Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens

Kirk W. Deitsch, Sheila A. Lukehart‡ and James R. Stringer§*

Abstract | The complex relationships between infectious organisms and their hosts often reflect the continuing struggle of the pathogen to proliferate and spread to new hosts, and the need of the infected individual to control and potentially eradicate the infecting population. This has led, in the case of mammals and the pathogens that infect them, to an 'arms race', in which the highly adapted mammalian immune system has evolved to control the proliferation of infectious organisms and the pathogens have developed correspondingly complex genetic systems to evade this immune response. We review how bacterial, protozoan and fungal pathogens from distant evolutionary lineages have evolved surprisingly similar mechanisms of antigenic variation to avoid eradication by the host immune system and can therefore maintain persistent infections and ensure their transmission to new hosts.

Antigenic variation
Changes in the antigenic molecules of an invasive organism exposed to the immune system over the course of an infection. This can incorporate mechanisms of phase variation, DNA recombination, epigenetic modifications or mutually exclusive expression.

**Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, New York 10021, USA.*
‡Departments of Medicine and Global Health, University of Washington, Seattle, Washington 98195, USA.
§Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA.
Correspondence to K.W.D.
e-mail: kwd2001@med.cornell.edu
doi:10.1038/nrmicro2145
Published online 8 June 2009

The evolution of mammalian species has resulted in the development of large, multicellular organisms that, in addition to replicating in their own right, also provide an environment for the proliferation of many other species, particularly single-celled organisms that inhabit various niches within, and on the surface of, mammals. It has been estimated that the average human contains tenfold more bacterial cells than human cells[1]. Although the relationship between the host organism and its resident microorganisms is often commensal or symbiotic, many microbial species have evolved to have a detrimental or even lethal effect on their mammalian hosts. Mammals have responded by developing an extremely complex, multifaceted immune system that enables the infected individual to recognize, control and, in some cases, eradicate detrimental microorganisms. The microorganisms have, in turn, evolved correspondingly complex methods for avoiding destruction, resulting in an intricate balance of host–pathogen interactions that we are only beginning to understand.

Infectious microorganisms — viral, bacterial, fungal or protozoan — all face similar challenges when infecting a susceptible host. First, they must avoid mechanical clearance to successfully colonize their preferred tissue or niche, a process that frequently involves the production of specific adhesive molecules that use various host ligands as anchors. In addition, they must either avoid recognition by the immune system through the use of hypervariable surface molecules that allow them to multiply undetected (at least temporarily) or, once recognized, they must be able to avoid destruction by various components of the innate and acquired immune response. This common need to evade the host immune system has resulted in the evolution of remarkably similar survival strategies among pathogens, even those from distant evolutionary lineages. One such strategy is antigenic variation: the capacity of an infecting organism to systematically alter the proteins displayed to the host immune system, such that the host is confronted with a continually changing population that is difficult or impossible to eliminate. The term antigenic variation is generally used to encompass both phase variation (the on–off expression of a particular antigen) and true antigenic variation (the expression of alternative forms of a particular antigen).

Antigenic variation has been extensively studied in a number of microbial systems, leading to the generation of several models of the mechanisms underlying this phenomenon. In recent years, the availability of extensive genome sequence data and improvements in the tools available to study non-model pathogenic organisms have shed new light on old paradigms, providing greater insight into how pathogens avoid the immune systems of their mammalian hosts. In this Review we

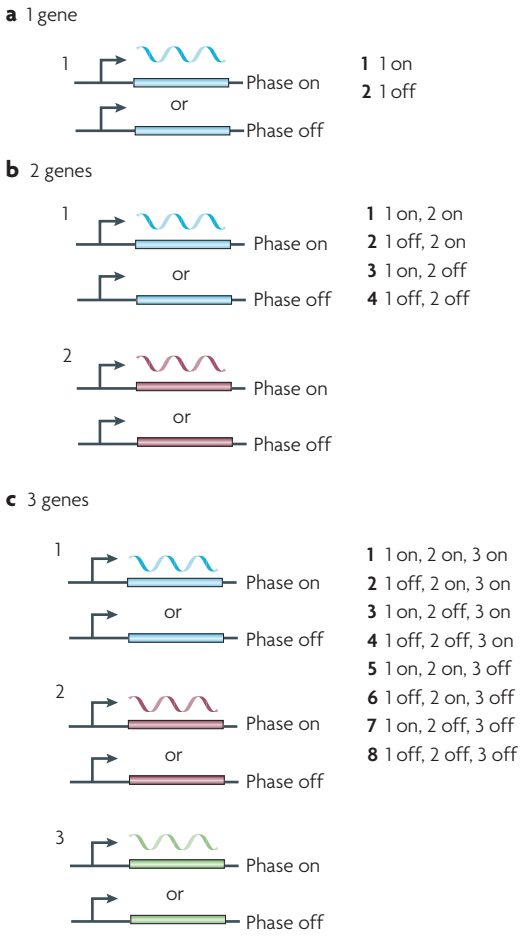

Figure 1 | **Relationship between the number of phase variant genes and the number of phenotypes.** An increase in the number of phase variant genes can contribute to a large increase in the number of possible phenotypes. The total number of phenotypes can be calculated as 2^N (*N* represents the number of independently regulated genes), so in part **a**, where 1 gene is shown, the total number of phenotypes is 2 (2^1), in part **b**, where 2 genes are shown, the total number of phenotypes is 4 (2^2) and in part **c**, where 3 genes are shown, the total number of phenotypes is 8 (2^3).

highlight several recent examples in bacteria, protozoa and fungi which illustrate common themes that are repeatedly observed despite significant evolutionary distance between the various pathogens.

Gene families and variant phenotypes

Antigenic variation in microorganisms is created by two general types of mechanism: genetic mechanisms and epigenetic mechanisms. Genetic events (mutation and recombination) change the DNA sequence of an antigen-encoding gene or its regulatory elements, thereby altering either the level of expression of the gene or the amino acid sequence of the gene product. By contrast, epigenetic mechanisms affect the expression of a gene without altering its primary nucleotide sequence. Whether genetic or epigenetic, the mechanisms underlying antigenic variation described here occur at specific loci, occur frequently and are readily reversible; these features distinguish these systems from antigenic variation caused by random spontaneous mutation, as is more typical of some viruses, such as HIV-1.

A simple form of antigenic variation is often termed phase variation because it was first recognized by observing switching between two alternative phenotypes (phases) among the cells in a clonal population of bacteria. In general, one phase variant differs from another by virtue of the particular cell surface markers expressed (for example, pili), such that a cell surface marker is present in one phase and not in the alternative phase. In some cases, more than one gene in a family can be regulated by phase variation, in which case each gene switches on and off independently. The possible number of phenotypes that can be created by *N* phase-variable genes is therefore 2^N (FIG. 1).

Although as few as 7 phase-variable genes can produce more than 100 different phenotypes, even more phenotypes can be generated by systems that use families of related DNA sequences. Some of these systems display a phenomenon called mutually exclusive expression in which any family member can be expressed but only one member of the family is expressed in a particular microbial cell. The number of different phenotypes is therefore at least as high as the number of family members. However, the number of possible phenotypes can be much greater than the number of family members if recombination occurs among family members, as recombination can produce virtually unlimited diversity through the production of chimeric sequences.

Mechanisms of phase variation

Phase variation in bacteria was first identified in the expression of flagella in *Salmonella*, and a DNA inversion mechanism was implicated[2]. Perhaps the most extensively studied organisms that undergo phase variation, however, are the pathogenic *Neisseria* spp., including *Neisseria meningitidis* and *Neisseria gonorrhoeae*. Both of these species produce surface proteins that function in adherence to host cells and perhaps also in tissue tropism[3–6], in particular pili and the Opa proteins. Antigenic variation mechanisms that were first defined in these bacteria, including phase variation by slipped-strand mispairing during DNA replication and gene conversion, have provided a conceptual framework for many recent advances in our understanding of the mechanisms responsible for antigenic variation in other prokaryotes and eukaryotes. Several recent examples of phase variation mechanisms are described below.

Phase variation through transcriptional regulation. *Treponema pallidum*, the causative agent of syphilis, is a difficult organism to study owing to our inability to culture or genetically manipulate this bacterium. However, recent work has shown that there are many parallels between *T. pallidum* phase variation and the phase variation mechanisms originally described in *Neisseria* spp. The *T. pallidum* repeat (*tpr*) gene family encodes 12 antigens, some of which are predicted to be located in the outer membrane and are expressed during syphilis infection, as shown by the development of humoral[7–9] and

Phase variation
Regulation of gene expression in which an individual gene switches between 'on' and 'off' states. This can be regulated at the level of either transcription initiation or RNA translation.

Epigenetic
Inheritance of particular patterns of gene expression that is not based on changes in DNA sequence. This phenomenon is often associated with DNA modifications (in particular DNA methylation) and/or with alterations in chromatin structure. Post-translational modifications to histones are a well-studied example of chromatin marks associated with epigenetic inheritance.

Mutually exclusive expression
The expression of a single gene from a multicopy gene family. Typically, switches in gene expression do not require DNA recombination and are strictly coordinated so that activation of one gene involves the simultaneous silencing of the previously active gene.

Gene conversion
Also called duplicative transposition. The copying of an entire gene or segment of a gene from one position in the genome to a different position in the genome. The silent copy of the gene is often referred to as the donor, and gene conversion results in its duplication within the genome.

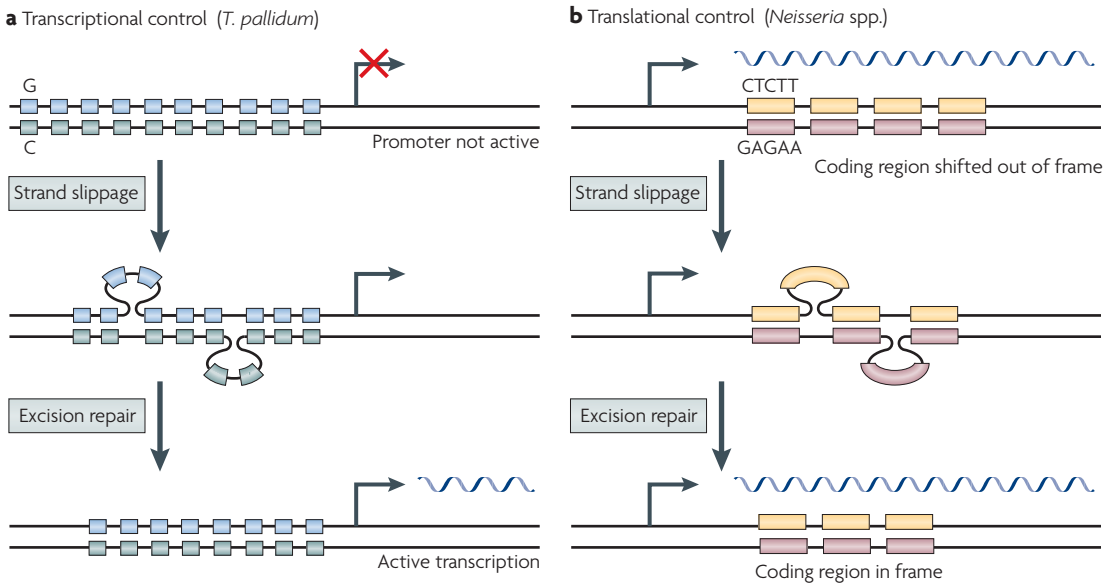

Figure 2 | **Phase variation through slipped-strand mispairing.** During DNA replication, tandem DNA repeats can 'slip', leading to changes in the number of repeats and consequent changes in the transcription or translation of the gene product. **a** | Some *tpr* genes of *Treponema pallidum* are regulated at the level of transcription through changes in the number of Gs found upstream of the transcription start site. **b** | Similar mispairing of CTCTT repeats in the open reading frame can regulate Opa expression in *Neisseria* spp. at the level of translation.

cellular[10] immune responses. The genes encoding subfamilies I (*tprC*, *tprD*, *tprF* and *tprI*) and II (*tprE*, *tprG* and *tprJ*) of the *tpr* family have homopolymeric guanosine (poly-G) repeats of lengths varying from 7 bp to 12 bp immediately upstream of the transcriptional start sites[11,12]. The length of the poly-G tract dramatically affects transcription of the *tprE* gene, and the *tprG/F* and *tprJ/I* operons[12]: loci with poly-G tracts of 8 or fewer Gs are transcribed, whereas transcription of genes with longer poly-G tracts is reduced by ~95–100% (FIG. 2a). Changes in the number of G repeats are thought to occur through slipped-strand mispairing during DNA replication, which closely parallels the mechanism originally described for regulating the transcription of PilC (poly-C)[13,14] and Por (poly-G)[15] in *Neisseria* spp.

Phase variation can also result from epigenetic changes, and examples are found in both bacterial and eukaryotic pathogens. For example, the expression of pili on the surface of pathogenic *Escherichia coli* allows adhesion in the urogenital tract and is dependent on the expression of the *pap* operon. Transcription of this operon is regulated by methylation of specific sites in the regulatory region of the locus[16,17], and the activity of the methyltransferase at these specific sites is influenced by the richness of the media in which the bacteria are grown[18]. When the bacteria are grown in an environment rich in amino acids (such as human urine), methylation rates are increased and switching to the 'phase on' phenotype is favoured[19], which facilitates attachment of the bacteria to the bladder epithelium. The eukaryotic pathogen *Candida glabrata* can similarly respond to its environment through expression of the gene *epa6*, which encodes a surface protein that binds to receptors on the epithelium of the genital tract[20]. In this case, however, the cells are reacting to the low level of NAD^+ found in urine, which influences the activity of the NAD^+-dependent histone deacetylase Sir2, which in turn results in an altered chromatin structure and activation of the *epa6* gene[20]. Thus, both *E. coli* and *C. glabrata* have independently evolved the capacity to sense the environment of the genito-urinary tract and respond epigenetically to express the appropriate surface antigens that facilitate infection.

Phase variation through translational regulation. Phase variation can also occur at the translational level in pathogens using various mechanisms, including slipped-strand mispairing, early ribosome dissociation and mRNA instability. Phase variation regulated at the level of translation was first described as a consequence of slipped-strand mispairing of CTCTT repeat units in the *opa* genes of neisseriae[21,22] (FIG. 2b). The phase 'on' state allows these bacteria to specifically adhere to certain tissues through attachment to specific surface receptors on host cells. Recently, phase variation mediated by altered translational efficiency was similarly shown in several eukaryotic pathogens, again demonstrating how frequently similar mechanisms have evolved independently in distant evolutionary lineages. For example, a mechanism that is remarkably similar to *opa* phase variation in neisseriae is displayed by the fungus *Pneumocystis carinii*, an organism that lives in the lungs of rats. The genome of *P. carinii* is AT-rich, yet long poly-G tracts occur in the middle of some *msr* genes, which belong to a large family of genes that encode variable surface glycoproteins[23]. The number of GC base pairs in these

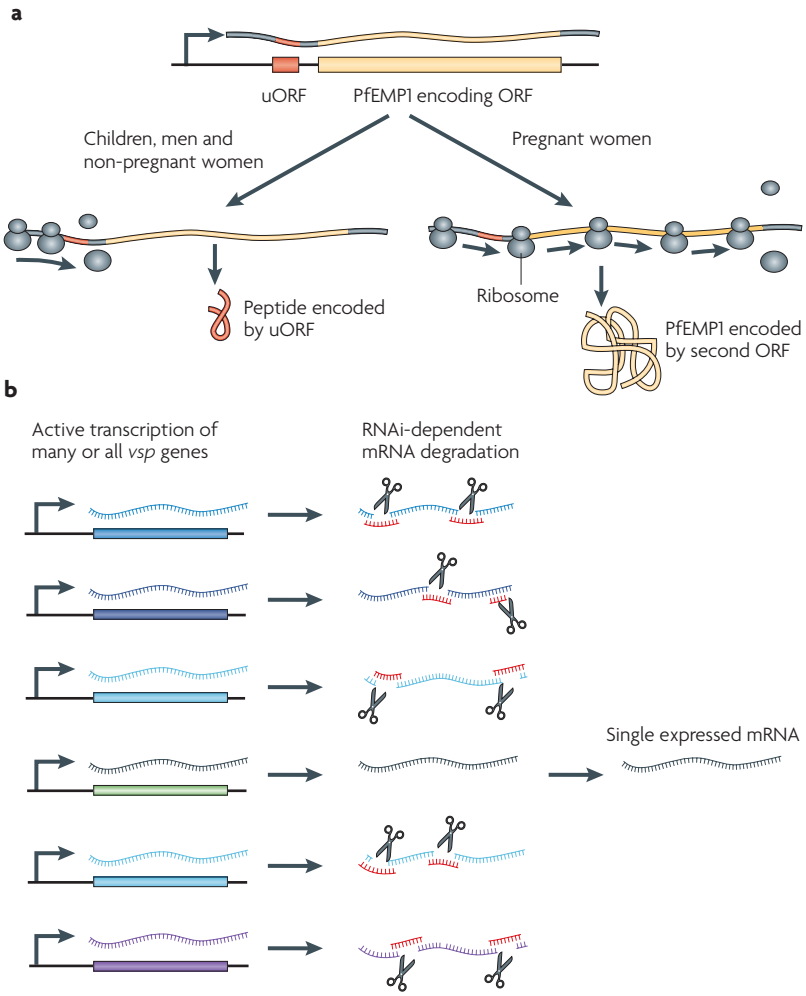

Figure 3 | **Regulation of phase variation at the level of mRNA translation.** **a** | The *var2csa* gene of *Plasmodium falciparum* is transcribed into an mRNA with two open reading frames (ORFs). The mRNA is bound by the large and small subunits of the ribosome, which moves along the transcript (arrows) until it initiates translation at a start codon. Only the upstream ORF (uORF) is translated when the parasites infect children, men or non-pregnant women; the ribosome is thought to dissociate from the mRNA before reaching the second ORF, preventing expression of the encoded form of *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). The second ORF is translated only when parasites infect pregnant women, and the encoded form of PfEMP1 functions only in the presence of a placenta. **b** | Mutually exclusive expression of the *vsp* genes of *Giardia lamblia* is achieved through the RNA interference (RNAi) pathway. Small RNAs (red) target the RNAi degradation machinery (scissors) to mRNAs from all but a single *vsp* genes. Only mRNA from this gene escapes degradation and is translated into protein, although the mechanism that enables this escape is unknown.

tracts seems to vary as a result of spontaneous slipped-strand mispairing events, which cause frameshift mutations that truncate the MSR protein and thereby contribute to antigenic variation in *P. carinii* populations.

Plasmodium falciparum is a protozoan parasite that invades the circulating red blood cells of its human host, causing malaria. These parasites express the highly variable *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), which is expressed on the surface of parasitized red blood cells. PfEMP1 binds tightly to host endothelial cell receptors, effectively removing the infected cells from the circulation and thereby avoiding immune clearance by the spleen[24]. PfEMP1 is encoded by a large, polymorphic gene family called *var*, with each individual *var* gene encoding a different form of PfEMP1. The *var* genes are expressed one at a time, and therefore transcriptional switching between genes leads to antigenic variation. However, one particular *var* gene, *var2csa*, encodes a form of PfEMP1 that specifically binds to a receptor found in the placenta[25]. To avoid expressing this protein when no placenta is available (that is, when infecting men, children or non-pregnant women), the transcript derived from this gene is also subject to translational regulation, and this results in a similar type of phase variation (either 'on' or 'off') to that observed in bacterial systems. In the case of *var2csa*, phase variation is mediated by a small upstream open reading frame (upstream ORF; uORF) that exists between the 5′ end of the message and the start methionine of the PfEMP1 coding region[26]. Translation of the uORF causes dissociation of the ribosome from the mRNA before the ribosome reaches the PfEMP1 start codon, which silences expression of the protein (FIG. 3a). It is not yet known how this repression is reversed when parasites infect pregnant women, but it might involve alternative phosphorylation of a ribosomal initiation factor, as has been observed in yeast[27].

Giardia lamblia is a protozoan parasite that can infect the digestive tracts of most mammals, causing severe diarrhoea in susceptible individuals. The primary antigen exposed on its surface is called variant-specific surface protein (VSP), and the genome contains many genes encoding different forms of VSP. Recent work shows that parasites actively transcribe multiple members of the *vsp* gene family; however, most of these transcripts are degraded through the RNA interference pathway and are therefore maintained in the equivalent of the phase 'off' state[28]. Only transcripts from a single gene are stable and are translated into protein, and the stable transcripts change over the course of an infection, leading to antigenic variation. The mechanism that maintains the stability of a single transcript is unknown (FIG. 3b).

Large, hypervariable gene families

Phase variation is a simple method for avoiding immune recognition, yet many microorganisms have evolved more sophisticated systems that rely on large, multi-copy repertoires of sequences in which each individual sequence encodes an antigenically distinct surface protein. These repertoires can consist of either large families of similar, complete genes that are expressed singly; or, alternatively, they can comprise multiple copies of partial gene sequences that can be recombined into a single expression site. Either way, sequence repertoires provide the capacity to express many functionally similar but antigenically distinct surface proteins that can be expressed at different times over the course of an infection.

Large families of antigen-encoding sequences are observed in bacterial, protozoan and fungal pathogens (TABLE 1), indicating that expansion of the number of antigen-encoding genes is a common evolutionary

Table 1 | **Multicopy sequence repertoires encoding variant surface antigens**

Species	Gene	Copy number	Encoded surface antigen
Bacterial systems			
Borrelia hermsii	*vsp*	~30	Variable short protein
	vlp	~30	Variable long protein
Neisseria meningitidis	*pil*	~19	Type IV pili
	opa	4–5	Opacity proteins
Neisseria gonorrhoeae	*pil*	~19	Type IV pili
	opa	11	Opacity proteins
Treponema pallidum	*tpr*	12	*T. pallidum* repeat proteins A–L; some are predicted to be surface exposed
Protozoan systems			
Trypanosoma brucei	*vsg*	~1,000	Variant surface glycoprotein
Babesia bovis	*vesα* and *vesβ*	~130–160	Variant expressed surface antigen
Giardia lamblia	vsp	~150	Variant surface protein
Plasmodium falciparum	var	~60	*P. falciparum* erythrocyte membrane protein 1
Fungal systems			
Pneumocystis carinii	*msg*	~85	Major surface glycoprotein
	msr	~85	Major surface glycoprotein-related protein
	prt	~85	Protease 1
Candida albicans	*als*	9	Agglutinin-like sequence
Candida glabrata	*epa*	17–23	Epithelial adhesin

strategy. Classic examples of such repertoires in bacteria are typified by the subtelomeric *vsp* and *vlp* genes of *Borrelia hermsii* and the *pil* genes of *Neisseria* species[3,4]. Similarly, the family of *tpr* genes in *Treponema* spp. has expanded to include many similar but distinct copies, with individual members of subfamilies I and II containing substantial sequence variation in the central regions[29]. The prototypical example in protozoan parasites is the *vsg* gene family of *Trypanosoma brucei*, which has undergone a massive expansion to ~1,000 members in each parasite genome[30]. Smaller but similar hypervariable gene families are observed in the genomes of *Plasmodium* species[31–33], *Giardia* species[34] and *Babesia* species[35]. In the fungal pathogen *P. carinii*, genes from multicopy families are typically found in head-to-tail arrays adjacent to telomeres and include members of the *msg*, *msr* and *prt* gene families[23]. The *als* and *epa* genes of the distantly related fungi *Candida albicans* and *C. glabrata*, respectively, are similarly arranged in subtelomeric domains[36]. Several different mechanisms have evolved to control expression of these large families, some of which remain poorly understood.

Mechanisms involving DNA recombination. DNA recombination that leads to the movement of non-expressed coding regions into defined genomic expression sites is one of the most common molecular mechanisms used by both bacterial and eukaryotic pathogens to produce antigenic variation (FIG. 4a). The recombination event can involve the movement of entire genes or, alternatively, small portions of genes to create chimeric sequences. This recombination can be duplicative (gene conversion) or reciprocal. The best-studied bacterial systems that use gene conversion to generate antigenic variation are the *Borrelia* spp. (*B. hermsii, vlp* and *vsp*[37,38]; *Borrelia burgdorferi, vlsE*[39,40]) and *Neisseria* spp. (*pil*)[41–43]; a similar mechanism is used in the eukaryotic African trypanosomes (*vsg*)[44]. Gene conversion has more recently been recognized to have an important role in generating antigenic variation in the bacteria *Anaplasma marginale* (*msp*)[45,46], *T. pallidum* (*tprK*)[47], the fungal pathogen *P. carinii* (*msg*)[48,49] and the protozoan parasites *Babesia bovis* (*ves*)[50] and *P. falciparum* (*var*)[51,52]. Recombination through gene conversion is therefore one of the most widespread mechanisms of antigenic variation.

Mycoplasma spp. serve as excellent examples of organisms that use several alternative recombinatorial mechanisms for antigenic variation. In *Mycoplasma genitalium*, which is associated with urethritis, cervicitis and salpingitis, variation of the adhesion protein MgPa is generated by reciprocal recombination with defined donor loci[53] and possibly also through gene conversion[54]. Another recombinatorial type of antigenic variation involves the movement of promoter sequences and is best described in other species of *Mycoplasma*. Many phase-variable lipoproteins have been identified in *Mycoplasma* spp., including the VmpA and Avg proteins of *Mycoplasma agalactiae*[55,56], the Vsp proteins of *Mycoplasma bovis*[57,58], the LAMP (lipid-associated membrane protein) family of *Mycoplasma penetrans*[59] and the Vsa (variable surface antigen) of *Mycoplasma pulmonis*[60]. In all cases, phase variation is driven by inversion of a promoter sequence; however, numerous types of promoter inversion have been described. Two examples will be mentioned here (FIG. 4b). In *M. pulmonis*, the *vsa* locus contains one *vsa* expression site near several silent *vsa* genes, each of which lacks a functional promoter region and contains varying numbers of tandem repeat sequences and *vsa* recombination sites (*vrs* genes)[60]. Phase and antigenic variation occurs when the single expression unit (containing the promoter region plus the sequence coding for the conserved amino-terminal region of the Vsa) moves by recombination into the locus of a silent *vsa* gene; this occurs at a frequency of about 10^{-3} per colony-forming unit per generation, and the site-specific inversion is mediated by the HvsR recombinase[61]. The new Vsa has a shorter or longer number of tandem repeats, according to the length of the repeat region in the recombined silent gene. The number of repeats in these silent genes is thought to vary by slipped-strand mispairing. In *M. penetrans*, the *mpl* genes each encode different LAMP proteins and undergo phase variation at a rate estimated to be 10^{-2} to 10^{-4} per cell per generation[62]. In this organism, each *mpl* gene contains an independent promoter-like sequence, and the orientation of this region (promoter inversion) determines the on–off character of transcription. Notably, this contrasts with the *vsa* genes, in which a single promoter is moved from one pseudogene to another.

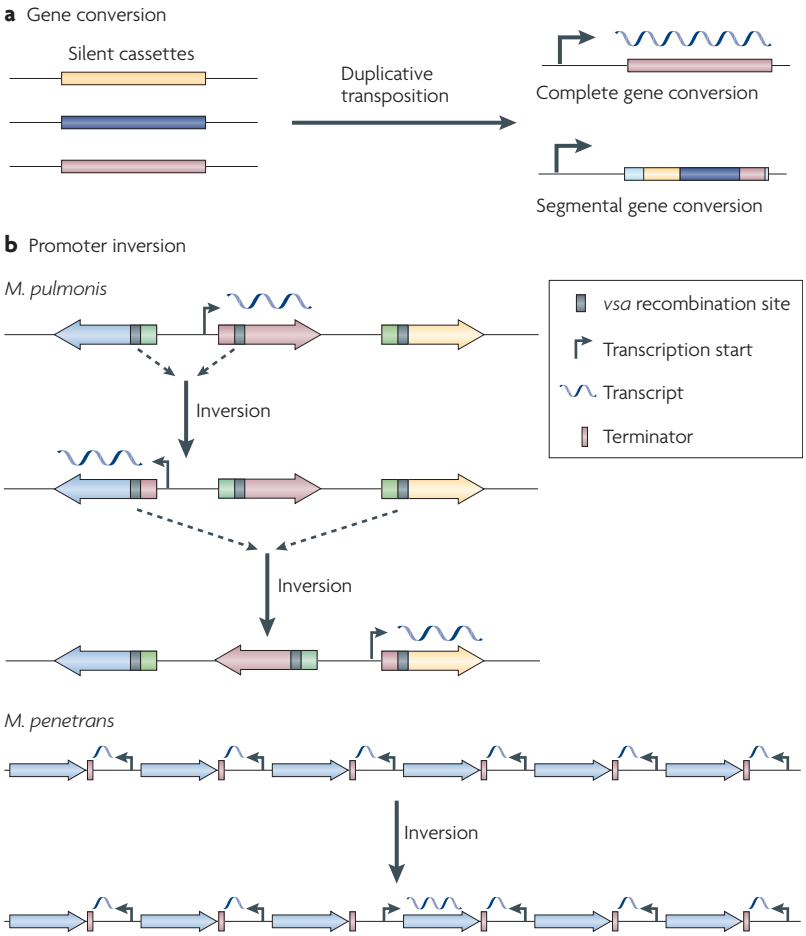

Figure 4 | **Antigenic variation through DNA recombination. a** | Gene conversion resulting in the duplication of a segment of a silent gene into an active expression site, as found in *Trypanosoma* spp., *Babesia* spp., *Pneumocystis carinii*, *Treponema pallidum* and *Borrelia* spp. The duplicated segment can include an entire open reading frame (complete gene conversion) or small regions of an open reading frame, creating a chimeric gene (segmental gene conversion). **b** | Promoter inversion. As shown for *Mycoplasma pulmonis*, inversion of a single promoter leads to expression of one *vsa* gene at a time (top). Inversion occurs by recombination between regions of sequence similarity (grey boxes). In the *mpl* genes of *Mycoplasma penetrans* (bottom), tandem arrays of genes are each regulated by separate promoters. In one orientation, a promoter drives the expression of a short transcript that is stopped by a transcriptional terminator and does not encode protein. When inverted, the promoter drives transcription through the *mpl* open reading frame, leading to expression of the encoded protein.

Antigenic variation through epigenetic modifications. Instead of using DNA sequence alterations, many systems instead rely on epigenetic modifications to control gene activation and silencing. Two examples of epigenetic transcriptional control were described above for *E. coli* and *C. glabrata*. There are many other examples, however, and they share common hallmarks, including histone modifications, the use of modified nucleotides, changes in chromatin structure and nuclear organization[63]. In bacterial systems DNA modification can contribute to phase variation[16,64]; by contrast, in eukaryotes such modifications often combine to ensure mutually exclusive expression. Substantial progress has been made in defining the epigenetic modifications that mark individual genes for either the active or silent states[65]; however, the mechanisms by which expression of an entire gene family is co-regulated, and by which a switch in the expressed gene is coordinated, largely remain a mystery.

Antigenic variation by African trypanosomes has long served as a paradigm for understanding this process in protozoan parasites. In addition to the recombinatorial mechanisms described above, trypanosomes can also alter *vsg* expression by changing which of several independent subtelomeric expression sites is actively transcribed, a process that is controlled epigenetically[66,67]. Chromatin condensation is observed at the silent expression sites, and chromatin remodelling enzymes, including a SWI2/SNF2 ATPase and the histone methyltransferase DOT1B, have recently been implicated in maintaining expression-site silencing[68,69]. In addition, a modified nucleotide referred to as DNA-J is preferentially incorporated throughout the silent expression sites[70,71], a phenomenon that is reminiscent of modifications such as DNA methylation at silent genes in bacteria and higher eukaryotes. These modifications might not regulate transcription initiation, however, as there is some evidence that expression is controlled at the level of transcription elongation, which is limited to a single expression site[72]. Interestingly, the active *vsg* expression site is transcribed by RNA polymerase I[73] and is localized within a specific extra-nucleolar region within the nucleus called the expression site body[74]. It has been suggested that this body can accommodate only a single *vsg* gene at a time, which would provide a potential mechanism for mutually exclusive expression.

Mutually exclusive expression of the *var* genes that encode PfEMP1 by *P. falciparum* seems to be regulated at the level of transcription initiation[75]. Switching between the active and silent states is associated with similar chromatin modifications to those seen in other eukaryotes; for example, histone acetylation is associated with active genes[76], whereas trimethylation of lysine 9 of histone H3 is found at silent loci[77,78]. Movement of a *var* gene to a specific subnuclear position has also been observed[79,80], indicating that nuclear structure might also affect the coordination of *var* gene expression in a similar manner to that described for trypanosomes. However, *var* genes are transcribed by RNA polymerase II[81], and experiments have shown that the expression site can accommodate more than one active *var* promoter at a time[82,83], leading to the conclusion that other mechanisms are involved in maintaining mutually exclusive expression. Other protozoan parasites, for example *Babesia* and *Giardia* spp., are likely to be similar to *Plasmodium* spp. in relying on RNA polymerase II for variant gene expression, and epigenetic modifications have been linked to *vsp* expression changes in *G. lamblia*[84]. In *C. glabrata*, the *epa* genes are silenced through the spreading of condensed telomeric chromatin into the subtelomeric chromosomal region in which the genes reside[85].

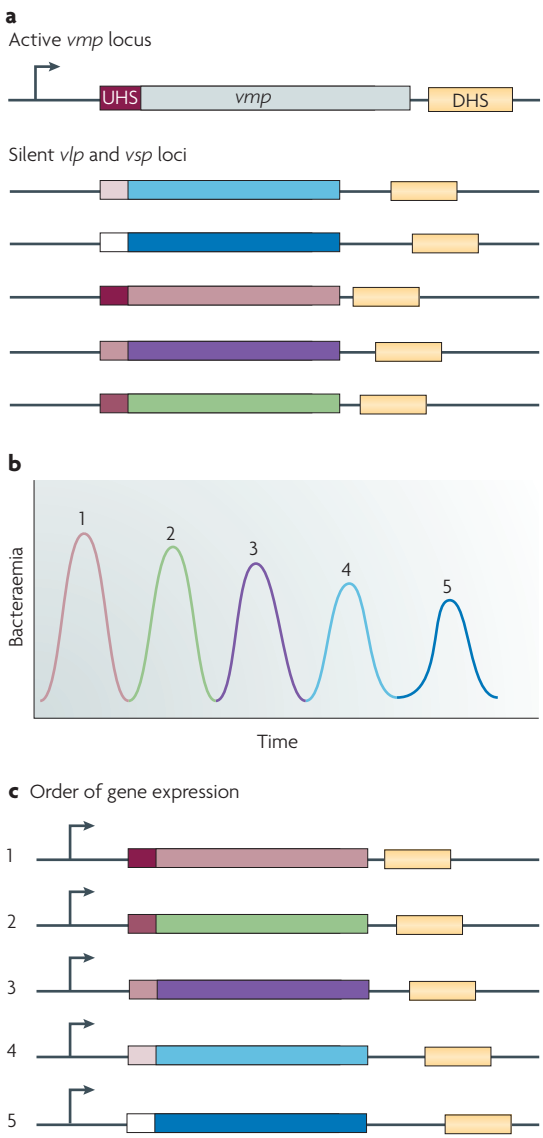

Figure 5 | **'Programmed' sequence changes in *Borrelia hermsii*.** a | The active *vmp* gene contains two elements that contribute to expression switching: the upstream and downstream homology sequences (UHSs and DHSs, respectively). These elements are also found in the silent *vlp* and *vsp* loci, with varying degrees of sequence identity (in the case of UHSs, this is represented by different colours). In addition, the distance between the DHS and the *vmp* coding region is also variable. **b** | In a mouse infection, waves of bacteria arise as antigenically distinct populations cause relapses. **c** | Examination of the expressed *vmp* gene in each wave shows that the order of expression is determined by the level of sequence identity within the UHS and the distance between the DHS and the coding region.

Programmed sequence change

Organisms that undergo antigenic variation by switching expression between members of large multicopy gene families raise the question of whether the individual genes are activated in a specific order or programmed sequence. It has been recognized for a number of years that, during a case of relapsing fever, there is a rough order of emergence of *B. hermsii* serotypes that arise by gene conversion during the subsequent relapses following infection with a given serotype[86]; however, the mechanism of this 'programming' has not been understood until recently. Dai and colleagues[87] identified a ≤62-nucleotide region located upstream of the *vmp* expression site that has varying homology with similar regions upstream of the silent *vlp* and *vsp* genes; these are termed upstream homology sequences (UHSs). They also identified 13 repetitive 214-bp sequences downstream of each silent gene and another downstream of the expression locus, termed downstream homology sequences (DHSs). Through careful analysis of variants obtained during mouse infections, Dai *et al.* identified the precise sites of recombination within these regions surrounding the new *vmp*. This and a subsequent study from the same group[88] demonstrated that the likelihood of recombination for a particular silent gene increased with higher homology between the expression site and donor site UHSs. Additionally, the shorter the distance from the donor gene sequence to the DHS, the more likely the donor gene was to be recombined into the expression site. Using these measures, Barbour *et al.*[88] developed a recombination likelihood hierarchy for 22 *vlp* and *vsp* genes; comparison with the actual genes expressed during relapses following infection supported their model. This switching hierarchy model explains, for the first time, the semi-programmed antigenic variation that has been observed for decades in the relapsing fever caused by *Borrelia* species (FIG. 5).

In African trypanosomes, which also use gene conversion events as a primary mechanism for antigenic variation, the likelihood of specific genes being involved in the conversion event depends, at least in part, on their level of sequence identity, resulting in a rough order of gene expression over the length of an infection[89]. In organisms that do not rely on recombinatorial mechanisms for antigenic variation, for example, *P. falciparum*, no such specific order is observed[90]. However, it seems that not all genes possess equal inherent on or off rates[90,91], and certain genes are therefore activated more frequently than others. This might explain why different subsets of *var* genes are found to be active in patients who are not immune to malaria compared with patients who have suffered several previous infections and as a result are likely to possess antibodies against more frequently activated genes[92–94].

Evolutionary advantages for pathogens

For pathogens, reproductive fitness and evolutionary success depend on their capacity to infect a host and to survive long enough to be transmitted to the next host. The most obvious advantage of antigenic variation is that it allows pathogens to avoid immune recognition and extend the length of an infection (BOX 1). However, other less well-recognized benefits are also important; for example, conservation of metabolic energy by

Box 1 | **Evolutionary success of an infectious agent and R_0**

The evolutionary success of an infectious agent depends on several factors, only some of which are determined by the pathogen. This concept has been elegantly expressed by Anderson and May[118], who quantified the success of an infectious agent according to the equation $R_0 = \beta cD$, in which R_0 is the average number of new infections caused by a single infected person. If R_0 =>1, the infection spreads and increases in incidence, and the pathogen is successful. In this equation, β is the inherent transmissibility of the agent, *c* is related to the number of transmission opportunities and *D* is the duration of infectiousness of the disease. Thus, any mechanism by which a pathogen can increase one or more of these factors will provide a selective advantage. The most obvious and best-recognized benefit of phase and antigenic variation is immune evasion — that is, the capacity to disarm the effect of acquired immunity by changing epitopes. Successful evasion of host immunity results in an increased value for *D*.

regulating expression of gene products that are no longer needed (for example, pili), and the generation of diverse surface structures. Such alterations in surface structures can facilitate adaptation to new environments by conferring a fitness advantage in certain environments, as with the binding of placental receptors in malaria[95] or the capacity of *E. coli*[19] or *C. glabrata*[20] to adhere to the genito-urinary epithelium.

Immune evasion. Although antigenic variation can help pathogens to avoid the innate immune system (for example, variant lipopolysaccharide can result in decreased sensitivity to cationic peptides), the main benefit of phase and antigenic variation is evasion of the developing acquired immune response of the infected host, specifically by interfering directly with antibody function. For some organisms, such as *B. burgdorferi* or *B. hermsii*, a specific antibody mediates the clearance of the organisms from the bloodstream[96–100]. The capacity to turn off or to alter the epitopes of surface antigens renders these antibodies ineffective, thereby facilitating survival. For other pathogens, such as *T. pallidum*, antibody (or antibody plus complement) bound to the pathogen surface can opsonize the organism so that it is more readily ingested and killed by macrophages, neutrophils or other phagocytic cells[101]. In the TprK antigen of *T. pallidum*, the variant portions (V regions) of the protein are antibody epitopes[102], and even minor sequence changes in the V regions can abrogate the capacity of antibodies to bind to these peptides[103]. Preventing opsonic antibody binding to TprK reduces phagocytosis of the organisms, which again facilitates survival. Malaria parasites live inside the red blood cells of their hosts, and antibodies to the variant surface antigens (for example, PfEMP1) alone do not kill parasites[104,105]. However, in combination with macrophages, antibodies against parasite-encoded surface antigens can lead to antibody-dependent cellular inhibition (ADCI), which kills parasites *in vitro*[106,107]. In addition to ADCI, antibodies against PfEMP1 can disrupt cytoadhesion[105], which is required to avoid circulation through the spleen, where infected cells are cleared. Antigenic variation therefore enables malaria parasites to avoid destruction by macrophages and splenic clearance, leading to longer persistence of the pathogen in the host. Similar mechanisms are seen in *Babesia* species[108].

Enhanced duration of the infectious stage. Persistence of a pathogen in its host is advantageous only if this persistence contributes to transmission. Many elegant antigenic variation systems are found in blood-borne pathogens, such as *Plasmodium* spp., *Anaplasma* spp., *Babesia* spp., *Trypanosoma* spp. and *Borrelia* spp., and antigenic variation in each of these examples contributes to survival in the host and more lengthy infections. Long-term survival in the bloodstream, however, does not directly contribute to the evolutionary success of the pathogen unless it relates to transmission. In the cases listed above, transmission occurs through the taking of a blood meal by an arthropod vector. The vector then transmits the infection to a new host through a bite. Increasing the persistence of the pathogen in the bloodstream through antigenic variation increases the likelihood of transmission via the vectors and contributes to the success of the pathogen.

Antigenic variation, however, also exists in organisms that are not transmitted by blood-feeding vectors. For example, syphilis is a persistent infection that disseminates through the bloodstream to distant tissues. As described above, TprK undergoes rapid sequence variation in *T. pallidum*, leading to chronic infection with bacteria that can persist in numerous tissues for decades. Yet syphilis is transmitted only during the primary and secondary stages, when skin lesions are present. What is the advantage to *T. pallidum* of TprK variation and decades-long infection? Although there are no vectors that transmit the infection following blood meals, the natural history of syphilis and knowledge of the mechanism of treponemal clearance from early lesions provides clues as to the possible advantages of antigenic variation in this organism. It is recognized that the infectious primary lesion (chancre) persists for 2–6 weeks. Treponemes are cleared from these lesions following opsonization by specific antibodies and phagocytosis by interferon-γ-activated macrophages[109–113]; as a result, antigenic variation can result in a lack of effective opsonization by TprK-specific antibodies, leading to increased duration of the infectious primary lesion. Furthermore, the rash of secondary syphilis (which is also infectious) generally develops weeks after resolution of the primary chancre and persists for weeks to months. It is thought that these lesions appear at skin sites that were seeded by haematogenous dissemination during the primary stage. Antigenic variation might therefore promote survival of treponemes in the bloodstream during dissemination, leading to development of the infectious secondary stage, as well as prolonged duration of the secondary rash. Although decades-long persistence of *T. pallidum* infection can result from antigenic variation, we propose that the real advantage of TprK variation for *T. pallidum* is the prolongation of the infectious primary and secondary stages. Other organisms, such as *Mycoplasma* spp. and *Neisseria* spp., which also have elegant systems for generating antigenic variation, are also not transmitted from the bloodstream. In these infections, persistence on mucosal surfaces is required for transmission, and antigenic variation can protect the organisms from immune clearance by mucosal antibody and opsonophagocytosis.

Reinfection and superinfection. The capacity of a given pathogen to infect a host that has resolved, or been cured of, prior infection (reinfection) or is persistently infected with the same organism (superinfection) provides two advantages to the pathogen. The first is that it preserves a large population of susceptible hosts. If initial infection results in development of species-wide protective immunity, the proportion of hosts in a population that are susceptible to infection by another strain of the same species is diminished. By contrast, when initial infection does not provide broad protection, the number of susceptible hosts remains larger. The second benefit is that it permits genetic exchange among strains. This occurs by transformation and conjugation in bacteria and by sexual recombination in eukaryotic pathogens. Antigenic variation is important for both reinfection and superinfection: during an initial infection, specific immunity develops against the subset of variants expressed by the first population of infecting organisms; however, if the host is challenged by a variant that it has not previously encountered, reinfection or superinfection can occur. Futse and colleagues[114] have examined the requirements for superinfection in *A. marginale*. During initial infection, gene conversion of the hypervariable region results in expression of a large number of Msp2 variants. Sequencing the complete repertoire of the *msp2* donor alleles revealed that the existence of as few as one unique donor allele was sufficient to permit reinfection or superinfection by that strain. Similarly, in regions of high malaria transmission, the existence of superinfection with *P. falciparum* is extremely high: a recent study showed that single individuals were simultaneously infected with up to seven genetically distinct parasite populations[115]. In addition, reinfection is virtually always possible, with most individuals never developing sterile immunity, even after repeated infections[116].

These observations led to the hypothesis that the duration of infection is likely to be determined by the number of donor alleles in the variant repertoire and the mechanism of molecular variation. For example, segmental gene conversion can theoretically generate more antigenic variants than promoter inversion as a consequence of the vastly larger number of potential chimeric sequences that can be expressed. By contrast, reinfection or superinfection may be less dependent on the number of donor alleles but more dependent on the diversity of donor alleles among strains. Studies have produced results that are consistent with this model. A recent attempt to measure the extent of diversity in the *var* gene family in *P. falciparum* (in which reinfection and superinfection are extremely common) from global parasite isolates was unable to detect significant overlap between individual parasite lines, despite having sequenced more than 8,000 variant sequences[117]. This tremendous diversity in the repertoire of surface antigens perhaps explains why sterile immunity to *P. falciparum* infection is difficult or impossible to achieve, regardless of the number of infections.

Conclusions

The level to which pathogenic organisms of disparate evolutionary origins have independently converged on similar strategies for avoiding destruction by the immune system is remarkable. By studying the mechanisms that bacterial, fungal and protozoan parasites have developed to maintain persistent infections, researchers have also gained a much deeper understanding of the intricacies of the mammalian immune system. Immunology, parasitology, mycology and bacteriology all converge at the focal point of host–pathogen interactions and provide fertile ground for breakthroughs in all of these fields of study. Such breakthroughs will, we hope, result in new intervention strategies for alleviating human diseases caused by these infectious agents.

1. Berg, R. D. The indigenous gastrointestinal microflora. *Trends Microbiol.* **4**, 430–435 (1996).
2. Zeig, J., Silverman, M., Hilmen, H. & Simon, M. Recombinational switching for gene expression. *Science* **196**, 170–175 (1977).
3. Merz, A. J. & So, M. Interactions of pathogenic *Neisseriae* with epithelial cell membranes. *Annu. Rev. Cell Dev. Biol.* **16**, 423–457 (2000).
4. Plant, L. & Jonsson, A. B. Contacting the host: insights and implications of pathogenic *Neisseria* cell interactions. *Scand. J. Infect. Dis.* **35**, 608–613 (2003).
5. Gray-Owen, S. D. Neisserial Opa proteins: impact on colonization, dissemination and immunity. *Scand. J. Infect. Dis.* **35**, 614–618 (2003).
6. Edwards, J. L. & Apicella, M. A. The molecular mechanisms used by *Neisseria gonorrhoeae* to initiate infection differ between men and women. *Clin. Microbiol. Rev.* **17**, 965–981 (2004).
7. Leader, B. T. *et al.* Antibody responses elicited against the *Treponema pallidum* repeat proteins differ during infection with different isolates of *Treponema pallidum* subsp. *pallidum*. *Infect. Immun.* **71**, 6054–6057 (2003).
8. Brinkman, M. B. *et al.* Reactivity of antibodies from syphilis patients to a protein array representing the *Treponema pallidum* proteome. *J. Clin. Microbiol.* **44**, 888–891 (2006).
9. McKevitt, M. *et al.* Genome scale identification of *Treponema pallidum* antigens. *Infect. Immun.* **73**, 4445–4450 (2005).
10. Giacani, L. *et al.* Quantitative analysis of *tpr* gene expression in *Treponema pallidum* isolates: differences among isolates and correlation with T-cell responsiveness in experimental syphilis. *Infect. Immun.* **75**, 104–112 (2007).
11. Fraser, C. M. *et al.* Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**, 375–388 (1998).
12. Giacani, L., Lukehart, S. & Centurion-Lara, A. Length of guanosine homopolymeric repeats modulates promoter activity of subfamily II *tpr* genes of *Treponema pallidum* ssp. *pallidum*. *FEMS Immunol. Med. Microbiol.* **51**, 289–301 (2007).
13. Jonsson, A., Nyberg, G. & Normark, S. Phase variation of gonococcal pili by frameshift mutation in *pilC*, a novel gene for pilus assembly. *EMBO J.* **10**, 477–488 (1991).
14. Koomey, J. M., Gotschlich, E. C., Robbins, K., Berstrom, S. & Swanson, J. Effects of *recA* mutations on pilus antigenic variation and phase transitions in *Neisseria gonorrhoeae*. *Genetics* **117**, 391–398 (1987).
15. van der Ende, A. *et al.* Variable expression of class 1 outer membrane protein in *Neisseria meningitidis* is caused by variation in the spacing between the −10 and −35 regions of the promoter. *J. Bacteriol.* **177**, 2475–2480 (1995).
16. Blyn, L. B., Braaten, B. A. & Low, D. A. Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J.* **9**, 4045–4054 (1990).
17. Hernday, A., Krabbe, M., Braaten, B. & Low, D. Self-perpetuating epigenetic pili switches in bacteria. *Proc. Natl Acad. Sci. USA* **99** (Suppl. 4), 16470–16476 (2002).
18. Peterson, S. N. & Reich, N. O. Competitive Lrp and Dam assembly at the pap regulatory region: implications for mechanisms of epigenetic regulation. *J. Mol. Biol.* **383**, 92–105 (2008).
19. Holden, N., Totsika, M., Dixon, L., Catherwood, K. & Gally, D. L. Regulation of P-fimbrial phase variation frequencies in *Escherichia coli* CFT073. *Infect. Immun.* **75**, 3325–3334 (2007).
20. Domergue, R. *et al.* Nicotinic acid limitation regulates silencing of *Candida* adhesins during UTI. *Science* **308**, 866–870 (2005).
Demonstrates how pathogenic fungi can react to environmental conditions to express appropriate surface proteins.
21. Stern, A., Brown, M., Nickel, P. & Meyer, T. F. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell* **47**, 61–67 (1986).
22. Stern, A., Nickel, P., Meyer, T. F. & So, M. Opacity determinants of *Neisseria gonorrhoeae*: gene expression and chromosomal linkage to the gonococcal pilus gene. *Cell* **37**, 447–456 (1984).
23. Keely, S. P. *et al.* Gene arrays at *Pneumocystis carinii* telomeres. *Genetics* **170**, 1589–1600 (2005).

Describes the complete sequences of five clusters of genes encoding surface antigens that have the potential to generate high-frequency antigenic variation, which is likely to be a strategy by which this parasitic fungus prolongs its survival in the rat lung.

24. Berendt, A. R., Ferguson, D. J. P. & Newbold, C. I. Sequestration in *Plasmodium falciparum* malaria: sticky cells and sticky problems. *Parasitol. Today* **6**, 247–254 (1990).
25. Salanti, A. *et al.* Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *J. Exp. Med.* **200**, 1197–1203 (2004).
26. Amulic, B., Salanti, A., Lavstsen, T., Nielsen, M. A. & Deitsch, K. W. An upstream open reading frame controls translation of *var2csa*, a gene implicated in placental malaria. *PLoS Pathog.* **5**, e1000256 (2009).
27. Morris, D. R. & Geballe, A. P. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell Biol.* **20**, 8635–8642 (2000).
28. Prucca, C. G. *et al.* Antigenic variation in *Giardia lamblia* is regulated by RNA interference. *Nature* **456**, 750–754 (2008).
This paper was the first to describe RNA interference as a mechanism for regulating mutually exclusive expression of a large multicopy gene family responsible for antigenic variation.
29. Gray, R. R. *et al.* Molecular evolution of the *tprC*, *D*, *I*, *K*, *G*, and *J* genes in the pathogenic genus *Treponema*. *Mol. Biol. Evol.* **23**, 2220–2233 (2006).
30. Taylor, J. E. & Rudenko, G. Switching trypanosome coats: what's in the wardrobe? *Trends Genet.* **22**, 614–620 (2006).
31. Baruch, D. I. *et al.* Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82**, 77–87 (1995).
32. Smith, J. D. *et al.* Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110 (1995).
33. Su, X. *et al.* A large and diverse gene family (*var*) encodes 200–350 kD proteins implicated in the antigenic variation and cytoadherence of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89–100 (1995).
References 31–33 describe the discovery of the *var* gene family, which encodes the primary surface antigen and virulence factor of the human malaria parasite *P. falciparum*.
34. Nash, T. E. Surface antigenic variation in *Giardia lamblia*. *Mol. Microbiol.* **45**, 585–590 (2002).
35. Allred, D. R. *et al.* The *ves* multigene family of *B. bovis* encodes components of rapid antigenic variation at the infected erythrocyte surface. *Mol. Cell* **5**, 153–162 (2000).
36. Filler, S. G. *Candida*–host cell receptor–ligand interactions. *Curr. Opin. Microbiol.* **9**, 333–339 (2006).
37. Kitten, T. & Barbour, A. G. Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. *Proc. Natl Acad. Sci. USA* **87**, 6077–6081 (1990).
38. Barbour, A. G., Burman, N., Carter, C. J., Kitten, T. & Bergstrom, S. Variable antigen genes of the relapsing fever agent *Borrelia hermsii* are activated by promoter addition. *Mol. Microbiol.* **5**, 489–493 (1991).
39. Zhang, J. R. & Norris, S. J. Genetic variation of the *Borrelia burgdorferi* gene *vlsE* involves cassette-specific, segmental gene conversion. *Infect. Immun.* **66**, 3698–3704 (1998).
40. Zhang, J. R. & Norris, S. J. Kinetics and *in vivo* induction of genetic variation of *vlsE* in *Borrelia burgdorferi*. *Infect. Immun.* **66**, 3689–3697 (1998).
41. Criss, A. K., Kline, K. A. & Seifert, H. S. The frequency and rate of pilin antigenic variation in *Neisseria gonorrhoeae*. *Mol. Microbiol.* **58**, 510–519 (2005).
42. Serkin, C. D. & Seifert, H. S. Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. *J. Bacteriol.* **180**, 1955–1958 (1998).
43. Haas, R. & Meyer, T. F. The repertoire of silent pilus genes in *Neisseria gonorrhoeae*: evidence for gene conversion. *Cell* **44**, 107–115 (1986).
44. Bernards, A. *et al.* Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. *Cell* **27**, 497–505 (1981).
Describes the correlation between transcription of an *msg* gene and its residence at the unique locus in the genome that promotes transcription of adjacent *msg* genes.
45. Barbet, A. F., Lundgren, A., Yi, J., Rurangirwa, F. R. & Palmer, G. H. Antigenic variation of *Anaplasma marginale* by expression of MSP2 mosaics. *Infect. Immun.* **68**, 6133–6138 (2000).
46. Brayton, K. A., Palmer, G. H., Lundgren, A., Yi, J. & Barbet, A. F. Antigenic variation of *Anaplasma marginale msp2* occurs by combinatorial gene conversion. *Mol. Microbiol.* **43**, 1151–1159 (2002).
47. Centurion-Lara, A. *et al.* Gene conversion: a mechanism for generation of heterogeneity in the *tprK* gene of *Treponema pallidum* during infection. *Mol. Microbiol.* **52**, 1579–1596 (2004).
48. Sunkin, S. M. & Stringer, J. R. Residence at the expression site is necessary and sufficient for the transcription of surface antigen genes of *Pneumocystis carinii*. *Mol. Microbiol.* **25**, 147–160 (1997).
49. Stringer, J. R. Antigenic variation in *Pneumocystis*. *J. Eukaryot. Microbiol.* **54**, 8–13 (2007).
50. al Khedery, B. & Allred, D. R. Antigenic variation in *Babesia bovis* occurs through segmental gene conversion of the *ves* multigene family, within a bidirectional locus of active transcription. *Mol. Microbiol.* **59**, 402–414 (2006).
Describes the identification of an expression site for the primary surface antigen of *B. bovis*, including the possibility that the genes encoding the two subunits are expressed from a single, bidirectional promoter.
51. Kraemer, S. M. *et al.* Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. *BMC Genomics* **8**, 45 (2007).
52. Frank, M. *et al.* Frequent recombination events generate diversity within the multi-copy variant antigen gene families of *Plasmodium falciparum*. *Int. J. Parasitol.* **38**, 1099–1109 (2008).
53. Iverson-Cabral, S. L., Astete, S. G., Cohen, C. R. & Totten, P. A. *mgpB* and *mgpC* sequence diversity in *Mycoplasma genitalium* is generated by segmental reciprocal recombination with repetitive chromosomal sequences. *Mol. Microbiol.* **66**, 55–73 (2007).
The first paper to describe segmental reciprocal recombination as a mechanism of antigenic variation.
54. Ma, L. *et al. Mycoplasma genitalium*: an efficient strategy to generate genetic variation from a minimal genome. *Mol. Microbiol.* **66**, 220–236 (2007).
55. 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).
56. Glew, M. D. *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).
57. Lysnyansky, I., Rosengarten, R. & Yogev, D. Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. *J. Bacteriol.* **178**, 5395–5401 (1996).
58. 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).
59. Neyrolles, O. *et al.* Antigenic characterization and cytolocalization of P35, the major *Mycoplasma penetrans* antigen. *Microbiology* **145**, 343–355 (1999).
60. Bhugra, B., Voelker, L. L., Zou, N., Yu, H. & Dybvig, K. Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. *Mol. Microbiol.* **18**, 703–714 (1995).
61. Sitaraman, R., Denison, A. M. & Dybvig, K. A unique, bifunctional site-specific DNA recombinase from *Mycoplasma pulmonis*. *Mol. Microbiol.* **46**, 1033–1040 (2002).
62. Roske, K. *et al.* Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect. Immun.* **69**, 7642–7651 (2001).
63. Goldmit, M. & Bergman, Y. Monoallelic gene expression: a repertoire of recurrent themes. *Immunol. Rev.* **200**, 197–214 (2004).
64. Casadesus, J. & Low, D. Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* **70**, 830–856 (2006).
65. Hakimi, M. A. & Deitsch, K. W. Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. *Curr. Opin. Microbiol.* **10**, 357–362 (2007).
66. Cross, G. A. M. Antigenic variation in trypanosomes: secrets surface slowly. *Bioessays* **18**, 283–291 (1996).
67. Greaves, D. R. & Borst, P. *Trypanosoma brucei* variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion. *J. Mol. Biol.* **197**, 471–483 (1987).
68. Hughes, K. *et al.* A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. *EMBO J.* **26**, 2400–2410 (2007).
69. Figueiredo, L. M., Janzen, C. J. & Cross, G. A. A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biol.* **6**, e161 (2008).
70. Gommers-Ampt, J. H. *et al.* β-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*. *Cell* **75**, 1129–1136 (1993).
71. Van Leeuwen, F. *et al.* Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. *Genes Dev.* **11**, 3232–3241 (1997).
72. Vanhamme, L. *et al.* Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol. Microbiol.* **36**, 328–340 (2000).
73. Gunzl, A. *et al.* RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot. Cell* **2**, 542–551 (2003).
74. Navarro, M. & Gull, K. A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* **414**, 759–763 (2001).
Reports the first evidence for a specific subnuclear expression site where active transcription of the gene encoding the primary cell surface antigen takes place; provides the basis for a model for mutually exclusive expression that relies on only a single antigen-encoding gene having access to the expression site at a time.
75. Scherf, A. *et al.* Antigenic variation in malaria: *in situ* switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J.* **17**, 5418–5426 (1998).
76. Freitas-Junior, L. H. *et al.* Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**, 25–36 (2005).
77. Chookajorn, T. *et al.* Epigenetic memory at malaria virulence genes. *Proc. Natl Acad. Sci. USA* **104**, 899–902 (2007).
78. Lopez-Rubio, J. J. *et al.* 5′ flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol.* **66**, 1296–1305 (2007).
79. Duraisingh, M. T. *et al.* Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* **121**, 13–24 (2005).
References 76 and 79 provide the first descriptions of chromatin modifications associated with mutually exclusive *var* gene expression in malaria parasites.
80. Ralph, S. A., Scheidig-Benatar, C. & Scherf, A. Antigenic variation in *Plasmodium falciparum* is associated with movement of *var* loci between subnuclear locations. *Proc. Natl Acad. Sci. USA* **102**, 5414–5419 (2005).
81. Kyes, S. *et al. Plasmodium falciparum var* gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. *Mol. Microbiol.* **63**, 1237–1247 (2007).
82. Dzikowski, R. *et al.* Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep.* **8**, 959–965 (2007).
83. Dzikowski, R. & Deitsch, K. W. Active transcription is required for maintenance of epigenetic memory in the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* **382**, 288–297 (2008).
84. Kulakova, L., Singer, S. M., Conrad, J. & Nash, T. E. Epigenetic mechanisms are involved in the control of *Giardia lamblia* antigenic variation. *Mol. Microbiol.* **61**, 1533–1542 (2006).
85. De Las Peñas, A. *et al.* Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to *RAP1*- and *SIR*-dependent transcriptional silencing. *Genes Dev.* **17**, 2245–2258 (2003).
86. Stoenner, H. G., Dodd, T. & Larsen, C. Antigenic variation of *Borrelia hermsii*. *J. Exp. Med.* **156**, 1297–1311 (1982).
87. Dai, Q. *et al.* Antigenic variation by *Borrelia hermsii* occurs through recombination between extragenic repetitive elements on linear plasmids. *Mol. Microbiol.* **60**, 1329–1343 (2006).

88. Barbour, A. G., Dai, Q., Restrepo, B. I., Stoenner, H. G. & Frank, S. A. Pathogen escape from host immunity by a genome program for antigenic variation. *Proc. Natl Acad. Sci. USA* **103**, 18290–18295 (2006).
89. Marcello, L. & Barry, J. D. From silent genes to noisy populations-dialogue between the genotype and phenotypes of antigenic variation. *J. Eukaryot. Microbiol.* **54**, 14–17 (2007).
90. Frank, M., Dzikowski, R., Amulic, B. & Deitsch, K. Variable switching rates of malaria virulence genes are associated with chromosomal position. *Mol. Microbiol.* **64**, 1486–1498 (2007).
91. Horrocks, P., Pinches, R., Christodoulou, Z., Kyes, S. A. & Newbold, C. I. Variable *var* transition rates underlie antigenic variation in malaria. *Proc. Natl Acad. Sci. USA* **101**, 11129–11134 (2004).
92. Jensen, A. T. *et al. Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. *J. Exp. Med.* **199**, 1179–1190 (2004).
93. Rottmann, M. *et al.* Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infect. Immun.* **74**, 3904–3911 (2006).
94. Kaestli, M. *et al.* Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. *J. Infect. Dis.* **193**, 1567–1574 (2006).
95. Salanti, A. *et al.* Selective upregulation of a single distinctly structured *var* gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol. Microbiol.* **49**, 179–191 (2003).
96. Cadavid, D., Thomas, D. D., Crawley, R. & Barbour, A. G. Variability of a bacterial surface protein and disease expression in a possible mouse model of systemic Lyme borreliosis. *J. Exp. Med.* **179**, 631–642 (1994).
97. Alugupalli, K. R. *et al.* The resolution of relapsing fever borreliosis requires IgM and is concurrent with expansion of B1b lymphocytes. *J. Immunol.* **170**, 3819–3827 (2003).
98. Fikrig, E. *et al.* Sera from patients with chronic Lyme disease protect mice from Lyme borreliosis. *J. Infect. Dis.* **169**, 568–574 (1994).
99. Barthold, S. W. & Bockenstedt, L. K. Passive immunizing activity of sera from mice infected with *Borrelia burgdorferi*. *Infect. Immun.* **61**, 4696–4702 (1993).
100. Johnson, R. C., Kodner, C. & Russell, M. Passive immunization of hamsters against experimental infection with the Lyme disease spirochete. *Infect. Immun.* **53**, 713–714 (1986).
101. Lukehart, S. A. & Miller, J. N. Demonstration of the *in vitro* phagocytosis of *Treponema pallidum* by rabbit peritoneal macrophages. *J. Immunol.* **121**, 2014–2024 (1978).
102. Morgan, C. A., Molini, B. J., Lukehart, S. A. & Van Voorhis, W. C. Segregation of B and T cell epitopes of *Treponema pallidum* repeat protein K to variable and conserved regions during experimental syphilis infection. *J. Immunol.* **169**, 952–957 (2002).
103. LaFond, R. E., Molini, B. J., Van Voorhis, W. C. & Lukehart, S. A. Antigenic variation of TprK V regions abrogates specific antibody binding in syphilis. *Infect. Immun.* **74**, 6244–6251 (2006).
104. McGregor, I. A. Studies in the acquisition of immunity of *Plasmodium falciparum* infections in Africa. *Trans. R. Soc. Trop. Med. Hyg.* **58**, 80–92 (1964).
105. Baruch, D. I., Gormley, J. A., Ma, C., Howard, R. J. & Pasloske, B. L. *Plasmodium falciparum* erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc. Natl Acad. Sci. USA* **93**, 3497–3502 (1996).
106. Bouharoun-Tayoun, H., Oeuvray, C., Lunel, F. & Druilhe, P. Mechanisms underlying the monocyte-mediated antibody-dependent killing of *Plasmodium falciparum* asexual blood stages. *J. Exp. Med.* **182**, 409–418 (1995).
107. Badell, E. *et al.* Human malaria in immunocompromised mice: an *in vivo* model to study defense mechanisms against *Plasmodium falciparum*. *J. Exp. Med.* **192**, 1653–1660 (2000).
108. O'Connor, R. M. & Allred, D. R. Selection of *Babesia bovis*-infected erythrocytes for adhesion to endothelial cells coselects for altered variant erythrocyte surface antigen isoforms. *J. Immunol.* **164**, 2037–2045 (2000).
109. Baker-Zander, S. A., Shaffer, J. M. & Lukehart, S. A. Characterization of the serum requirement for macrophage-mediated killing of *Treponema pallidum* ssp. *pallidum*: relationship to the development of opsonizing antibodies. *FEMS Immunol. Med. Microbiol.* **6**, 273–279 (1993).
110. Lukehart, S. A., Shaffer, J. M. & Baker-Zander, S. A. A subpopulation of *Treponema pallidum* is resistant to phagocytosis: possible mechanism of persistence. *J. Infect. Dis.* **166**, 1449–1453 (1992).
111. Sell, S., Gamboa, D., Baker-Zander, S. A., Lukehart, S. A. & Miller, J. N. Host response to *Treponema pallidum* in intradermally-infected rabbits: evidence for persistence of infection at local and distant sites. *J. Invest. Dermatol.* **75**, 470–475 (1980).
112. Lukehart, S. A., Baker-Zander, S. A., Lloyd, R. M. & Sell, S. Characterization of lymphocyte responsiveness in early experimental syphilis. II. Nature of cellular infiltration and *Treponema pallidum* distribution in testicular lesions. *J. Immunol.* **124**, 461–467 (1980).
113. Baker-Zander, S. A. & Lukehart, S. A. Macrophage-mediated killing of opsonized *Treponema pallidum*. *J. Infect. Dis.* **165**, 69–74 (1992).
114. Futse, J. E., Brayton, K. A., Dark, M. J., Knowles, D. P. Jr & Palmer, G. H. Superinfection as a driver of genomic diversification in antigenically variant pathogens. *Proc. Natl Acad. Sci. USA* **105**, 2123–2127 (2008).
115. Montgomery, J. *et al.* Differential *var* gene expression in the organs of patients dying of *falciparum* malaria. *Mol. Microbiol.* **65**, 959–967 (2007).
116. Miller, L. H., Good, M. F. & Milon, G. Malaria pathogenesis. *Science* **264**, 1878–1883 (1994).
117. Barry, A. E. *et al.* Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*. *PLoS Pathog.* **3**, e34 (2007).
118. Anderson, R. M. & May, R. M. Population biology of infectious diseases: part I. *Nature* **280**, 361–367 (1979).

Acknowledgements

The authors thank S. Frank for valuable discussions concerning the possible role of antigenic variation in the duration of infectious syphilis lesions. Work in the laboratory of K.W.D. is supported by a grant from the US National Institutes of Health (AI 52390) and the United States–Israel Binational Science Foundation. The Department of Microbiology and Immunology at Weill Medical College of Cornell University acknowledges the support of the William Randolph Hearst Foundation. K.W.D. is a Stavros S. Niarchos Scholar. Work in the laboratory of S.A.L. is supported by the National Institutes of Health (AI42143 and AI63940) and work in the laboratory of J.R.S. is supported by a grant from the National Institute of Allergy and Infectious Diseases (5R01AI036701-14).

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
msp2 | *tprC* | *tprD* | *tprF* | *tprI* | *tprE* | *tprG* | *tprJ*
Entrez Genome Project: http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj
Anaplasma marginale | *Babesia bovis* | *Borrelia hermsii* | *Borrelia burgdorferi* | *Candida albicans* | *Candida glabrata* | *Escherichia coli* | *Giardia lamblia* | *Mycoplasma agalactiae* | *Mycoplasma bovis* | *Mycoplasma genitalium* | *Mycoplasma penetrans* | *Mycoplasma pulmonis* | *Neisseria gonorrhoeae* | *Neisseria meningitidis* | *Plasmodium falciparum* | *Pneumocystis carinii* | *Treponema pallidum* | *Trypanosoma brucei*
UniProtKB: http://ca.expasy.org/sprot
MgPa | Sir2 | TprK

FURTHER INFORMATION

Kirk W. Deitsch's homepage: http://www.med.cornell.edu/research/kdeitsch/
Sheila A. Lukehart's homepage: http://depts.washington.edu/deptgh/faculty_staff/fac_bio.php?url=http://sph.washington.edu/faculty/fac_bio.asp?url_ID=Lukehart_Sheila
James R. Stringer's homepage: http://molgen.uc.edu/node/170

ALL LINKS ARE ACTIVE IN THE ONLINE PDF