 17 *modH* alleles [93].

Many Type I R-M systems contain multiple variable *hsdS* loci that contain IRs, and shuffle between multiple allelic variants. The specificity of Type I R-M systems is dictated by the encoded specificity gene, *hsdS*; if the expressed HsdS protein changes, so does the sequence methylated. Therefore, rather than Type III *mod* genes that reversibly switch ON-OFF, these Type I methyltransferases are always expressed, but as multiple allelic variants, dependent on the sequence of the expressed *hsdS* gene. These phase-variable Type I methyltransferases have been termed 'inverting' systems, as their specificity varies by DNA inversions [101]. Inverting Type I methyltransferases have been described and studied in *Streptococcus pneumoniae* [91] and *Streptococcus suis* [102], and observed in *Listeria monocytogenes* and *Enterococcus faecalis* [101], suggesting these systems are widespread in the bacterial domain. The inverting Type I methyltransferase in *S. pneumoniae* shuffles between six methyltransferase specificities (SpnD39III HsdS alleles A-F) [91], and is in part catalysed by a recombinase encoded at the Type I locus [103]. Phase-variation of the SpnD39III system results in the variable expression of the pneumococcal capsule [91], one of the main virulence factors of this organism. In *S. suis*, the inverting Type I system shuffles between four unique HsdS proteins [102].

A single Type I R-M system that varies by SSR tract variation has been identified in the human pathogen *N. gonorrhoeae* [104]. Variation in length of the SSR tract results in a truncated or full-length HsdS protein

being expressed, resulting in two different methyltransferase specificities [104]. However, apart from the SpnD39III system in *S. pneumoniae*, none of these Type I systems have been shown to result in differential gene regulation commensurate with methyltransferase phase-variation.

Summary/Conclusion

Many host-adapted bacterial pathogens use phase-variation as a strategy to generate phenotypic variation within a population, which allows adaptation to changing host micro-environments, and evasion of the immune system. Phase-variable gene expression is rapid and reversible — variants better able to adapt to particular conditions may be at a disadvantage in others, meaning a selection and counter-selection for variants is continually occurring in a bacterial population as it interacts with the host. A thorough and comprehensive understanding of phase-variable gene expression will allow the generation of improved vaccines and treatments; although it appears counter-intuitive to use proteins that show unstable expression in a vaccine, phase-variable proteins are often highly immunogenic, and if we understand their regulation, and the conditions in which they are required, we will have a further tool in our arsenal to combat many important pathogens.

Perspectives

- **The importance of the field**: Understanding phase-variable genes in bacteria is key to understanding bacterial adaptation to changing conditions within the host, and in determining the stably expressed protein and antigenic repertoire of organisms encoding phase-variable genes.

- **Summary of current thinking**: Although phase-variable genes, and the genes controlled by phasevarions, are not ideal vaccine candidates as their expression is not stable, they can be included in multi-subunit vaccines if their expression is high and/or essential under certain circumstances, or they are highly immunogenic. In order to design stable and efficacious vaccines and treatments, a thorough understanding of the conditions influencing phase-variable gene expression is required, including pro- and anti-selective pressures, *in vivo* niches where the genes may be required, and the mode of phase-variable expression

- **Future directions**: Study of *in vivo* selective pressures will be key for understanding phase-variable gene expression, but these are not usually easily replicated *in vitro*. However, a thorough investigation of phase-variable gene expression using multiple variable experimental conditions will provide the best possible information on which phase-variable genes can be included in rationally designed vaccines. This could potentially lead to the generation of vaccines that contain only highly immunogenic phase-variable proteins, that are essential for important aspects of infection, thereby reducing the chances that all genes are switched OFF during infection. This would decrease the chances of vaccine escape.

Abbreviations

IR, inverted repeat; LOS, lipooligosaccharide; NTHi, non-typeable *Haemophilus influenzae*; ORF, open reading frame; SSR, simple DNA sequence repeat.

Author Contribution

J.M.A., Z.N.P., and G.T. wrote the article and prepared figures. K.L.S. critically read the manuscript, and provided valuable feedback and input

Acknowledgements

This work was supported by the Australian Research Council (ARC) Discovery Project 180100976 to J.M.A.; Australian National Health and Medical Research Council (NHMRC) Project Grant 1099279 to K.L.S. and J.M.A.; and Garnett Passe & Rodney Williams Memorial Foundation (GPRWMF) Grant-In-Aid (Supplementation) to K.L.S. and J.M.A.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References

1 Moxon, R., Bayliss, C. and Hood, D. (2006) Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* **40**, 307–333 https://doi.org/10.1146/annurev.genet.40.110405.090442

2 Moxon, E.R. and Thaler, D.S. (1997) Microbial genetics. The tinkerer's evolving tool-box. *Nature* **387**, 659–662 https://doi.org/10.1038/42607

3 Moxon, E.R., Rainey, P.B., Nowak, M.A. and Lenski, R.E. (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**, 24–33 https://doi.org/10.1016/S0960-9822(00)00005-1

4 Atack, J.M., Winter, L.E., Jurcisek, J.A., Bakaletz, L.O., Barenkamp, S.J. and Jennings, M.P. (2015) Selection and counter-selection of Hia expression reveals a key role for phase-variable expression of this adhesin in infection caused by non-typeable *Haemophilus influenzae*. *J. Infect. Dis.* **212**, 645–653 https://doi.org/10.1093/infdis/jiv103

5 Blyn, L.B., Braaten, B.A. and Low, D.A. (1990) Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J.* **9**, 4045–4054 https://doi.org/10.1002/j.1460-2075.1990.tb07626.x

6 Richardson, A.R. and Stojiljkovic, I. (1999) Hmbr, a hemoglobin-binding outer membrane protein of *Neisseria meningitidis*, undergoes phase variation. *J. Bacteriol.* **181**, 2067–2074 PMID:10094683

7 Ren, Z., Jin, H., Whitby, P.W., Morton, D.J. and Stull, T.L. (1999) Role of CCAA nucleotide repeats in regulation of hemoglobin and hemoglobin-haptoglobin binding protein genes of *Haemophilus influenzae*. *J. Bacteriol.* **181**, 5865–5870 PMID:10482534

8 Fox, K.L., Atack, J.M., Srikhanta, Y.N., Eckert, A., Novotny, L.A., Bakaletz, L.O. et al. (2014) Selection for phase variation of LOS biosynthetic genes frequently occurs in progression of non-typeable *Haemophilus influenzae* infection from the nasopharynx to the middle ear of human patients. *PLoS ONE* **9**, e90505 https://doi.org/10.1371/journal.pone.0090505

9 Poole, J., Foster, E., Chaloner, K., Hunt, J., Jennings, M.P., Bair, T. et al. (2013) Analysis of nontypeable *Haemophilus influenzae* phase variable genes during experimental human nasopharyngeal colonization. *J. Infect. Dis.* **208**, 720–727 https://doi.org/10.1093/infdis/jit240

10 Wanford, J.J., Green, L.R., Aidley, J. and Bayliss, C.D. (2018) Phasome analysis of pathogenic and commensal *Neisseria* species expands the known repertoire of phase variable genes, and highlights common adaptive strategies. *PLoS ONE* **13**, e0196675 https://doi.org/10.1371/journal.pone.0196675

11 Langereis, J.D. and Weiser, J.N. (2014) Shielding of a lipooligosaccharide IgM epitope allows evasion of neutrophil-mediated killing of an invasive strain of nontypeable *Haemophilus influenzae*. *mBio* **5**, e01478-14 https://doi.org/10.1128/mBio.01478-14

12 Fox, K.L., Yildirim, H.H., Deadman, M.E., Schweda, E.K., Moxon, E.R. and Hood, D.W. (2005) Novel lipopolysaccharide biosynthetic genes containing tetranucleotide repeats in *Haemophilus influenzae*, identification of a gene for adding *O*-acetyl groups. *Mol. Microbiol.* **58**, 207–216 https://doi.org/10.1111/j.1365-2958.2005.04814.x

13 Lewis, L.A. and Ram, S. (2014) Meningococcal disease and the complement system. *Virulence* **5**, 98–126 https://doi.org/10.4161/viru.26515

14 Mubaiwa, T.D., Semchenko, E.A., Hartley-Tassell, L.E., Day, C.J., Jennings, M.P. and Seib, K.L. (2017) The sweet side of the pathogenic *Neisseria*: the role of glycan interactions in colonisation and disease. *Pathog. Dis.* **75**, ftx063 https://doi.org/10.1093/femspd/ftx063

15 Giuliani, M.M., Adu-Bobie, J., Comanducci, M., Aricò, B., Savino, S., Santini, L. et al. (2006) A universal vaccine for serogroup B meningococcus. *Proc. Natl Acad. Sci. U.S.A.* **103**, 10834–10839 https://doi.org/10.1073/pnas.0603940103

16 Fagnocchi, L., Biolchi, A., Ferlicca, F., Boccadifuoco, G., Brunelli, B., Brier, S. et al. (2013) Transcriptional regulation of the *nadA* gene in *Neisseria meningitidis* impacts the prediction of coverage of a multicomponent meningococcal serogroup B vaccine. *Infect. Immun.* **81**, 560–569 https://doi.org/10.1128/IAI.01085-12

17 Winter, L.E. and Barenkamp, S.J. (2010) Construction and immunogenicity of recombinant adenovirus vaccines expressing the HMW1, HMW2, or Hia adhesion protein of nontypeable *Haemophilus influenzae*. *Clin. Vaccine Immunol.* **17**, 1567–1575 https://doi.org/10.1128/CVI.00115-10

18 Tan, A., Atack, J.M., Jennings, M.P. and Seib, K.L. (2016) The capricious nature of bacterial pathogens: phasevarions and vaccine development. *Front. Immunol.* **7**, 586 https://doi.org/10.3389/fimmu.2016.00586

19 Bayliss, C.D., Bidmos, F.A., Anjum, A., Manchev, V.T., Richards, R.L., Grossier, J.-P. et al. (2012) Phase variable genes of *Campylobacter jejuni* exhibit high mutation rates and specific mutational patterns but mutability is not the major determinant of population structure during host colonization. *Nucleic Acids Res.* **40**, 5876–5889 https://doi.org/10.1093/nar/gks246

20 Cox, E.C. (1976) Bacterial mutator genes and the control of spontaneous mutation. *Annu. Rev. Genet.* **10**, 135–156 https://doi.org/10.1146/annurev.ge.10.120176.001031

21 Farabaugh, P.J., Schmeissner, U., Hofer, M. and Miller, J.H. (1978) Genetic studies of the lac repressor. *J. Mol. Biol.* **126**, 847–863 https://doi.org/10.1016/0022-2836(78)90023-2

22 Jiang, X., Hall, A.B., Arthur, T.D., Plichta, D.R., Covington, C.T., Poyet, M. et al. (2019) Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. *Science* **363**, 181–187 https://doi.org/10.1126/science.aau5238

23 Casadesús, J. and Low, D.A. (2013) Programmed heterogeneity: epigenetic mechanisms in bacteria. *J. Biol. Chem.* **288**, 13929–13935 https://doi.org/10.1074/jbc.R113.472274

24 Atack, J.M., Tan, A., Bakaletz, L.O., Jennings, M.P. and Seib, K.L. (2018) Phasevarions of bacterial pathogens: methylomics sheds new light on old enemies. *Trends Microbiol.* **26**, 715–726 https://doi.org/10.1016/j.tim.2018.01.008

25 St Geme, III, J.W., Kumar, V.V., Cutter, D. and Barenkamp, S.J. (1998) Prevalence and distribution of the *hmw* and *hia* genes and the HMW and Hia adhesins among genetically diverse strains of nontypeable *Haemophilus influenzae*. *Infect. Immun.* **66**, 364–368 PMID:9423882

26 Dawid, S., Barenkamp, S.J. and St. Geme, J.W. (1999) Variation in expression of the *Haemophilus influenzae* HMW adhesins: a prokaryotic system reminiscent of eukaryotes. *Proc. Natl Acad. Sci. U.S.A.* **96**, 1077–1082 https://doi.org/10.1073/pnas.96.3.1077

27 St Geme, J.W. (1994) The HMW1 adhesin of nontypeable *Haemophilus influenzae* recognizes sialylated glycoprotein receptors on cultured human epithelial cells. *Infect. Immun.* **62**, 3881–3889 PMID:8063405

28 Atack, J.M., Day, C.J., Poole, J., Brockman, K.L., Bakaletz, L.O., Barenkamp, S.J. et al. (2018) The HMW2 adhesin of non-typeable *Haemophilus influenzae* is a human-adapted lectin that mediates high-affinity binding to 2-6 linked N-acetylneuraminic acid glycans. *Biochem. Biophys. Res. Commun.* **503**, 1103–1107 https://doi.org/10.1016/j.bbrc.2018.06.126

29 Winter, L.E. and Barenkamp, S.J. (2003) Human antibodies specific for the high-molecular-weight adhesion proteins of nontypeable *Haemophilus influenzae* mediate opsonophagocytic activity. *Infect. Immun.* **71**, 6884–6891 https://doi.org/10.1128/IAI.71.12.6884-6891.2003
30 Winter, L.E. and Barenkamp, S.J. (2006) Antibodies specific for the high-molecular-weight adhesion proteins of nontypeable *Haemophilus influenzae* are opsonophagocytic for both homologous and heterologous strains. *Clin. Vaccine Immunol.* **13**, 1333–1342 https://doi.org/10.1128/CVI.00221-06
31 Goodwin, A.C., Weinberger, D.M., Ford, C.B., Nelson, J.C., Snider, J.D., Hall, J.D. et al. (2008) Expression of the *Helicobacter pylori* adhesin SabA is controlled via phase variation and the ArsRS signal transduction system. *Microbiology* **154**, 2231–2240 https://doi.org/10.1099/mic.0.2007/016055-0
32 Åberg, A., Gideonsson, P., Vallström, A., Olofsson, A., Öhman, C., Rakhimova, L. et al. (2014) A repetitive DNA element regulates expression of the *Helicobacter pylori* sialic acid binding adhesin by a rheostat-like mechanism. *PLoS Pathog.* **10**, e1004234 https://doi.org/10.1371/journal.ppat.1004234
33 Stern, A., Brown, M., Nickel, P. and Meyer, T.F. (1986) Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47**, 61–71 https://doi.org/10.1016/0092-8674(86)90366-1
34 Bhat, K.S., Gibbs, C.P., Barrera, O., Morrison, S.G., Jähnig, F., Stern, A. et al. (1991) The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes. *Mol. Microbiol.* **5**, 1889–1901 https://doi.org/10.1111/j.1365-2958.1991.tb00813.x
35 Wachter, J. and Hill, S. (2016) Positive selection pressure drives variation on the surface-exposed variable proteins of the pathogenic *Neisseria*. *PLoS ONE* **11**, e0161348 https://doi.org/10.1371/journal.pone.0161348
36 Benz, I. and Schmidt, M.A. (2011) Structures and functions of autotransporter proteins in microbial pathogens. *Int. J. Med. Microbiol.* **301**, 461–468 https://doi.org/10.1016/j.ijmm.2011.03.003
37 Henderson, I.R., Navarro-Garcia, F., Desvaux, M., Fernandez, R.C. and Ala'Aldeen, D. (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol. Mol. Biol. Rev.* **68**, 692–744 https://doi.org/10.1128/MMBR.68.4.692-744.2004
38 Metruccio, M.M.E., Pigozzi, E., Roncarati, D., Berlanda Scorza, F., Norais, N., Hill, S.A. et al. (2009) A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. *PLoS Pathog.* **5**, e1000710 https://doi.org/10.1371/journal.ppat.1000710
39 Oldfield, N.J., Matar, S., Bidmos, F.A., Alamro, M., Neal, K.R., Turner, D.P. et al. (2013) Prevalence and phase variable expression status of two autotransporters, NalP and MspA, in carriage and disease isolates of *Neisseria meningitidis*. *PLoS ONE* **8**, e69746 https://doi.org/10.1371/journal.pone.0069746
40 Echenique-Rivera, H., Muzzi, A., Del Tordello, E., Seib, K.L., Francois, P., Rappuoli, R. et al. (2011) Transcriptome analysis of *Neisseria meningitidis* in human whole blood and mutagenesis studies identify virulence factors involved in blood survival. *PLoS Pathog.* **7**, e1002027 https://doi.org/10.1371/journal.ppat.1002027
41 Turner, D.P., Marietou, A.G., Johnston, L., Ho, K.K., Rogers, A.J., Wooldridge, K.G. et al. (2006) Characterization of MspA, an immunogenic autotransporter protein that mediates adhesion to epithelial and endothelial cells in *Neisseria meningitidis*. *Infect Immun* **74**, 2957–2964 https://doi.org/10.1128/IAI.74.5.2957-2964.2006
42 Arenas, J., Paganelli, F.L., Rodríguez-Castaño, P., Cano-Crespo, S., van der Ende, A., van Putten, J.P.M. et al. (2016) Expression of the gene for autotransporter AutB of *Neisseria meningitidis* affects biofilm formation and epithelial transmigration. *Front. Cell. Infect. Microbiol.* **6**, 162 https://doi.org/10.3389/fcimb.2016.00162
43 Rudel, T., Scheurerpflug, I. and Meyer, T.F. (1995) Neisseria PilC protein identified as type-4 pilus tip-located adhesin. *Nature* **373**, 357–359 https://doi.org/10.1038/373357a0
44 Rytkönen, A., Albiger, B., Hansson-Palo, P., Källström, H., Olcén, P., Fredlund, H. et al. (2004) *Neisseria meningitidis* undergoes PilC phase variation and PilE sequence variation during invasive disease. *J. Infect. Dis.* **189**, 402–409 https://doi.org/10.1086/381271
45 Jonsson, A.B., Nyberg, G. and Normark, S. (1991) Phase variation of gonococcal pili by frameshift mutation in pilC, a novel gene for pilus assembly. *EMBO J.* **10**, 477–488 https://doi.org/10.1002/j.1460-2075.1991.tb07970.x
46 Helm, R.A. and Seifert, H.S. (2010) Frequency and rate of pilin antigenic variation of *Neisseria meningitidis*. *J. Bacteriol.* **192**, 3822–3823 https://doi.org/10.1128/JB.00280-10
47 Seifert, H.S. (1996) Questions about gonococcal pilus phase- and antigenic variation. *Mol. Microbiol.* **21**, 433–440 https://doi.org/10.1111/j.1365-2958.1996.tb02552.x
48 Sechman, E.V., Rohrer, M.S. and Seifert, H.S. (2005) A genetic screen identifies genes and sites involved in pilin antigenic variation in *Neisseria gonorrhoeae*. *Mol. Microbiol.* **57**, 468–483 https://doi.org/10.1111/j.1365-2958.2005.04657.x
49 Wörmann, M.E., Horien, C.L., Bennett, J.S., Jolley, K.A., Maiden, M.C., Tang, C.M. et al. (2014) Sequence, distribution and chromosomal context of class I and class II pilin genes of *Neisseria meningitidis* identified in whole genome sequences. *BMC Genomics* **15**, 253 https://doi.org/10.1186/1471-2164-15-253
50 Criss, A.K., Kline, K.A. and Seifert, H.S. (2005) The frequency and rate of pilin antigenic variation in *Neisseria gonorrhoeae*. *Mol. Microbiol.* **58**, 510–519 https://doi.org/10.1111/j.1365-2958.2005.04838.x
51 Xu, J. and Seifert, H.S. (2018) Analysis of pilin antigenic variation in *Neisseria meningitidis* by next-generation sequencing. *J Bacteriol* **200**, e0045-18 https://doi.org/10.1128/JB.00465-18
52 Mehr, I.J. and Seifert, H.S. (1998) Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation and DNA repair. *Mol. Microbiol.* **30**, 697–710 https://doi.org/10.1046/j.1365-2958.1998.01089.x
53 Cahoon, L.A. and Seifert, H.S. (2009) An alternative DNA structure is necessary for pilin antigenic variation in *Neisseria gonorrhoeae*. *Science* **325**, 764–767 https://doi.org/10.1126/science.1175653
54 Zöllner, R., Oldewurtel, E.R., Kouzel, N. and Maier, B. (2017) Phase and antigenic variation govern competition dynamics through positioning in bacterial colonies. *Sci. Rep.* **7**, 12151 https://doi.org/10.1038/s41598-017-12472-7
55 Abraham, J.M., Freitag, C.S., Clements, J.R. and Eisenstein, B.I. (1985) An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc. Natl Acad. Sci. U.S.A.* **82**, 5724–5727 https://doi.org/10.1073/pnas.82.17.5724
56 Anjuwon-Foster, B.R. and Tamayo, R. (2018) Phase variation of *Clostridium* difficile virulence factors. *Gut Microbes* **9**, 76–83 https://doi.org/10.1080/19490976.2017.1362526
57 Anjuwon-Foster, B.R. and Tamayo, R. (2017) A genetic switch controls the production of flagella and toxins in *Clostridium* difficile. *PLoS Genet.* **13**, e1006701 https://doi.org/10.1371/journal.pgen.1006701

58 Lewis, L.A., Gipson, M., Hartman, K., Ownbey, T., Vaughn, J. and Dyer, D.W. (1999) Phase variation of HpuAB and HmbR, two distinct haemoglobin receptors of *Neisseria meningitidis* DNM2. *Mol. Microbiol.* **32**, 977–989 https://doi.org/10.1046/j.1365-2958.1999.01409.x

59 Evans, N.J., Harrison, O.B., Clow, K., Derrick, J.P., Feavers, I.M. and Maiden, M.C. (2010) Variation and molecular evolution of HmbR, the *Neisseria meningitidis* haemoglobin receptor. *Microbiology* **156**, 1384–1393 https://doi.org/10.1099/mic.0.036475-0

60 Stojiljkovic, I., Hwa, V., de Saint Martin, L., O'Gaora, P., Nassif, X., Heffron, F. et al. (1995) The *Neisseria meningitidis* haemoglobin receptor: its role in iron utilization and virulence. *Mol. Microbiol.* **15**, 531–541 https://doi.org/10.1111/j.1365-2958.1995.tb02266.x

61 Bidmos, F.A., Chan, H., Praekelt, U., Tauseef, I., Ali, Y.M., Kaczmarski, E.B. et al. (2015) Investigation into the antigenic properties and contributions to growth in blood of the meningococcal haemoglobin receptors, HpuAB and HmbR. *PLoS ONE* **10**, e0133855 https://doi.org/10.1371/journal.pone.0133855

62 Tauseef, I., Harrison, O.B., Wooldridge, K.G., Feavers, I.M., Neal, K.R., Gray, S.J. et al. (2011) Influence of the combination and phase variation status of the haemoglobin receptors HmbR and HpuAB on meningococcal virulence. *Microbiology* **157**, 1446–1456 https://doi.org/10.1099/mic.0.046946-0

63 Lucidarme, J., Findlow, J., Chan, H., Feavers, I.M., Gray, S.J., Kaczmarski, E.B. et al. (2013) The distribution and 'in vivo' phase variation status of haemoglobin receptors in invasive meningococcal serogroup B disease: genotypic and phenotypic analysis. *PLoS ONE* **8**, e76932 https://doi.org/10.1371/journal.pone.0076932

64 Seale, T.W., Morton, D.J., Whitby, P.W., Wolf, R., Kosanke, S.D., VanWagoner, T.M. et al. (2006) Complex role of hemoglobin and hemoglobin-haptoglobin binding proteins in *Haemophilus influenzae* virulence in the infant rat model of invasive infection. *Infect. Immun.* **74**, 6213–6225 https://doi.org/10.1128/IAI.00744-06

65 Graham, M.R. and Lo, R.Y. (2002) A putative iron-regulated TonB-dependent receptor of *Mannheimia* (*Pasteurella*) haemolytica A1: possible mechanism for phase variation. *Vet. Microbiol.* **84**, 53–67 https://doi.org/10.1016/S0378-1135(01)00415-1

66 Gulati, S., Shaughnessy, J., Ram, S. and Rice, P.A. (2019) Targeting lipooligosaccharide (LOS) for a gonococcal vaccine. *Front. Immunol.* **10**, 321 https://doi.org/10.3389/fimmu.2019.00321

67 Weiser, J.N., Maskell, D.J., Butler, P.D., Lindberg, A.A. and Moxon, E.R. (1990) Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide. *J. Bacteriol.* **172**, 3304–3309 https://doi.org/10.1128/jb.172.6.3304-3309.1990

68 van Belkum, A., Scherer, S., van Leeuwen, W., Willemse, D., van Alphen, L. and Verbrugh, H. (1997) Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. *Infect. Immun.* **65**, 5017–5027 PMID:9393791

69 van Belkum, A., Scherer, S., van Alphen, L. and Verbrugh, H. (1998) Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.* **62**, 275–293 PMID:9618442

70 Weiser, J.N., Lindberg, A.A., Manning, E.J., Hansen, E.J. and Moxon, E.R. (1989) Identification of a chromosomal locus for expression of lipopolysaccharide epitopes in *Haemophilus influenzae*. *Infect. Immun.* **57**, 3045–3052 PMID:2476397

71 High, N.J., Deadman, M.E. and Moxon, E.R. (1993) The role of a repetitive DNA motif (5′-CAAT-3′) in the variable expression of the *Haemophilus influenzae* lipopolysaccharide epitope α Gal(1-4)β Gal. *Mol. Microbiol.* **9**, 1275–1282 https://doi.org/10.1111/j.1365-2958.1993.tb01257.x

72 Fox, K.L., Cox, A.D., Gilbert, M., Wakarchuk, W.W., Li, J., Makepeace, K. et al. (2006) Identification of a bifunctional lipopolysaccharide sialyltransferase in *Haemophilus influenzae*: incorporation of disialic acid. *J. Biol. Chem.* **281**, 40024–40032 https://doi.org/10.1074/jbc.M602314200

73 Jarosik, G.P. and Hansen, E.J. (1994) Identification of a new locus involved in expression of *Haemophilus influenzae* type B lipooligosaccharide. *Infect. Immun.* **62**, 4861–4867 PMID:7523298

74 Hood, D.W., Deadman, M.E., Jennings, M.P., Bisercic, M., Fleischmann, R.D., Venter, J.C. et al. (1996) DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc. Natl Acad. Sci. U.S.A.* **93**, 11121–11125 https://doi.org/10.1073/pnas.93.20.11121

75 Dixon, K., Bayliss, C.D., Makepeace, K., Moxon, E.R. and Hood, D.W. (2007) Identification of the functional initiation codons of a phase-variable gene of *Haemophilus influenzae*, lic2A, with the potential for differential expression. *J. Bacteriol.* **189**, 511–521 https://doi.org/10.1128/JB.00815-06

76 Phillips, Z.N., Brizuela, C., Jennison, A.V., Staples, M., Grimwood, K., Seib, K.L. et al. (2019) Analysis of invasive non-typeable *Haemophilus influenzae* isolates reveals a selection for the expression state of particular phase-variable lipooligosaccharide biosynthetic genes. *Infect. Immun.* **87**, e00093-19 https://doi.org/10.1128/IAI.00093-19

77 Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D. et al. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665–668 https://doi.org/10.1038/35001088

78 Linton, D., Gilbert, M., Hitchen, P.G., Dell, A., Morris, H.R., Wakarchuk, W.W. et al. (2000) Phase variation of a β-1,3 galactosyltransferase involved in generation of the ganglioside GM1-like lipo-oligosaccharide of *Campylobacter jejuni*. *Mol. Microbiol.* **37**, 501–514 https://doi.org/10.1046/j.1365-2958.2000.02020.x

79 Moran, A.P., Prendergast, M.M. and Appelmelk, B.J. (1996) Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunol. Med. Microbiol.* **16**, 105–115 https://doi.org/10.1111/j.1574-695X.1996.tb00127.x

80 Yuki, N., Taki, T., Inagaki, F., Kasama, T., Takahashi, M., Saito, K. et al. (1993) A bacterium lipopolysaccharide that elicits Guillain-Barre syndrome has a GM1 ganglioside-like structure. *J. Exp. Med.* **178**, 1771–1775 https://doi.org/10.1084/jem.178.5.1771

81 Parker, C.T., Gilbert, M., Yuki, N., Endtz, H.P. and Mandrell, R.E. (2008) Characterization of lipooligosaccharide-biosynthetic loci of *Campylobacter jejuni* reveals new lipooligosaccharide classes: evidence of mosaic organizations. *J. Bacteriol.* **190**, 5681–5689 https://doi.org/10.1128/JB.00254-08

82 Gilbert, M., Karwaski, M.F., Bernatchez, S., Young, N.M., Taboada, E., Michniewicz, J. et al. (2002) The genetic bases for the variation in the lipo-oligosaccharide of the mucosal pathogen, *Campylobacter jejuni*. Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide. *J. Biol. Chem.* **277**, 327–337 https://doi.org/10.1074/jbc.M108452200

83 Parker, C.T., Horn, S.T., Gilbert, M., Miller, W.G., Woodward, D.L. and Mandrell, R.E. (2005) Comparison of *Campylobacter jejuni* lipooligosaccharide biosynthesis loci from a variety of sources. *J. Clin. Microbiol.* **43**, 2771–2781 https://doi.org/10.1128/JCM.43.6.2771-2781.2005

84 Dupont, C., Armant, D.R. and Brenner, C.A. (2009) Epigenetics: definition, mechanisms and clinical perspective. *Semin. Reprod. Med.* **27**, 351–357 https://doi.org/10.1055/s-0029-1237423

85 Blow, M.J., Clark, T.A., Daum, C.G., Deutschbauer, A.M., Fomenkov, A., Fries, R. et al. (2016) The epigenomic landscape of prokaryotes. *PLoS Genet.* **12**, e1005854 https://doi.org/10.1371/journal.pgen.1005854

86 Adhikari, S. and Curtis, P.D. (2016) DNA methyltransferases and epigenetic regulation in bacteria. *FEMS Microbiol. Rev.* **40**, 575–591 https://doi.org/10.1093/femsre/fuw023

87 Hernday, A., Krabbe, M., Braaten, B. and Low, D. (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc. Natl Acad. Sci. U.S.A.* **99**, 16470–16476 https://doi.org/10.1073/pnas.182427199
88 Henderson, I.R. and Owen, P. (1999) The major phase-variable outer membrane protein of *Escherichia coli* structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and oxyR. *J. Bacteriol.* **181**, 2132–2141 PMID:10094691
89 Blakeway, L.V., Power, P.M., Jen, F.E., Worboys, S.R., Boitano, M., Clark, T.A. et al. (2014) Modm DNA methyltransferase methylome analysis reveals a potential role for *Moraxella catarrhalis* phasevarions in otitis media. *FASEB J.* **28**, 5197–5207 https://doi.org/10.1096/fj.14-256578
90 Fox, K.L., Dowideit, S.J., Erwin, A.L., Srikhanta, Y.N., Smith, A.L. and Jennings, M.P. (2007) *Haemophilus influenzae* phasevarions have evolved from type III DNA restriction systems into epigenetic regulators of gene expression. *Nucleic Acids Res.* **35**, 5242–5252 https://doi.org/10.1093/nar/gkm571
91 Manso, A.S., Chai, M.H., Atack, J.M., Furi, L., De Ste Croix, M., Haigh, R. et al. (2014) A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat. Commun.* **5**, 5055 https://doi.org/10.1038/ncomms6055
92 Srikhanta, Y.N., Dowideit, S.J., Edwards, J.L., Falsetta, M.L., Wu, H.-J., Harrison, O.B. et al. (2009) Phasevarions mediate random switching of gene expression in pathogenic *Neisseria*. *PLoS Pathog.* **5**, e1000400 https://doi.org/10.1371/journal.ppat.1000400
93 Srikhanta, Y.N., Gorrell, R.J., Steen, J.A., Gawthorne, J.A., Kwok, T., Grimmond, S.M. et al. (2011) Phasevarion mediated epigenetic gene regulation in *Helicobacter pylori*. *PLoS ONE* **6**, e27569 https://doi.org/10.1371/journal.pone.0027569
94 Srikhanta, Y.N., Maguire, T.L., Stacey, K.J., Grimmond, S.M. and Jennings, M.P. (2005) The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc. Natl Acad. Sci. U.S.A.* **102**, 5547–5551 https://doi.org/10.1073/pnas.0501169102
95 Atack, J.M., Srikhanta, Y.N., Fox, K.L., Jurcisek, J.A., Brockman, K.L., Clark, T.A. et al. (2015) A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable *Haemophilus influenzae*. *Nat. Commun.* **6**, 7828 https://doi.org/10.1038/ncomms8828
96 Phillips, Z.N., Husna, A.U., Jennings, M.P., Seib, K.L. and Atack, J.M. (2019) Phasevarions of bacterial pathogens - phase-variable epigenetic regulators evolving from restriction-modification systems. *Microbiology* https://doi.org/10.1099/mic.0.000805
97 Rao, D.N., Dryden, D.T. and Bheemanaik, S. (2014) Type III restriction-modification enzymes: a historical perspective. *Nucleic Acids Res.* **42**, 45–55 https://doi.org/10.1093/nar/gkt616
98 Srikhanta, Y.N., Fox, K.L. and Jennings, M.P. (2010) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.* **8**, 196–206 https://doi.org/10.1038/nrmicro2283
99 Atack, J.M., Yang, Y., Seib, K.L., Zhou, Y. and Jennings, M.P. (2018) A survey of type III restriction-modification systems reveals numerous, novel epigenetic regulators controlling phase-variable regulons; phasevarions. *Nucleic Acids Res.* **46**, 3532–3542 https://doi.org/10.1093/nar/gky192
100 Tan, A., Hill, D.M., Harrison, O.B., Srikhanta, Y.N., Jennings, M.P., Maiden, M.C. et al. (2016) Distribution of the type III DNA methyltransferases modA, modB and modD among *Neisseria meningitidis* genotypes: implications for gene regulation and virulence. *Sci. Rep.* **6**, 21015 https://doi.org/10.1038/srep21015
101 De Ste Croix, M., Vacca, I., Kwun, M.J., Ralph, J.D., Bentley, S.D., Haigh, R. et al. (2017) Phase-variable methylation and epigenetic regulation by type I restriction–modification systems. *FEMS Microbiol. Rev.* **41**, S3–S15 https://doi.org/10.1093/femsre/fux025
102 Atack, J.M., Weinert, L.A., Tucker, A.W., Husna, A.U., Wileman, T.M., N, F.H. et al. (2018) *Streptococcus* suis contains multiple phase-variable methyltransferases that show a discrete lineage distribution. *Nucleic Acids Res.* **46**, 11466–11476 https://doi.org/10.1093/nar/gky913
103 Li, J., Li, J.-W., Feng, Z., Wang, J., An, H., Liu, Y. et al. (2016) Epigenetic switch driven by DNA inversions dictates phase variation in *Streptococcus pneumoniae*. *PLoS Pathog.* **12**, e1005762 https://doi.org/10.1371/journal.ppat.1005762
104 Adamczyk-Poplawska, M., Lower, M. and Piekarowicz, A. (2011) Deletion of one nucleotide within the homonucleotide tract present in the hsdS gene alters the DNA sequence specificity of type I restriction-modification system NgoAV. *J. Bacteriol.* **193**, 6750–6759 https://doi.org/10.1128/JB.05672-11