Phase and antigenic variation in mycoplasmas

Christine Citti†[1,2], Laurent-Xavier Nouvel[1,2] & Eric Baranowski[1,2]
[1]INRA, UMR 1225, F-31076 Toulouse, France [2]Université de Toulouse, ENVT, UMR 1225, F-31076 Toulouse, France
†Author for correspondence: Tel.: +33 561 193 856 ■ Fax: +33 561 193 273 ■ c.citti@envt.fr

With their reduced genome bound by a single membrane, bacteria of the *Mycoplasma* species represent some of the simplest autonomous life forms. Yet, these minute prokaryotes are able to establish persistent infection in a wide range of hosts, even in the presence of a specific immune response. Clues to their success in host adaptation and survival reside, in part, in a number of gene families that are affected by frequent, stochastic genotypic changes. These genetic events alter the expression, the size and the antigenic structure of abundant surface proteins, thereby creating highly versatile and dynamic surfaces within a clonal population. This phenomenon provides these wall-less pathogens with a means to escape the host immune response and to modulate surface accessibility by masking and unmasking stably expressed components that are essential in host interaction and survival.

Review
Future Microbiology

Mycoplasmas as minimal, successful pathogens

Bacteria of the Class *Mollicutes* have often been portrayed as the simplest self-replicative life form because of their minute cell size, their total lack of a cell wall, the paucity of their metabolic pathways and their reduced genome. These features best describe the *Mollicutes* of the *Mycoplasmataceae* family, which contains one of the smallest free-living organisms, *Mycoplasma genitalium* the genome of which barely contains 500 protein-coding genes [1]. Further subdivided into the *Mycoplasma* and *Ureaplasma* genuses, this family encompasses the vast majority of animal and human species of the *Mollicutes*, with more than 100 well-described species [2–4]. Several of these are of particular concern for the medical and for the veterinary fields. In this review, the term 'Mycoplasma', which usually refers to *Mollicutes*, will be restricted to designate members of the *Mycoplasmataceae* family.

A large number of *Mycoplasma* species are successful pathogens of humans and a wide range of animals (Table 1) in which they often cause chronic infections that result in morbid rather than mortal diseases. These simple bacteria have adopted a parasitic lifestyle and live in close contact to the host, with a predilection for the mucosal surfaces of the respiratory and genital tracts. While current advances in molecular biology, genomics and proteomics continue to take advantage of their reduced genome to understand broad biological concepts, factors involved in mycoplasma-host colonization, virulence and pathogenesis remain largely unknown.

Like many other parasites, mycoplasmas are faced with challenging, complex host environments to which they have to adapt and survive. More classical bacteria have evolved several mechanisms to cope and adjust to changing and hostile surroundings, one of which is based on classical gene regulation where changes in phenotypes are driven by gene activation/repression with feedback loops [5,6]. A drawback of this strategy is that it relies on a repertoire of specific sensing mechanisms that might not be suited to respond to new, challenging threats or conditions. These systems usually involve a cascade of events and consequently a number of regulatory and sensing genes that are costly for the organisms, and appear to be lacking in the mycoplasma reduced genomes. Another strategy is to generate population diversity through high-frequency, random mutation so that at least one variant may survive in a particular context [7,8]. Here, the main drawback would be the need for advanced maintenance and repair systems to prevent deleterious mutations in essential genes, and once again, these processes are limited in mycoplasmas. An alternative that has been adopted by these reduced bacteria is to couple particular gene subsets with rapid and reversible genetic changes that would produce a large number of cell-surface variants in an isogenic population. For this purpose, mycoplasmas have evolved a number of combinatorial genetic mechanisms directed towards phase and size variation of major surface components so that when faced with unpredicted challenges, one variant may prevail.

In the absence of a cell wall, the surface of the mycoplasma membrane constitutes a critical interface in the infectious process, mediating basic functions such as the transport of nutrients as well as more complex interactions with the host cells and the host immune defenses. Therefore, understanding how and when phase

Keywords
■ antigenic variation ■ immune evasion ■ lipoproteins ■ minimal cell ■ mycoplasmas ■ phase variation ■ population dynamics ■ surface diversity

future medicine part of fsg

and antigenic variations occur at this particular location can provide valuable insights into the population dynamics and the strategies used by mycoplasmas to establish successful infections.

In this review, the genetic systems involving related, large gene families and underlying high-frequency surface variation in mycoplasmas will be described and discussed in the context of their biological significance during the infectious process.

Mechanisms governing phase & antigenic variation of surface proteins

Mycoplasmas possess a number of sophisticated systems [9–12] that associate gene families of related, surface lipoproteins with a genetic mechanism that drives surface variation with a high frequency. So far, two types of ON/OFF molecular switches have been described that affect the expression status of particular genes and are either based on spontaneous mutations in regions prone to DNA slippage by nucleotide insertion/deletion in simple sequence repeats (homo- or heteropolymeric nucleotide tracts or short tandem repeats), or on DNA rearrangements involving site-specific recombinase. A third mechanism involves unidirectional (gene conversion) or reciprocal recombination and has a similar impact, although it affects the structure of a particular gene rather than its expression *per se* (Figure 1 & Table 2).

Molecular switches involving DNA slippage (Figure 1A) were first described in the swine pathogen *Mycoplasma hyorhinis* [13,14]. This species possesses a genetic system composed of three to eight genes that are clustered on the chromosome and encode related lipoproteins abundantly expressed at the cell surface, namely the Vlps [15–17]. Each *vlp* gene represents a single transcription unit that contains a polyA tract lying in the promoter region between the -35 and the -10 sequences. When the polyA tract is 17 bp long, the downstream *vlp* is transcribed and translated. Insertion or deletion of a single nucleotide is sufficient to abolish *vlp* transcription, most likely because of the critical contribution of the polyA tract to intrinsic local DNA curvature and to transcription initiation [14]. Overall, this genetic system offers an efficient ON/OFF switch that operates independently for each *vlp* gene in a noncoordinated manner and results in Vlp oscillation in expression within propagating populations, with one cell expressing a single or a combination of any Vlp product [18]. A similar system is present in the avian pathogen *Mycoplasma gallisepticum*: the *vlhA* gene family that comprises approximately 40 related lipoprotein genes distributed within several clusters [19–21]. Phase variation in VlhA expression is also governed at the transcription level, although mutations occur in a trinucleotide repeat (GAA), located at the 5′ end of the promoter of individual *vlhA* genes, upstream of the -35 box. When 12 GAA repeats are present, transcription of the downstream *vlhA* is allowed but the mechanism by which this element controls transcription is not clear. Unlike for the Vlp system, data obtained so far suggest that an unknown mechanism prevents the expression of more than one *vlhA* gene in a given cell [20,22–24].

The second type of molecular switch associated with phase variation in mycoplasmas is based on DNA rearrangements (Figure 1B). Several genetic systems have been described (Table 2) that use a cut-and-paste mechanism,

Table 1. General features of major pathogenic mycoplasma species in which the occurrence of phase variable gene family has been described.

Mycoplasma species	Phylogenetic group	Host tropism	Disease/pathology	Genome accession number
M. capricolum subsp. *capricolum*	Spiroplasma (Mycoides cluster)	Small ruminants	Contagious agalactia	NC_007633
M. hominis	Hominis	Human	Opportunistic, urogenital tract	NC_013511
M. hyorhinis	Hominis	Swine	Respiratory disease, arthritis	NA
M. synoviae	Hominis	Avian	Respiratory disease, arthritis	NC_007294
M. pulmonis	Hominis	Rodent	Respiratory disease	NC_002771
M. agalactiae	Hominis	Small ruminants	Contagious agalactia	NC_009497
M. bovis	Hominis	Large ruminants	Respiratory disease, arthritis, mastitis	NA
Ureaplasma parvum	Pneumoniae	Human	Urogenital infections	NC_002162
M. penetrans	Pneumoniae	Human	Unknown (AIDS-associated)	NC_004432
M. genitalium	Pneumoniae	Human	Genital infections	NC_000908
M. gallisepticum	Pneumoniae	Avian	Respiratory disease	NC_004829

NA: Not applicable.

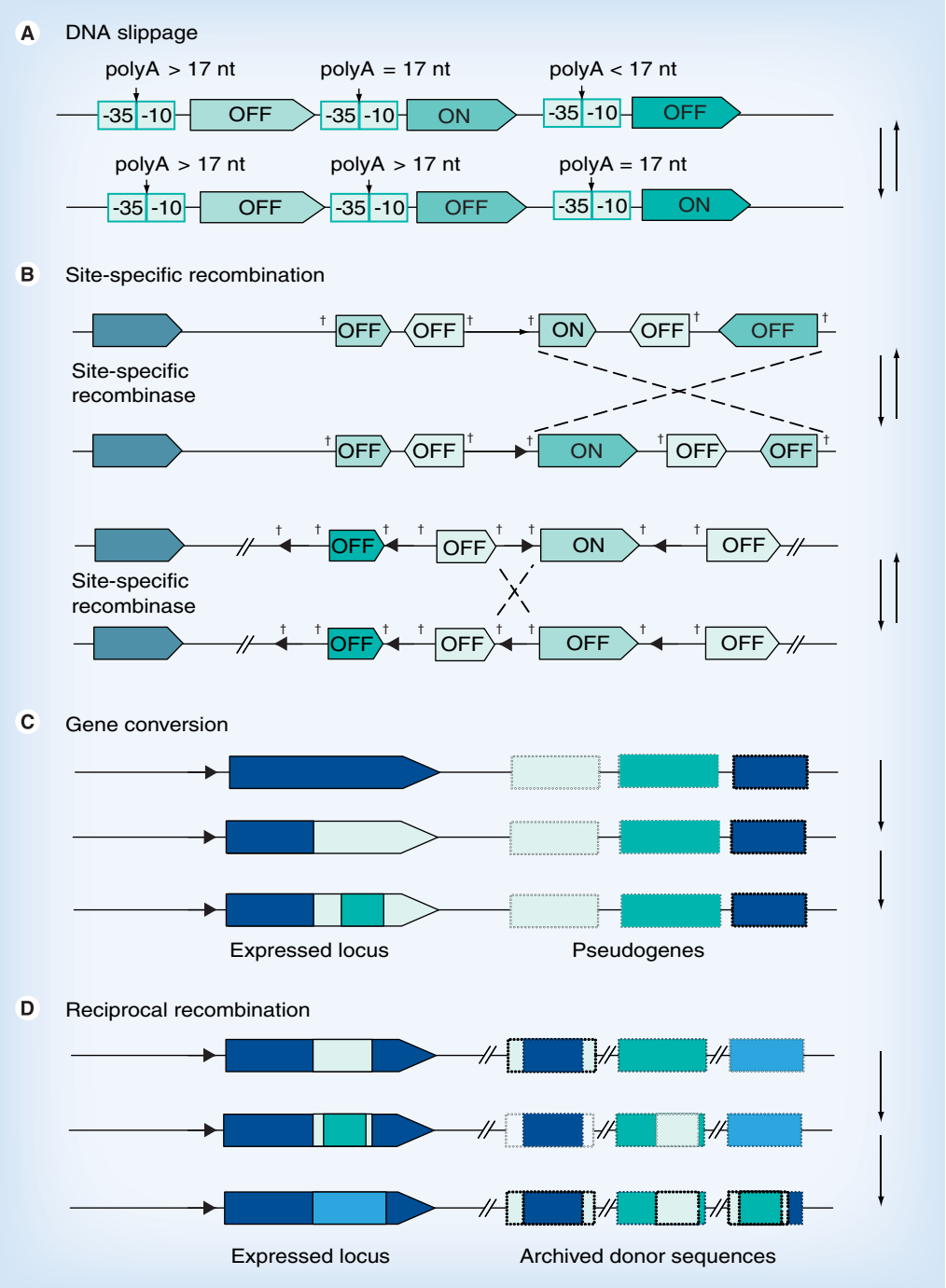

Figure 1. Main genetic mechanisms governing high-frequency phase- and antigenic-variation of mycoplasma surface proteins. (A) ON/OFF switching by DNA slippage using *Mycoplasma hyorhinis vlp* genes as an example. Spontaneous, high-frequency mutations occur in the polyA tract located between the -35 and the -10 promoter region of each *vlp* gene. Transcription of any *vlp* gene only takes place when the polyA is 17 nt long. **(B)** ON–OFF switching by site-specific recombination using the situation encountered in the *vpma* system of *Mycoplasma agalactiae* (top) or the *mpl* system of *Mycoplasma penetrans*. In each system, a recombinase recognizes specific sites and catalyzes DNA rearrangements. **(C)** Antigenic variations by unidirectional recombinations (gene conversion), as occur in *Mycoplasma synoviae* in between an expressed locus and reservoir sequences located in relatively closed clusters. **(D)** Antigenic variations by reciprocal recombinations, as occur in *Mycoplasma genitalium* in between an expressed locus and archived donor sequences distributed all over the chromosome. Large arrows represent gene coding sequences; black arrow heads represent promoters except in **(A)** where promoters are detailed; dotted lines indicate DNA inversion.
†Short sequences targeted by the recombinase.

involving a site-specific recombinase to alternate silent structural genes behind a functional promoter. The *vsp*, the *vsa* and the *vpma* systems of *M. bovis*, *Mycoplasma pulmonis* and *Mycoplasma agalactiae,* respectively, all include clusters of variable genes in which a single gene is placed behind a promoter and is expressed while the others are silent [10,25–30]. All these systems possess a site-specific recombinase that recognizes a specific short DNA sequence located upstream of each variable gene as a target for DNA rearrangements. Indeed, phase-locking of Vpma expression in the *M. agalactiae* type strain was achieved by disrupting the *xer1* recombinase gene located next to the *vpma* locus [31]. In *M. agalactiae*, the number of different Vpma products that can be expressed by any given cell is directly linked to its gene repertoire and, more specifically, to the number of variable gene clusters. For instance, the *M. agalactiae* type strain that contains a single cluster of six *vpma* genes [32] can only express one *vpma* per cell and a total of six *vpma* combinations. By contrast, any given cell of the strain 5632 can express two Vpmas and at least 91 Vpma combinations because it possesses 23 *vpma* genes distributed into two separate clusters, each containing a functional *xer1* gene [33]. Concomitant expression of variable surface proteins also occurs in the human pathogen *Mycoplasma penetrans*, which possesses a family of immunodominant lipoproteins [34] encoded by 38 *mpl* genes distributed in three clusters [35]. Unlike the previous situation encountered in *M. agalactiae*, each *mpl* gene is equipped with its own invertible promoter [36]. ON/OFF switching in expression of each *mpl* gene is driven by flipping of its promoter and one additional difference with the *M. agalactiae* system resides in the location of the recombinase mediating these events, which was found 40 kb away from the nearer *mpl* gene cluster [37].

A third type of genetic mechanisms (Figure 1C) has been described more recently and involves high-frequency unidirectional (gene conversion) or reciprocal gene recombination for generating antigenic variation in *Mycoplasma synoviae* [38] and in *M. genitalium* [39,40], respectively. The *vlhA* family of *M. synoviae* is composed of a single *vlhA* gene, homolog to the *vlhA* genes of *M. gallisepticum* (see above) and of several *vlhA* pseudogenes that lack 5′ end coding sequences. Extensive antigenic variation of the *M. synoviae* VlhA product is the result of unidirectional recombination events occurring between the entire functional *vlhA* copy and the pseudogenes, with duplication of the pseudogene sequence and loss of the corresponding region in the previously expressed gene [38,41]. Unlike *M. gallisepticum,* in which the *vlhA* genes are distributed in several loci around the chromosome, the expressed *vlhA* gene and pseudogenes

Table 2. Main mechanisms governing phase- and antigenic-variation of mycoplasma surfaces.

Type of variation	Genetic event	Mycoplasma species (gene family)	Ref.
ON/OFF switching	DNA slippage involving SSR in promoter regions	*M. hyorhinis* (*vlp*)	[13,14]
		M. gallisepticum (*vlhA*)†	[19,20]
		M. capricolum subsp. *capricolum* (*vmc*)	[55]
	Site-specific recombination (gene rearrangement)	*M. pulmonis* (*vsa*)	[27,28]
		M. bovis (*vsp*)	[25,26]
		M. agalactiae (*vpma*)‡	[29,31,62]
		Ureaplasma parvum (*mba*)	[63]
	Site-specific recombination (promoter inversion)	*M. penetrans* (*mpl*)	[34,36,37]
Size variation	DNA slippage involving short direct repeats within CDSs	*M. agalactiae* (*vpma*)	[33]
		M. hyorhinis (*vlp*)	[13,17,48]
		M. pulmonis (*vsa*)	[64]
		M. bovis (*vsp*)	[65]
		Ureaplasma urealyticum (*mba*)	[66]
Domain shuffling	Gene conversion (unidirectional)	*M. synoviae* (*vlhA*)	[38]
	Reciprocal recombination	*M. genitalium (mgp)*	[39,40]
	DNA recombination§	*M. agalactiae* (*vpma*)	[33]
		M. bovis (*vsp*)	[67]
Other	Gene or locus duplication§	*M. agalactiae* (*vpma*)	[33]

†*The* vlhA *gene family was previously designated as* pMGA.
‡*Some of the* vpma *genes were also designated as* avg [68].
§*Expected to occur at low frequency (rare events).*
CDS: Coding sequence; SSR: Single sequence repeat.

are confined to a 114 kb region suggesting that in this system a relative proximity may facilitate site-specific recombinations [42]. The strategy used by the human pathogen *M. genitalium* for generating antigenic diversity is also based on recombination. The genome of this mycoplasma species contains multiple repeats of sequences termed MgPar that have homology to *mgpB* and *mgpC*, two genes coding immunogenic adhesion proteins. This system is so far unique among bacteria in that it generates high-frequency antigenic variation via reciprocal recombinations between the *mgpB* and *mgpC* genes and the pool of MgPar archived donor sequences [39,40]. Unlike the situation encountered in *M. synoviae*, the MgPar sequences are arranged in several clusters dispersed on the chromosome. Whether differences in variable gene organizations in the two mycoplasma species reflect two distinct genetic mechanisms for unidirectional and reciprocal recombinations is as yet unknown.

Extent of the genetic events contributing to mycoplasma surface variation

The majority of the highly variable components described in mycoplasmas are surface exposed, encoded by multigene families and strongly recognized by the host humoral immune response [10]. All contain a signal peptide followed by a lipobox at their N-terminal while the rest of the molecule is made of blocks of amino acids that can be directly repeated or shared by other members of the family.

These structural features are illustrated by the Vpma family of *M. agalactiae*, for which three strain repertoires have been fully sequenced and account for 18 distinct *vpma* genes recorded in this species (Figure 2) [29,33,43]. The primary amino acid sequence of the mature Vpma products is suggestive of several types of genetic events that contribute to generate diversity, in addition to phase variation. One of these events is the expansion and contraction of the C-terminal sequence, possibly via a DNA slippage mechanism involving short direct repeats as illustrated by the allelic version of VpmaD (Vpma-D1 and -D2) or VpmaF (VpmaF1 and F2). This genetic event results in size variation of particular variable products and has also been reported to occur with a number of other mycoplasma variable systems such as for the Vlp of *M. hyorhinis*, the Vsa of *M. pulmonis* and the Vsp of *M. bovis*. For *M. agalactiae*, the frequency of this event is not known, but in *M. hyorhinis* the occurrence of size variants of individual Vlp in clonal populations is high [18]. Another possible variation that may affect Vpma structure is the shuffling of a number of motifs, most likely by homologous recombination as suggested by comparing the primary amino acid sequences of VpmaU with VpmaZ and VpmaV. A third genetic event contributing to surface diversity is gene duplication followed by antigenic drift resulting in the emergence of allelic versions and new products. For instance VpmaD1, VpmaD2 and VpmaI are likely to be the products of duplication events followed by mutational events that have resulted in size variants of the same gene (Vpma D1 and VpmaD2) and of an allelic version (VpmaI).

Intrastrain differences in the size of variable gene repertoires have been observed for most mycoplasma systems and generate an additional level of diversity. In *M. agalactiae* strain 5632, duplication of seven *vpma* genes at a second locus was responsible for a dramatic increase in surface diversification with 91 possible Vpma combinations compared with only six in the type strain, which only possesses a single six-*vpma* gene cluster [33]. Contraction and expansion of the variable gene repertoires are expected to be rare events and responsible for intrastrain diversity rather than intraclonal variation. However, as shown for *M. agalactiae*, these events may have a drastic impact on the possible antigenic mosaics displayed at the surface by one species.

All these events are superimposed to phase variation of the individual products and result in generating a highly dynamic repertoire of major immunodominant surface proteins. Phase and size variations have been shown to spontaneously occur *in vivo* and *in vitro* with a high frequency (10^{-2} to 10^{-5} events/cell/generation, *in vitro*) and to be reversible with the same frequency so that populations resulting from the expansion of a single cell are never pure (Figure 3A). This phenomenon is best illustrated by colony blotting using specific antibodies directed towards a variable surface protein, which shows the occurrence of multiple sectors as the outcome of ON and OFF switching in expression of the targeted epitope (Figure 3B).

Overall, the capacity of mycoplasmas to produce a large number of surface antigenic mosaics results from the association between contingency genes and multiple genetic events (e.g., ON–OFF switches, size variation, structural variation, and more rarely, gene duplication, gene loss) that occur in a noncoordinated, combinatorial manner (Figure 4). For most systems described in this review, the number of possible variants that can be generated by one particular system is difficult to assess but is expected to be remarkably high. For instance,

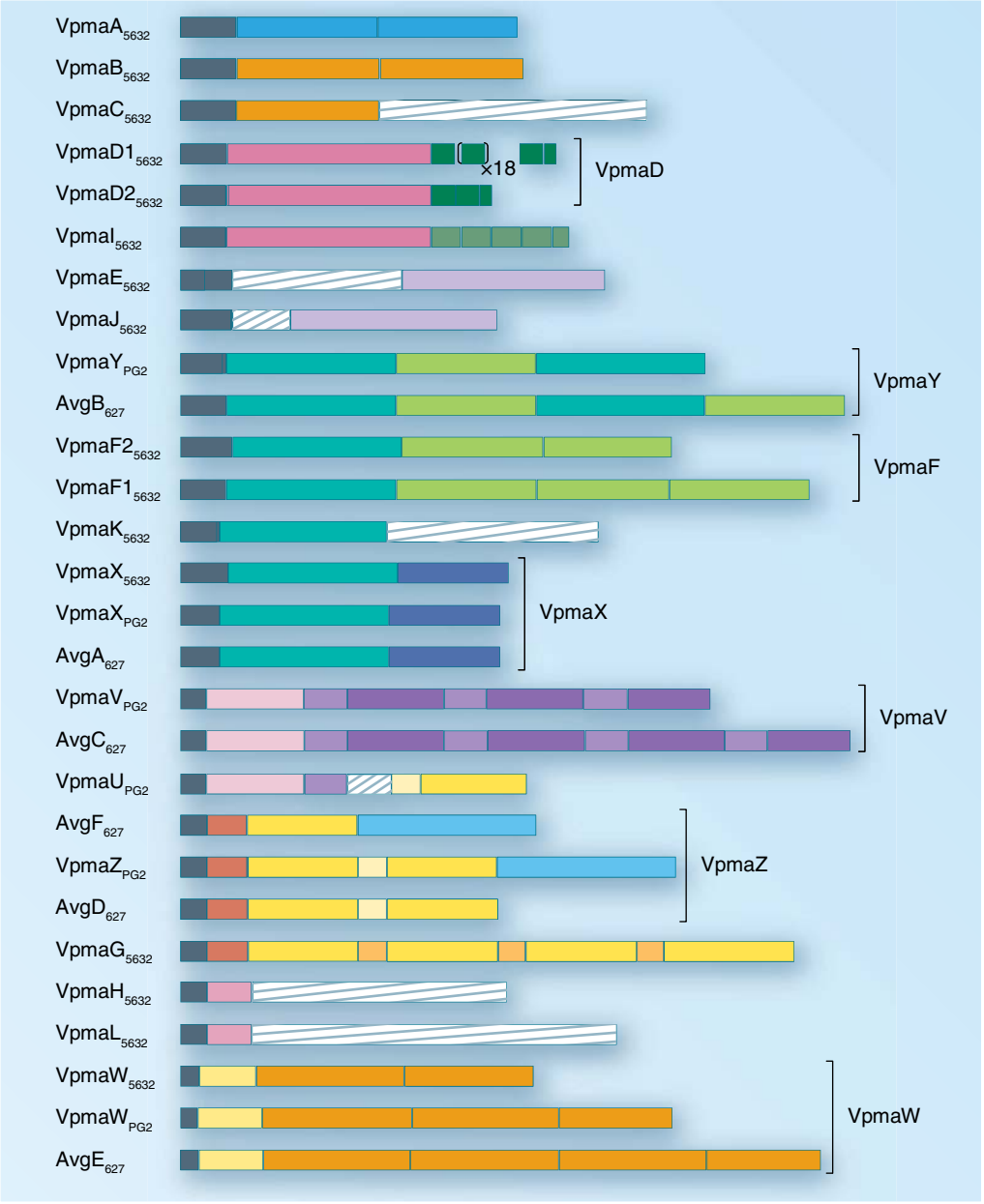

Figure 2. Mosaic structures of the Vpma variable products in *Mycoplasma agalactiae*. Mature Vpma sequences (not including the signal sequence) are shown as long rectangles, with a same color representing highly similar blocks of amino-acids that are either repeated within one Vpma or shared by other Vpmas, while hatched regions correspond to unique sequences. The *vpma* repertoire of *M. agalactiae* has been sequenced in three strains (PG2, 5632 and an isolate, namely isolate 627 [43], in which *vpma* were designated *avg*) and is composed of 28 *vpma* genes that differ by at least one nucleotide. The primary amino acid sequences of the corresponding products revealed only 18 Vpma products, with allelic versions occurring in one (Vpma D, I and F), two (VpamY and V) or three (VmpaX, Z and W) strains.

with approximately 40–70 *vlhA* pseudogenes identified in *M. synoviae*, the diversity generated by this pathogen has been estimated to exceed 10^5 VlhA variants [44]. In addition, *M. genitalium*, with 4% of its genome composed of MgPar reapeated sequences, is thought to generate an almost unlimited number of variants despite its minimized genome [40].

Biological significance of mycoplasma surface variations

After processing of the signal peptide and membrane anchorage via the lipid moiety, mature, phase–variable products are totally exposed at the mycoplasma cell surface and, in absence of a cell wall, are in close contact to the host environment. Interestingly, high-frequency phase and size

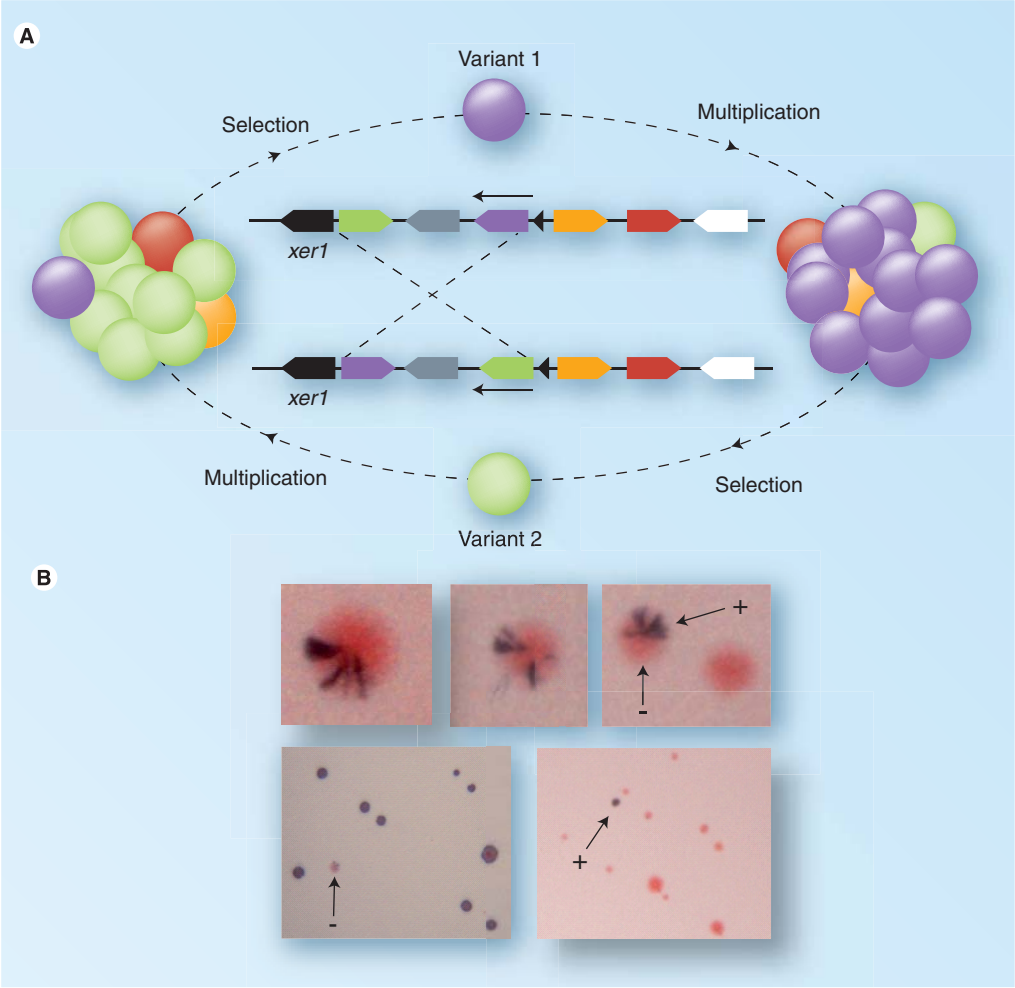

Figure 3. High-frequency phase variation of surface-exposed proteins in isogenic mycoplasma populations. (A) The outcome of frequent DNA rearrangements in generating surface diversity in an isogenic population using the *vpma* system of *Mycoplasma agalactiae* as an example. These events are reversible with the same frequency so that a single clone, regardless of whether it has been selected by a particular condition, has the capacity to produce the full repertoire of possible variants during multiplication. **(B)** Colonies *M. agalactiae* immunostained with Vpma-specific antibodies followed by unspecific incubation in Ponceau Red. Mycoplasma cells reacting positively with the antibodies appear as purple while those stained by the Ponceau Red were not immunostained. Top pictures represent individual colonies displaying sectors as a result of Vpma high-frequency phase variation while those at the bottom represent colonies from two isogenic variants predominantly expressing a different Vpma.

variation of surface components have only been described in *Mollicutes* that colonize immunocompentent hosts (*Mycoplasma* and *Ureaplasma*), and have yet to be reported in those infecting plants or insects (*Spiroplasma* and *Phytoplasma* species). This observation is not surprising taking into account that antigenic variation is an old theme in the field of infectious diseases [8,45] and it is assumed that this phenomenon largely contributes to escaping the host immune response.

Although some species have been shown to establish facultative intracellular residence [46,47], mycoplasmas have to multiply and disseminate in the extracellular compartment of their host and must withstand recognition by the host immune response. More specifically, in the absence of a cell wall, their cell surface must avoid antibody recognition that may interfere with essential membrane-associated functions.

The importance of antigenic variation in providing mycoplasmas with a means of escaping host antibody damage was first supported by experimental evidence from *M. hyorhinis* [48]. Serum antibodies collected from an infected swine were shown to inhibit the growth of *M. hyorhinis* clonal variants expressing a short Vlp while those displaying long versions of the same Vlp were resistant. This behavior

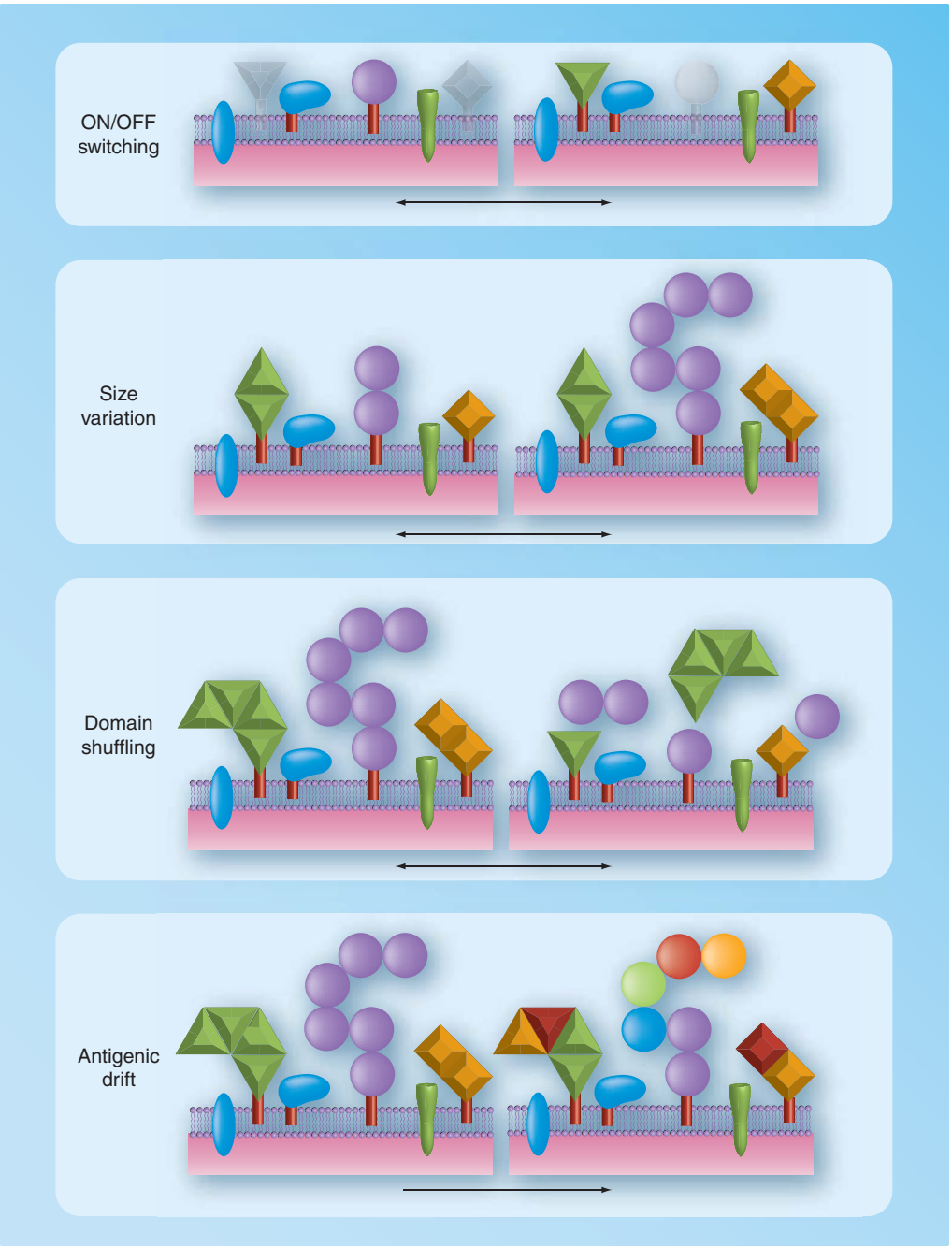

Figure 4. Antigenic variation in mycoplasmas. Major antigenic changes are reported to affect the mycoplasma surface. The wall-less mycoplasma surface is illustrated by a lipid bilayer in which surface-exposed proteins and lipoproteins are anchored via their transmembrane domain or their lipid moieties, respectively. At least four major processes have been shown to modulate mycoplasma surface architecture and antigenic structure. ON/OFF switching is responsible for major antigenic changes occurring at high frequency within mycoplasma clonal populations (10^{-2} to 10^{-5}/cell/generation). This type of variation relates to the stochastic expression of several lipoprotein genes. The modulation of lipoprotein C-terminal domains, leading to changes in protein length at the cell surface, is another common mechanism of antigenic variation in mycoplasmas. ON/OFF switching and size variation of surface lipoproteins are involved in a masking/unmasking processes modulating mycoplasma interactions with the host environment, as well as antibody binding to conserved surface-exposed structures. Additional mechanisms generating surface diversity in mycoplasmas include domain shuffling and antigenic drift. Domain shuffling is the outcome of DNA recombination events occurring within gene families, while antigenic drift results from cumulative changes in lipoprotein amino acid sequences. With the exception of the latter phenomenon, all these changes may occur with a high frequency.

was independent of the presence of complement and of the type of the Vlp expressed by the mycoplasma. Further propagation of susceptible variants in the presence of *M. hyorhinis* host antibodies resulted in the selection of resistant populations that express a long Vlp as a result of a mutation-selection process. Emergence of a Vlp resistant phenotype depends on the most favorable mutation pathway that would rapidly lead to expressing a long Vlp. Depending on the genetic background of the parent, this consists of selecting variants that have either turned on a silent, long *vlp* gene or expanded the number of repeated motifs of an expressed, short *vlp* to elongate its corresponding C-terminal product. Detailed analyses revealed that the Vlps are not the target of the inhibition and that long Vlps are likely to mask critical, as yet unknown, components.

The influence of host antibodies on phase variable products remains to be investigated in many mycoplasma systems but a number of other studies point towards a general role of these sophisticated variable systems in avoidance of the host immune system. This was elegantly shown for the *vsa* gene family of the murine pathogen *M. pulmonis* in which phase variation is driven by site-specific DNA rearrangements [27,49]. More specifically, *in vivo* experiments using wild-type mice versus mice deficient in B and T cells ($RAG^{-/-}$) indicate that the adaptive immune system exerts a selection pressure on Vsa profile shifting [49]. In addition, the temporal antigenic shifts of the *M. genitalium mgpB* and *mgpC* sequences observed in women with persistent infections are consistent with a successive selection of variants driven by host immune selection pressure [39].

Studies performed with *M. pulmonis* also point towards a role for Vsa size variation in several biological processes. More precisely, *M. pulmonis* variants displaying short Vsa, with a few tandem repeats composing their C-terminal end, adsorb to erythrocytes, form biofilms and are efficiently killed by complement while those with long Vsa do not [50–52]. Considering these data and the role played by Vlp size variation in protecting essential components located at the *M. hyorhinis* surface (above), a more encompassing function has been proposed for these variable systems [48,52]. In the absence of a cell wall, abundant related liporoteins, by varying in nature, size and composition, would provide a flexible shield that can rapidly change surface accessibility by masking and unmasking other stable components, thereby altering the overall thickness of the surface architecture. In this generic view, hypervariable lipoprotein families would modulate the interactions between the mycoplasma surface and the host cell, and the host environment, via a series of stochastic variations and selection processes. In addition to their nonspecific roles in avoiding the host immune response and in modulating surface accessibility, highly variable surface proteins may have additional overlapping functions, as suggested for the Vsp of *M. bovis*, which are thought to participate in adhesion to the host cell [53]. However, whether specific Vsp profiles correlate with the mycoplasma's ability to colonize particular host niches is not known.

Advances in genomics & the discovery of new additional, variable systems

Currently, approximately 30 mycoplasma genomes are available in public databases and many more are expected to be sequenced within the coming years. Whether *in silico* analyses will uncover new systems depends, in part, on what is currently known on the mechanisms of antigenic variation and on the gene targeted by these variations.

For instance, *in silico* detection in mycoplasma genomes of homopolymeric nucleotide tracts at the 5′ end of genes encoding putative lipoproteins could contribute to the identification of new putative phase-variable systems. A recent comparative genomic and proteomic analysis of two *M. agalactiae* strains points towards a number of lipoproteins, whose expression might be subjected to phase variation via DNA slippage in intragenic simple sequence repeats [54]. These analyses also identified a new gene family encoding a related, predicted lipoproteins, namely the *drp* genes, members of which are scattered on the chromosome and differently expressed by the two strains, although no clear ON/OFF switching mechanism has been identified. Whether these surface proteins only participate in generating intrastrain diversity or whether they are truly undergoing high-frequency phase- or antigenic-variation in isogenic population remains to be addressed.

Genome-based approaches have previously been successful in defining molecular players involved in phase- or antigenic-variation in mycoplasmas by allowing either the understanding of the intrinsic mechanism of a particular system or the discovery of novel phase-variable genes. For instance in *M. penetrans*, the expression of each *mpl* gene is driven by an invertible promoter (Figure 1). Scanning of the entire *M. penetrans* genome has identified two putative tyrosine site-specific recombinases, one of which was experimentally shown to catalyze the specific promoter inversion [37]. Another example of the contribution of genomics towards

variable systems is the comparison between whole *Mycoplasma capricolum* subsp. *capricolum* and *Mycoplasma mycoides* subsp. *mycoides* Small Colony biotype that revealed phase variation in expression of a six-gene family encoding surface lipoproteins in *M. capricolum* subsp. *capricolum* [55] when combined with proteomic data. Finally, a recent comparative genomic analysis performed with two *M. agalactiae* strains has identified an additional 12 variable *vpma* genes in this species [33].

The ability to discover new variable systems or mechanisms by mining mycoplasma genomes will increase as more variable systems and the underlying genetic switches are characterized in mycoplasmas and other bacteria.

Conclusion

A number of mycoplasma species are able to establish successful infection in a wide range of hosts and to cause diseases that are often chronic, even in the presence of a specific immune response. These persistent infections suggest that the interaction between these simple bacteria and their host is far more complex than first expected. Clues to their success in host colonization and survival may reside in the mycoplasma's ability to rapidly alter the antigenic make up of their surface via a number of sophisticated genetic systems. These complex systems are based on stochastic events that affect the expression and the structure of a number of abundant surface lipoproteins so that variants spontaneously emerge in clonal populations at a rate of 10^{-2} to 10^{-5} events/cell/generation. Several studies suggest that this high surface variability is a key element of the strategy evolved by mycoplasma species for avoiding the host immune response. It is no coincidence that most vaccine trials against mycoplasmas have been disappointing and mycoplasma high-frequency antigenic variation is one major contributor to the complexity in implementing preventive control measures.

With their reduced genome bound by a single membrane, mycoplasmas represent some of the simplest autonomous life forms. The plasticity of their surface architecture might then be more than a means to circumvent the immune response and may constitute a global, stochastic modulator of surface accessibility by masking and unmasking constantly expressed components, some of which might be crucial for mycoplasma survival [48]. While dynamic, phenotypic alteration of the mycoplasma surface architecture may compensate the lack of regulatory system found in more classical bacteria, it may also cover and uncover molecules [56–58] that are engaged in close interaction with the host cell, such as cell adhesion. Phase- and size-variable genes that do not belong to the family encoding surface proteins were not included in this review but several have been characterized and shown to participate in host interactions [59] and, presumably, in modulating environmental responses to nutrients [60]. From these various aspects, mycoplasma phase and antigenic surface variation has emerged in these minimal pathogens as a key contributor to the success of the infection process.

Future perspective

All mechanisms known to generate phase variation in bacteria have been described to occur in mycoplasmas except for that involving DNA site-specific methylation. Genomic data indicate the presence of several DNA methylase genes in these organisms but whether changes in methylation occur as an ON/OFF switch mechanism remains to be investigated.

Ongoing discussions on the random nature of rapid phase, size and antigenic variations in bacteria all point toward the need for re-examining these systems in their natural environments [45,61]. While under steady state conditions, a given variable system will generate variants at a constant rate, the question of whether its intrinsic mechanism can be modulated by environmental changes would be important to assess for mycoplasmas.

Mycoplasma genetic systems dedicated to high-frequency antigenic variation and the host immune system have co-evolved side by side and one intriguing question is whether the immune response contributes to selecting the best-fitted mycoplasma variants, as one adverse outcome of this long-term relationship.

Acknowledgement

We thank E Sagné for excellent technical assistance and P Giammarinaro and E Dordet-Frisoni for critical reading of the manuscript.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Mycoplasmas as minimal, successful pathogens

- Mycoplasmas are often portrayed as the simplest self-replicative life form because of their minute cell size, their total lack of a cell wall, the paucity of their metabolic pathways and their reduced genome.
- Several mycoplasma species are successful pathogens of humans and animals in which they often cause persistent infections of the respiratory and genital tracts.

Mechanisms governing phase & antigenic variation of surface proteins

- Mycoplasmas possess a number of sophisticated systems that associate gene families of related, surface lipoproteins with genetic mechanisms that randomly switch their expression ON or OFF (phase variation) and/or affect their structure with a high frequency.
- Two types of ON/OFF molecular switches have been described and are either based on spontaneous mutations in regions prone to DNA slippage by nucleotides insertion/deletion in simple sequence repeats, or on DNA rearrangements involving site-specific recombinase via a cut-and-paste mechanism.
- A third mechanism involves unidirectional (gene conversion) or reciprocal recombination and affects the structure of particular genes.

Extent of the genetic events contributing to mycoplasma surface variation

- In addition to phase variation, other super-imposed, combinatorial genetic events independently contribute to produce in mycoplasma an almost unlimited number of surface antigenic mosaics (e.g., high-frequency size variation, domain shuffling, gene duplication followed by antigenic drift, gene loss).

Biological significance of mycoplasma surface variations

- In the absence of a cell wall, the mycoplasma's ability to rapidly alter its surface antigen make-up has emerged as a key element for avoiding the host immune response and might be a major contributor to the difficulties in developing vaccines.
- The high variability of the mycoplasma surface might be more than a means to circumvent the immune response and may constitute a global, stochastic modulator of surface accessibility by masking and unmasking steadily expressed components, some of which might be crucial for mycoplasma–host interaction and survival.

Advances in genomics & the discovery of new additional, variable systems

- Currently, approximately 30 mycoplasma genomes are available and data mining has already been successful in discovering novel phase variable genes or underlying intrinsic mechanisms.
- The identifying of new variable genes and the underlying genetic switches will increase as more genome becomes available and as more variable systems are defined in other bacteria.

Bibliography

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1. Fraser CM, Gocayne JD, White O *et al.*: The minimal gene complement of *Mycoplasma genitalium*. *Science* 270, 397–403 (1995).
2. Razin S, Yogev D, Naot Y: Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.* 62, 1094–1156 (1998).
3. Rosengarten R, Citti C, Much P *et al.*: The changing image of mycoplasmas: from innocent bystanders to emerging and reemerging pathogens in human and animal diseases. *Contrib. Microbiol.* 8, 166–185 (2001).
4. Frey J: Mycoplasmas of animals. In: *Molecular Biology and Pathogenicity of Mycoplasmas.* Razin S, Herrmann R (Eds). Kluwer Academic/Plenum Publishers, New York, USA, 73–90 (2002).
5. Rambow-Larsen AA, Petersen EM, Gourley CR, Splitter GA: *Brucella* regulators: self-control in a hostile environment. *Trends Microbiol.* 17, 371–377 (2009).
6. Parker CT, Sperandio V: Cell-to-cell signalling during pathogenesis. *Cell. Microbiol.* 11, 363–369 (2009).
7. Moxon R, Bayliss C, Hood D: Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* 40, 307–333 (2006).

▪▪ **Review of the role of simple sequence repeats in bacteria phase variation.**

8. van der Woude MW, Baumler AJ: Phase and antigenic variation in bacteria. *Clin. Microbiol. Rev.* 17, 581–611, table of contents (2004).

▪▪ **Comprehensive review on phase variation in bacteria.**

9. Wise KS: Adaptive surface variation in mycoplasmas. *Trends Microbiol.* 1, 59–63 (1993).
10. Citti C, Browning GF, Rosengarten R: Phenotypic diversity and cell invasion in host subversion by pathogenic Mycoplasmas. In: *Mycoplasmas: Molecular Biology, Pathogenicity and Strategies for Control*. Blanchard A, Browning GF (Eds). Horizon bioscience, Norfolk, UK, 439–484 (2005).
11. Citti C, Rosengarten R: Mycoplasma genetic variation and its implication for pathogenesis. *Wien Klin Wochenschr.* 109, 562–568 (1997).
12. Rosengarten R, Citti C, Glew M *et al.*: Host-pathogen interactions in *Mycoplasma* pathogenesis: virulence and survival strategies of minimalist prokaryotes. *Int. J. Med. Microbiol.* 290, 15–25 (2000).
13. Yogev D, Rosengarten R, Watson-McKown R, Wise KS: Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5′ regulatory sequences. *EMBO J.* 10, 4069–4079 (1991).
14. Citti C, Wise KS: *Mycoplasma hyorhinis vlp* gene transcription: critical role in phase variation and expression of surface lipoproteins. *Mol. Microbiol.* 18, 649–660 (1995).
15. Rosengarten R, Wise KS: Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. *Science* 247, 315–318 (1990).

▪ **First description of phase variation in mycoplasmas.**

16. Yogev D, Watson-McKown R, Rosengarten R, Im J, Wise KS: Increased structural and combinatorial diversity in an extended family of genes encoding Vlp surface proteins of *Mycoplasma hyorhinis*. *J. Bacteriol.* 177, 5636–5643 (1995).
17. Citti C, Watson-McKown R, Droesse M, Wise KS: Gene families encoding phase- and size-variable surface lipoproteins of *Mycoplasma hyorhinis*. *J. Bacteriol.* 182, 1356–1363 (2000).

18. Rosengarten R, Wise KS: The Vlp system of *Mycoplasma hyorhinis*: combinatorial expression of distinct size variant lipoproteins generating high-frequency surface antigenic variation. *J. Bacteriol.* 173, 4782–4793 (1991).
19. Markham PF, Glew MD, Sykes JE *et al.*: The organisation of the multigene family which encodes the major cell surface protein, pMGA, of *Mycoplasma gallisepticum*. *FEBS Lett*. 352, 347–352 (1994).
20. Glew MD, Baseggio N, Markham PF, Browning GF, Walker ID: Expression of the *pMGA* genes of *Mycoplasma gallisepticum* is controlled by variation in the GAA trinucleotide repeat lengths within the 5′ noncoding regions. *Infect. Immun.* 66, 5833–5841 (1998).
21. Papazisi L, Gorton TS, Kutish G *et al.*: The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R(low). *Microbiology* 149, 2307–2316 (2003).
22. Glew MD, Browning GF, Markham PF, Walker ID: *pMGA* phenotypic variation in *Mycoplasma gallisepticum* occurs *in vivo* and is mediated by trinucleotide repeat length variation. *Infect. Immun.* 68, 6027–6033 (2000).
23. Liu L, Dybvig K, Panangala VS, van Santen VL, French CT: GAA trinucleotide repeat region regulates *M9*/*pMGA* gene expression in *Mycoplasma gallisepticum*. *Infect. Immun.* 68, 871–876 (2000).
24. Liu L, Panangala VS, Dybvig K: Trinucleotide GAA repeats dictate *pMGA* gene expression in *Mycoplasma gallisepticum* by affecting spacing between flanking regions. *J. Bacteriol.* 184, 1335–1339 (2002).
25. Lysnyansky I, Sachse K, Rosenbusch R, Levisohn S, Yogev D: The *vsp* locus of *Mycoplasma bovis*: gene organization and structural features. *J. Bacteriol.* 181, 5734–5741 (1999).
26. Lysnyansky I, Ron Y, Yogev D: Juxtaposition of an active promoter to *vsp* genes via site-specific DNA inversions generates antigenic variation in *Mycoplasma bovis*. *J. Bacteriol.* 183, 5698–5708 (2001).
27. Bhugra B, Voelker LL, Zou N, Yu H, Dybvig K: Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. *Mol. Microbiol.* 18, 703–714 (1995).

▪ **The first description of site-specific DNA inversion as a mechanism for high-frequency phase variation in mycoplasmas.**

28. Shen X, Gumulak J, Yu H *et al.*: Gene rearrangements in the *vsa* locus of *Mycoplasma pulmonis*. *J. Bacteriol.* 182, 2900–2908 (2000).
29. Glew MD, Marenda M, Rosengarten R, Citti C: Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus. *J. Bacteriol.* 184, 5987–5998 (2002).
30. Glew MD, Marenda M, Bergonier D, Rosengarten R, Citti C: The *vpma* gene repertoires in *Mycoplasma agalactiae*. In: *Mycoplasmas of Ruminants: Pathogenicity, Diagnostics, Epidemiology and Molecular Genetics (COST Action 826 - EUR 19693)*. Poveda JB *et al.* (Eds). Office for Official Publications of the European Communities (2001).
31. Chopra-Dewasthaly R, Citti C, Glew MD *et al.*: Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation. *Mol. Microbiol.* 67, 1196–1210 (2008).

▪ **Demonstrates the role of a site-specific recombinase in mediating phase variation in mycoplasmas.**

32. Sirand-Pugnet P, Lartigue C, Marenda M *et al.*: Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS Genet.* 3, e75 (2007).
33. Nouvel LX, Marenda M, Sirand-Pugnet P *et al.*: Occurrence, plasticity, and evolution of the *vpma* gene family, a genetic system devoted to high-frequency surface variation in *Mycoplasma agalactiae*. *J. Bacteriol.* 191, 4111–4121 (2009).
34. Roske K, Blanchard A, Chambaud I *et al.*: Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect. Immun.* 69, 7642–7651 (2001).
35. Sasaki Y, Ishikawa J, Yamashita A *et al.*: The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. *Nucleic Acids Res.* 30, 5293–5300 (2002).
36. Horino A, Sasaki Y, Sasaki T, Kenri T: Multiple promoter inversions generate surface antigenic variation in *Mycoplasma penetrans*. *J. Bacteriol.* 185, 231–242 (2003).

▪ **Describes promoter flip-flop as a mechanism for high-frequency phase variation in mycoplasmas.**

37. Horino A, Kenri T, Sasaki Y, Okamura N, Sasaki T: Identification of a site-specific tyrosine recombinase that mediates promoter inversions of phase-variable *mpl* lipoprotein genes in *Mycoplasma penetrans*. *Microbiology* 155, 1241–1249 (2009).
38. Noormohammadi AH, Markham PF, Kanci A, Whithear KG, Browning GF: A novel mechanism for control of antigenic variation in the haemagglutinin gene family of *Mycoplasma synoviae*. *Mol. Microbiol.* 35, 911–923 (2000).

▪ **First report of gene conversion as a mechanism underlying high-frequency antigenic variation in mycoplasma.**

39. Iverson-Cabral SL, Astete SG, Cohen CR, Totten PA: *mgpB* and *mgpC* sequence diversity in *Mycoplasma genitalium* is generated by segmental reciprocal recombination with repetitive chromosomal sequences. *Mol. Microbiol.* 66, 55–73 (2007).

▪ **First report of reciprocal recombination as a mechanism underlying high-frequency antigenic variation in bacteria.**

40. Ma L, Jensen JS, Myers L *et al.*: *Mycoplasma genitalium*: an efficient strategy to generate genetic variation from a minimal genome. *Mol. Microbiol.* 66, 220–236 (2007).
41. Noormohammadi AH, Markham PF, Duffy MF, Whithear KG, Browning GF: Multigene families encoding the major hemagglutinins in phylogenetically distinct mycoplasmas. *Infect. Immun.* 66, 3470–3475 (1998).
42. Allen JL, Noormohammadi AH, Browning GF: The vlhA loci of *Mycoplasma synoviae* are confined to a restricted region of the genome. *Microbiology* 151, 935–940 (2005).
43. Flitman-Tene R, Mudahi-Orenstein S, Levisohn S, Yogev D: Variable lipoprotein genes of *Mycoplasma agalactiae* are activated *in vivo* by promoter addition via site-specific DNA inversions. *Infect. Immun.* 71, 3821–3830 (2003).
44. Noormohammadi AH: Role of phenotypic diversity in pathogenesis of avian mycoplasmosis. *Avian Pathol.* 36, 439–444 (2007).
45. van der Woude MW: Re-examining the role and random nature of phase variation. *FEMS Microbiol. Lett.* 254, 190–197 (2006).

▪▪ **Re-examination of the biological roles of phase variation in light of the data obtained on commensal and nonpathogenic bacteria, and the integration of this phenomenon into bacterial regulatory networks.**

46. Baseman JB, Lange M, Criscimagna NL, Giron JA, Thomas CA: Interplay between mycoplasmas and host target cells. *Microb. Pathog.* 19, 105–116 (1995).
47. Winner F, Rosengarten R, Citti C: *In vitro* cell invasion of *Mycoplasma gallisepticum*. *Infect. Immun.* 68, 4238–4244 (2000).
48. Citti C, Kim MF, Wise KS: Elongated versions of Vlp surface lipoproteins protect *Mycoplasma hyorhinis* escape variants from growth-inhibiting host antibodies. *Infect. Immun.* 65, 1773–1785 (1997).

▪▪ ***In vitro* demonstration of the role of size- and phase-variation in protecting mycoplasmas against host-antibody damage via masking of essential surface components.**

49. Denison AM, Clapper B, Dybvig K: Avoidance of the host immune system through phase variation in *Mycoplasma pulmonis*. *Infect. Immun.* 73, 2033–2039 (2005).
50. Simmons WL, Bolland JR, Daubenspeck JM, Dybvig K: A stochastic mechanism for biofilm formation by *Mycoplasma pulmonis*. *J. Bacteriol.* 189, 1905–1913 (2007).
51. Simmons WL, Dybvig K: The Vsa proteins modulate susceptibility of *Mycoplasma pulmonis* to complement killing, hemadsorption, and adherence to polystyrene. *Infect. Immun.* 71, 5733–5738 (2003).
52. Simmons WL, Denison AM, Dybvig K: Resistance of *Mycoplasma pulmonis* to complement lysis is dependent on the number of Vsa tandem repeats: shield hypothesis. *Infect. Immun.* 72, 6846–6851 (2004).

▪▪ **Evidence for the role of size variation in modulating the mycoplasma surface.**

53. Sachse K, Helbig JH, Lysnyansky I *et al.*: Epitope mapping of immunogenic and adhesive structures in repetitive domains of *Mycoplasma bovis* variable surface lipoproteins. *Infect. Immun.* 68, 680–687 (2000).
54. Nouvel LX, Sirand-Pugnet P, Marenda MS *et al.*: Comparative genomic and proteomic analyses of two *Mycoplasma agalactiae* strains: clues to the macro- and micro-events that are shaping mycoplasma diversity. *BMC Genomics* 11, 86 (2010).

▪ **Comparative genomics and proteomics analyses highlighting potential new phase variable systems in mycoplasmas.**

55. Wise KS, Foecking MF, Roske K *et al.*: Distinctive repertoire of contingency genes conferring mutation- based phase variation and combinatorial expression of surface lipoproteins in *Mycoplasma capricolum* subsp. *capricolum* of the *Mycoplasma mycoides* phylogenetic cluster. *J. Bacteriol.* 188, 4926–4941 (2006).
56. Zhang Q, Wise KS: Coupled phase-variable expression and epitope masking of selective surface lipoproteins increase surface phenotypic diversity in *Mycoplasma hominis*. *Infect. Immun.* 69, 5177–5181 (2001).
57. Theiss PM, Kim MF, Wise KS: Differential protein expression and surface presentation generate high-frequency antigenic variation in *Mycoplasma fermentans*. *Infect. Immun.* 61, 5123–5128 (1993).
58. Theiss P, Karpas A, Wise KS: Antigenic topology of the P29 surface lipoprotein of *Mycoplasma fermentans*: differential display of epitopes results in high-frequency phase variation. *Infect. Immun.* 64, 1800–1809 (1996).
59. Zhang Q, Wise KS: Localized reversible frameshift mutation in an adhesin gene confers a phase-variable adherence phenotype in mycoplasma. *Mol. Microbiol.* 25, 859–869 (1997).
60. Theiss P, Wise KS: Localized frameshift mutation generates selective, high-frequency phase variation of a surface lipoprotein encoded by a mycoplasma ABC transporter operon. *J. Bacteriol.* 179, 4013–4022 (1997).
61. Bayliss CD: Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. *FEMS Microbiol. Rev.* 33, 504–520 (2009).

▪▪ **Review that places a particular emphasis on the involvement of genome maintenance pathways in determining phase variation rates.**

62. Glew MD, Papazisi L, Poumarat F *et al.*: Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*. *Infect. Immun.* 68, 4539–4548 (2000).
63. Zimmerman CU, Stiedl T, Rosengarten R, Spergser J: Alternate phase variation in expression of two major surface membrane proteins (MBA and UU376) of *Ureaplasma parvum* serovar 3. *FEMS Microbiol. Lett.* 292, 187–193 (2009).
64. Simmons WL, Zuhua C, Glass JI *et al.*: Sequence analysis of the chromosomal region around and within the V-1-encoding gene of *Mycoplasma pulmonis*: evidence for DNA inversion as a mechanism for V-1 variation. *Infect. Immun.* 64, 472–479 (1996).
65. Behrens A, Heller M, Kirchhoff H, Yogev D, Rosengarten R: A family of phase- and size-variant membrane surface lipoprotein antigens (Vsps) of *Mycoplasma bovis*. *Infect. Immun.* 62, 5075–5084 (1994).
66. Teng LJ, Zheng X, Glass JI *et al.*: *Ureaplasma urealyticum* biovar specificity and diversity are encoded in multiple-banded antigen gene. *J. Clin. Microbiol.* 32, 1464–1469 (1994).
67. Lysnyansky I, Ron Y, Sachse K, Yogev D: Intrachromosomal recombination within the *vsp* locus of *Mycoplasma bovis* generates a chimeric variable surface lipoprotein antigen. *Infect. Immun.* 69, 3703–3712 (2001).
68. Flitman-Tene R, Levisohn S, Lysnyansky I, Rapoport E, Yogev D: A chromosomal region of *Mycoplasma agalactiae* containing *vsp*-related genes undergoes *in vivo* rearrangement in naturally infected animals. *FEMS Microbiol. Lett.* 191, 205–212 (2000).