Published in final edited form as:
Nat Rev Microbiol. 2009 July ; 7(7): 493–503. doi:10.1038/nrmicro2145.

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

Kirk W. Deitsch[1], **Sheila A. Lukehart**[2], and **James R. Stringer**[3]

[1]Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY 10021, USA

[2]Departments of Medicine and Global Health, University of Washington, Seattle, Washington 98195, USA

[3]Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, OH 45220, USA.

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. In the case of mammals and the pathogens that infect them, a veritable “arms race” has ensued. A highly adapted 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. Here, 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 thus maintain persistent infections, thereby ensuring their transmission to new hosts.

Introduction

The evolution of mammalian species has resulted in the development of relatively large, multi-cellular organisms that, in addition to replicating in their own right, also serve as 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 10-fold more bacterial cells than human cells[1]. Although the relationship between the host organism and the 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 or ultimately eradicate detrimental organisms. 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, be they viral, bacterial, fungal or protozoan, all face similar challenges upon 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

Correspondence to: Kirk W. Deitsch.
kwd2001@med.cornell.edu.

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 alternatively, once recognized, they must be able to avoid destruction by various components of innate and acquired immunity. This shared need to evade a common assault has resulted in the evolution of remarkably similar survival strategies, even among pathogens from distant evolutionary lineages. One of these strategies is antigenic variation; the ability of an infecting organism to systematically alter the proteins displayed to the host's immune system, thus confronting the host 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 several models regarding the mechanisms underlying this phenomenon. In more recent years, with the availability of extensive genome sequence data and improvements in tools available to study non-model pathogenic organisms, studies 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 highlight several recent examples in bacteria, protozoa and fungi that serve to illustrate common themes that are repeatedly observed despite large evolutionary distances separating the various pathogens.

Gene families and variant phenotypes

Antigenic variation in microbes is created via two general types of mechanisms, genetic and epigenetic. 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 or the amino acid sequence of its 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 relatively frequently and are readily reversible, features that distinguish these systems from antigenic variation caused by random spontaneous mutation as is more typical of some viruses like HIV.

A relatively 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 state differs from the other by exhibiting a particular cell-surface marker (e.g. pili) that is not present in the alternative phase. In some cases, more than one gene in a family can be regulated through a phase variation mechanism, in which case each gene switches on and off independently. Therefore, the possible phenotypes that can be created by X phase variable genes is 2^X (Figure 1).

While as few as 7 phase-variable genes can produce more than 100 different phenotypes, even more phenotypes can be generated by systems that utilize 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 given microbial cell. Therefore, the number of different phenotypes is at least as high as the number of family members. However, the number of possible phenotypes may be much greater than the number of family members if recombination occurs among family members to produce virtually unlimited diversity through the production of chimeric sequences.

Mechanisms of phase variation

Phase variation in bacteria was first identified in flagella expression of Salmonella, and a DNA inversion mechanism was implicated[2]. Perhaps the most extensively studied organisms that undergo phase variation, however, are the pathogenic Neisseria, 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[3,4,5,6], in particular the genes involved in the production of pili and Opa. Phase variation mechanisms that were first defined in these bacteria, including slipped-strand mispairing during DNA replication and gene conversion events, have provided a conceptual framework for many recent advances in understanding the mechanisms of antigenic variation in other prokaryotes and eukaryotes. Several recently described 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 due to the inability to culture or genetically manipulate the bacterium. However, recent work has shown many parallels to mechanisms originally described in Neisseria. 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 evidenced by development of humoral[7,8,9] and cellular[10] immune responses. The genes encoding subfamilies I (*tprC, D, F, I*) and II (*tprE, G, J*) of the *tpr* family have homopolymeric guanosine (poly-G) repeats of lengths varying from 7 to 12 bases 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 *G/F* and *J/I* operons[12]: loci with poly-G tracts of ≤8 Gs are transcribed, while transcription of those genes with longer poly-G tracts is reduced by ~95-100% (Figure 2A). Changes in the number of G repeats are thought to occur via slipped-strand mispairing during DNA replication, a mechanism that closely parallels that originally described for regulating the transcription of PilC (polyC)[13,14] and Por (polyG)[15] in Neisseria.

Phase variation can also result from epigenetic changes without alterations to the genome, and examples are found in both bacterial and eukaryotic pathogens. For example, the expression of pili on the surface of pathogenic *E. coli* allows adhesion within the urogenital tract and is dependent on the expression of the *pap* operon. Transcription of this operon is regulated by methylation of specific sites within 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 favored[19], facilitating 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 instance however, the cells are reacting to the low level of NAD+ found in urine, which influences the activity of the NAD +-dependent histone deacetylase Sir2p, 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 ability to sense the environment of the genitourinary 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 a variety of 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] (Figure 2B). The "Phase ON" state allows these bacteria to specifically adhere to certain tissues through attachment to specific host cell surface receptors. 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 very 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 very rich in A and T, yet long poly-G tracts occur in the middle of some *msr* genes, a large gene family whose members encode variable surface glycoproteins[23]. The number of G:C basepairs in these tracts appear to vary due to spontaneous slipped-strand mispairing events, causing frameshift mutations and truncation of the MSR protein and thus contributing to antigenic variation in *P. carinii* populations.

P. falciparum is a protozoan parasite that invades the circulating red blood cells of its human host, causing malaria. These parasites anchor a highly variable protein called PfEMP1 in the host red blood cell membrane. PfEMP1 binds tightly to host endothelial 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. *var* genes are expressed one at a time, thus transcriptional switching between genes leads to antigenic variation. However one particular *var* gene, called *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 (i.e. when infecting men, children or non-pregnant women), the transcript derived from this gene is also under translational regulation leading to a type of phase variation (either ON or OFF) similar to that observed in bacterial systems. In the case of *var2csa*, this is mediated by a small upstream open reading frame (uORF) that exists between the 5' end of the message and the start methionine of the PfEMP1 coding region[26]. Translation of the uORF results in dissociation of the ribosome from the mRNA prior to reaching the PfEMP1 start codon and thus silences expression of the protein (Figure 3A). It is not yet known how this repression is reversed when parasites infect pregnant women, but it may 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 diarrhea in susceptible individuals. The primary antigen exposed on its surface is called the 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 RNAi pathway, and thus are 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 not yet known (Figure 3B).

Large, hyper-variable gene families

While phase variation is a simple method for avoiding antibody recognition, 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 which are expressed singly, or alternatively multiple copies of partial gene sequences that can be recombined into a single expression site. Either way, sequence repertoires

provide the capacity to express large numbers of functionally similar but antigenically distinct surface proteins that can be expressed at different times over the course of the infection.

Large families of antigen-encoding sequences are observed in bacterial, protozoan and fungal pathogens (see table 1), indicating that expansion of the number of antigen encoding genes is a common evolutionary 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 *Neisseriae* spp.[3,4]. Similarly, the family of *tpr* genes of *Treponema* spp. has expanded to 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 ~1000 members in each parasite genome[30]. Smaller but similar hypervariable gene families are observed in the genomes of *Plasmodium* spp.[31,32,33], *Giardia* spp.[34] and *Babesia* spp.[35] In the fungal pathogen *Pneumocystis carinii*, genes from multi-copy 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 *C. 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 that utilize 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 mediate antigenic variation (Figure 4A). The recombination event can involve the movement of entire genes or, alternatively, small portions of genes resulting in chimeric sequences. This recombination can be duplicative (gene conversion) or reciprocal. The best studied bacterial systems that utilize gene conversion for antigenic variation are the *Borrelia* (*B. hermsii, vlp, vsp*[37,38]; *B. burgdorferi, vlsE*[39,40]) and *Neisseria* (*pil*)[41,42,43]; a similar mechanism is used in the eukaryotic African Trypanosomes (*vsg*)[44]. Gene conversion has more recently been recognized to play major role in antigenic variation in the bacteria *Anaplasma marginale* (*msp*)[45,46], *T. pallidum* (*tprK*)[47], the fungal pathogen *Pneumocystis carnii* (*msg*)[48,49] and the protozoan parasites *Babesia bovis* (*ves*)[50] and *Plasmodium falciparum* (*var*)[51,52]. Thus recombination through gene conversion is one of the most wide-spread mechanisms of antigenic variation used by an extremely broad and diverse number of pathogens.

The Mycoplasma species 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 its 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 species, including the VmpA and Avg proteins of *Mycoplasma agalactiae*[55,56], the Vsp proteins of *Mycoplasma bovis*[57,58], the LAMP (lipid-associated membrane proteins or P35) family of *Mycoplasma penetrans*[59], and the Vsa (variable surface antigens) of *Mycoplasma pulmonis*[60]. In all cases, phase variation occurs via inversion of a promoter sequence, however multiple types of promoter inversion have been described. Two examples will be mentioned here (Figure 4B). In *M. pulmonis*, the *vsa* locus contains one *vsa* expression site near multiple silent *vsa* genes, each lacking a functional promoter region and containing varying numbers of tandem repeat sequences and *vsa* recombination sites (*vrs*)[60]. Phase and antigenic variation occurs when the single expression unit (containing the

promoter region plus the sequence coding for the conserved N-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 CFU 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. Note the contrast with the *vsa* genes in which a single promoter is moved from one pseudogene to another.

Antigenic variation through epigenetic modifications

Instead of utilizing 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*. Many other examples have also been described, the hallmarks of which include histone modifications, the use of modified nucleotides, changes in chromatin structure, and nuclear organization[63]. While in bacterial systems DNA modification can contribute to phase variation[16,64], in eukaryotes such modifications often combine to ensure mutually exclusive expression. Significant 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 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 remodeling enzymes, including a Swi2/Snf2 ATPase and the histone methyl transferase 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], reminiscent of modifications like 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 perhaps this body can accommodate only a single *vsg* gene at a time, thus providing a potential mechanism for mutually exclusive expression.

Mutually exclusive expression of the *var* genes (encoding PfEMP1) by *P. falciparum* appears to be regulated at the level of transcription initiation[75]. Switching between the active and silent states is associated with chromatin modifications similar to those seen in other eukaryotes, for example histone acetylation is associated with active genes[76] while tri-methylation of lysine 9 of histone H3 is found at silent loci[77,78] (Figure 5). Movement of a *var* gene to a specific subnuclear position has also been observed[79,80], suggesting that nuclear structure may also play a role in coordinating *var* gene expression similar 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 instance *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].

Programmed sequence change

In organisms that undergo antigenic variation by switching expression between members of large multi-copy gene families, one question is whether the individual genes are activated in a specific order or programmed sequence. In relapsing fever, it was recognized a number of years ago that there is a rough order of *B. hermsii* serotypes that arise via gene conversion during the subsequent relapses following infection with a given serotype[86], and 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 (UHS). They also identified 13 repetitive 214-bp sequences downstream of each silent gene and another downstream of the expression locus, termed downstream homology sequences or DHS. Through careful analysis of variants obtained during mouse infections, they identified the precise sites of recombination within these regions surrounding the new *vmp*. In this study and in a subsequent study from the same group[88], it was demonstrated that the likelihood of recombination for a given silent gene increased with higher homology between the expression site and donor site UHS. Additionally, the shorter the distance from the donor gene sequence to the DHS, the more likely the donor gene is 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 in the relapsing fever Borrelia for decades (Figure 6).

In African trypanosomes, which also utilize 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 degree 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 recombinational mechanisms for antigenic variation, for example *P. falciparum*, no such specific order is observed[90]. However, all genes do not appear to possess equal inherent "on" or "off" rates[90,91], and thus certain genes are likely to get activated more frequently than others. This might explain why different subsets of *var* genes are found to be active in non-immune malaria patients when compared to patients who have suffered several previous infections and thus are likely to possess antibodies against more frequently activated genes[92,93,94].

Evolutionary advantages for pathogens

For pathogens, reproductive fitness and evolutionary success depend on the ability to infect a host and to survive long enough to be transmitted to the next host. The most obvious advantage of antigenic variation for pathogens is the ability to avoid immune recognition and thus extend the length of an infection (see text box 1). However, other less well-recognized benefits are also significant, for example conservation of metabolic energy by regulating expression of genes whose products are no longer needed (e.g. pili), and generation of diversity of surface structures. Such alterations in surface structures may

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 ability to adhere to genitourinary epithelium by *E. coli*[19] or *C. glabrata*[20].

Immune evasion

While antigenic variation can help to avoid the innate immune system (for example variant LPS can result in decreased sensitivity to cationic peptides), the major benefit of phase and antigenic variation is evasion of the developing immune response of the infected host, specifically by interfering directly with antibody function. For some organisms, such as *B. burgdorferi* or *B. hermsii*, specific antibody mediates the clearance of the organisms from the bloodstream[96,97,98,99,100]. The ability to turn off or to alter the epitopes of surface antigens renders these antibodies ineffective, thus 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 very minor sequence changes in the V regions can abrogate the ability of antibodies to bind to these peptides[103]. Inability of opsonic antibody to bind TprK reduces phagocytosis of the organisms, again facilitating survival. Malaria parasites live inside the RBCs of their hosts, and antibodies to the variant surface antigens (PfEMP1) alone do not kill parasites[104,105]. However, in combination with macrophages, antibodies against parasite-encoded surface antigens can lead to antibody mediated cellular inhibition (ADCI), which does kill 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. Thus antigenic variation 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*[108].

Enhanced duration of the infectious stage

Persistence of a pathogen within its host is advantageous only if this persistence contributes to transmission. Many of the elegant antigenic variation systems that have been described are found in blood-borne pathogens such as Plasmodium, Anaplasma, Babesia, Trypanosoma, and Borrelia, and antigenic variation in each of these examples contributes to survival in the host and thus 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, thus leading to pathogen success.

Antigenic variation exists however in organisms that are not transmitted by blood feeding vectors. For example, syphilis is a persistent infection that disseminates via the bloodstream to distant tissues. As described above, TprK undergoes rapid sequence variation in *T. pallidum*, leading to chronic infection with bacteria persisting 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? While there are no vectors that transmit the infection following bloodmeals, 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 IFNγ-activated macrophages[109,110,111,112,113]; thus antigenic variation may result in lack of effective opsonization via anti-TprK antibodies leading to increased duration of the infectious primary lesion. Further, the rash of secondary syphilis (which is also infectious) generally appears weeks following resolution of the primary chancre and persists for weeks-to-months. It is hypothesized that these lesions appear at skin sites that were seeded via hematogenous dissemination during the primary stage. Thus, antigenic variation might 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 may result from antigenic variation, we proposed the real advantage of TprK variation for *T. pallidum* is the prolongation of the infectious primary and secondary stages. Other organisms such as *Mycoplasma* sp. and *Neisseria* sp., which also have elegant antigenic variation systems, are also not transmitted from the bloodstream. In these infections, persistence on mucosal surfaces is required for transmission, and antigenic variation may protect the organisms from immune clearance via mucosal antibody and opsonophagocytosis.

Re-infection and superinfection

The ability of a given pathogen to infect a host that has resolved (or been cured of) prior infection (i.e., re-infection) or who is persistently infected with the same organism (superinfection) provides two advantages to the pathogen: 1) by preserving 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. In contrast, when initial infection does not provide broad protection, the number of susceptible hosts remains larger. 2) by permitting genetic exchange among strains. This occurs in bacteria by transformation and conjugation and in eukaryotic pathogens by sexual recombination. Antigenic variation is important for both re-infection 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 later encounters a never-seen variant, re-infection or superinfection can occur. Futse and colleagues[114] have examined the requirements for superinfection in *A. marginale*. During initial infection, a large number of Msp2 variants are expressed as a result of gene conversion of the hypervariable region of this antigen. 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 re-infection or superinfection by that strain. Similarly, in regions of high malaria transmission, the existence of superinfection with *P. falciparum* is extremely high, with a recent study showing that single individuals were simultaneously infected with up to seven genetically distinct parasite populations[115]. In addition, re-infection 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 determined by the number of the 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 because of the vastly larger number of potential chimeric sequences that can be expressed. In contrast, re-infection or superinfection may be less dependent upon the number of donor alleles, but more dependent upon the diversity of donor alleles among strains. Consistent with this model, for *P. falciparum* in which re-infection and superinfection are extremely common, a recent attempt to measure the extent of diversity in the *var* gene family from global parasite isolates was unable to detect significant overlap between individual parasite lines despite having sequenced over 8000 variant sequences[117]. This tremendous degree of diversity within the repertoire of surface antigens

perhaps explains why sterile immunity to *P. falciparum* infection is difficult or impossible to obtain, regardless of the number of infections.

Conclusions

The degree 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. Thus immunology, parasitology, mycology and bacteriology all converge at the focal point of host/pathogen interactions, providing a fertile ground for breakthroughs in all of these fields of study. Such breakthroughs will hopefully result in new intervention strategies for alleviating human diseases caused by these infectious agents.

Acknowledgments

The authors would like to thank Dr. Steven Frank for valuable discussions concerning the possible role of antigenic variation in the duration of infectious syphilis lesions. Work in the laboratory of KWD is supported by a grant from the 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. KWD is a Stavros S. Niarchos Scholar. Work in the laboratory of SAL is supported by the National Institutes of Health (AI42143 and AI63940) and work in the laboratory of JRS is supported by a grant from the National Institute of Allergy and Infectious Diseases (5R01AI036701-14).

Biography

Kirk Deitsch earned his PhD investigating the molecular biology of mosquito reproduction at Michigan State University. He began working with the parasites that cause severe malaria as a postdoctoral fellow at the National Institutes of Health under the direction of Thomas Wellems. He is currently an Associate Professor in the Department of Microbiology and Immunology at Weill Medical College of Cornell University where he studies the molecular basis for antigenic variation and epigenetic regulation of gene expression in the malaria parasite *Plasmodium falciparum*.

Sheila Lukehart earned a PhD in Microbiology & Immunology at University of California, Los Angeles, where she began her work on syphilis with James N. Miller. Following postdoctoral training in immunopathology with Stewart Sell at UC San Diego, and in sexually transmitted infections with King Holmes at University of Washington, she joined the faculty at University of Washington where she is currently Professor of Medicine (Division of Allergy and Infectious Diseases). Her research focuses on the pathogenic treponemes, with an emphasis on pathogenesis and immunity in syphilis.

James Stringer earned a PhD investigating the transcription of Herpes Simplex virus at the University of California, Irvine. He studied SV40 virus while a post-doctoral fellow at Cold Spring Harbor Laboratory. He began studying Pneumocystis organisms at the University of Cincinnati, where he is currently Professor of Molecular Genetics, Biochemistry and Microbiology. His laboratory is investigating diversity, evolution and gene families in Pneumocystis species.

Glossary Terms

Phase variation	Regulation of gene expression in which an individual gene switches between “ON” and “OFF” states. This can be regulated at either the level of transcription initiation or RNA translation.
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.
Gene conversion	Also called duplicative transposition. The copying of an entire gene or segment of a gene from one position in the genome into another. The silent copy of the gene is often referred to as the “donor”, and gene conversion results in its duplication within the genome.
Mutually exclusive expression	The expression of a single gene from within a multi-copy 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.
Epigenetic	Inheritance of given patterns of gene expression which 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-translation modifications to histones are a well-studied example of chromatin marks associated with epigenetic inheritance.

References

1. Berg RD. The indigenous gastrointestinal microflora. Trends Microbiol. 1996; 4:430–435. [PubMed: 8950812]
2. Zeig J, Silverman M, Hilmen H, Simon M. Recombinational switching for gene expression. Science. 1977; 196:170–175. [PubMed: 322276]
3. Merz AJ, So M. Interactions of pathogenic neisseriae with epithelial cell membranes. Annu.Rev.Cell Dev.Biol. 2000; 16:423–457. [PubMed: 11031243]
4. Plant L, Jonsson AB. Contacting the host: insights and implications of pathogenic Neisseria cell interactions. Scand.J.Infect.Dis. 2003; 35:608–613. [PubMed: 14620143]
5. Gray-Owen SD. Neisserial Opa proteins: impact on colonization, dissemination and immunity. Scand.J.Infect.Dis. 2003; 35:614–618. [PubMed: 14620144]
6. Edwards JL, Apicella MA. The molecular mechanisms used by *Neisseria gonorrhoeae* to initiate infection differ between men and women. Clin.Microbiol.Rev. 2004; 17:965–981. [PubMed: 15489357]
7. Leader BT, et al. Antibody responses elicited against the *Treponema pallidum* repeat proteins differ during infection with different isolates of *Treponema pallidum subsp. pallidum*. Infect.Immun. 2003; 71:6054–6057. [PubMed: 14500529]
8. Brinkman MB, et al. Reactivity of antibodies from syphilis patients to a protein array representing the *Treponema pallidum* proteome. J.Clin.Microbiol. 2006; 44:888–891. [PubMed: 16517872]
9. McKevitt M, et al. Genome scale identification of *Treponema pallidum* antigens. Infect.Immun. 2005; 73:4445–4450. [PubMed: 15972547]
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. 2007; 75:104–112. [PubMed: 17030565]
11. Fraser CM, et al. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science. 1998; 281:375–388. [PubMed: 9665876]

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. 2007; 51:289–301. [PubMed: 17683506]
13. Jonsson A, Nyberg G, Normark S. Phase variation of gonococcal pili by frameshift mutation in *pilC*, a novel gene for pilus assembly. EMBO. 1991; 10:477–488.
14. Koomey JM, Gotschlich EC, Robbins K, Berstrom S, Swanson J. Effects of *recA* mutations on pilus antigenic variation and phase transitions in *Neisseria gonorrhoeae*. Genetics. 1987; 117:391–398. [PubMed: 2891588]
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. 1995; 177:2475–2480. [PubMed: 7730280]
16. Blyn LB, Braaten BA, Low DA. Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. EMBO J. 1990; 9:4045–4054. [PubMed: 2147413]
17. Hernday A, Krabbe M, Braaten B, Low D. Self-perpetuating epigenetic pili switches in bacteria. Proc Natl Acad Sci U S A. 2002; 99(Suppl 4):16470–16476. [PubMed: 12202745]
18. Peterson SN, Reich NO. Competitive Lrp and Dam assembly at the pap regulatory region: implications for mechanisms of epigenetic regulation. J Mol Biol. 2008; 383:92–105. [PubMed: 18706913]
19. Holden N, Totsika M, Dixon L, Catherwood K, Gally DL. Regulation of P-fimbrial phase variation frequencies in *Escherichia coli* CFT073. Infect Immun. 2007; 75:3325–3334. [PubMed: 17452474]
20. Domergue R, et al. Nicotinic acid limitation regulates silencing of Candida adhesins during UTI. Science. 2005; 308:866–870. [PubMed: 15774723] [This paper demonstrates how pathogenic yeast can react to environmental conditions to express appropriate surface proteins.]
21. Stern A, Brown M, Nickel P, Meyer TF. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. Cell. 1986; 47:61–67. [PubMed: 3093085]
22. Stern A, Nickel P, Meyer TF, So M. Opacity determinants of *Neisseria gonorrhoeae*: gene expression and chromosomal linkage to the gonococcal pilus gene. Cell. 1984; 37:447–456. [PubMed: 6144397]
23. Keely SP, et al. Gene arrays at *Pneumocystis carinii* telomeres. Genetics. 2005; 170:1589–1600. [PubMed: 15965256] [The Keely et al. paper 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 is able to prolong its survival in the rat lung. Some of the genes must move to a new locus to be transcribed, others are always transcribed but have long mononucleotide tracts in them. These unstable sequences can be expected to suffer frequent frameshift mutations, thereby modulating production of the antigen at the level of translation.]
24. Berendt AR, Ferguson DJP, Newbold CI. Sequestration in *Plasmodium falciparum* malaria: sticky cells and sticky problems. Parasitology Today. 1990; 6:247–254. [PubMed: 15463355]
25. Salanti A, et al. Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. J.Exp.Med. 2004; 200:1197–1203. [PubMed: 15520249]
26. Amulic B, Salanti A, Lavstsen T, Nielsen MA, Deitsch KW. An upstream open reading frame controls translation of *var2csa*, a gene implicated in placental malaria. PLoS.Pathog. 2009; 5:e1000256. [PubMed: 19119419] [This paper represents the first example of RNAi as a mechanism for regulating mutually exclusive expression of a large multi-copy gene family responsible for antigenic variation.]
27. Morris DR, Geballe AP. Upstream open reading frames as regulators of mRNA translation. Mol.Cell Biol. 2000; 20:8635–8642. [PubMed: 11073965]
28. Prucca CG, et al. Antigenic variation in *Giardia lamblia* is regulated by RNA interference. Nature. 2008; 456:750–754. [PubMed: 19079052]
29. Gray RR, et al. Molecular evolution of the tprC, D, I, K, G, and J genes in the pathogenic genus Treponema. Mol.Biol.Evol. 2006; 23:2220–2233. [PubMed: 16926243]
30. Taylor JE, Rudenko G. Switching trypanosome coats: what's in the wardrobe? Trends Genet. 2006; 22:614–620. [PubMed: 16908087]

31. Baruch DI, et al. Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell. 1995; 82:77–87. [PubMed: 7541722]
32. Smith JD, et al. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell. 1995; 82:101–110. [PubMed: 7606775]
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. 1995; 82:89–100. [PubMed: 7606788] [The three papers above (Baruch et al, Smith et al, and Su et al) describe the discovery of the *var* gene family that encodes the primary surface antigen and virulence factor of the human malaria parasite *Plasmodium falciparum*.]
34. Nash TE. Surface antigenic variation in *Giardia lamblia*. Mol.Microbiol. 2002; 45:585–590. [PubMed: 12139606]
35. Allred DR, et al. The *ves* multigene family of *B. bovis* encodes components of rapid antigenic variation at the infected erythrocyte surface. Mol.Cell. 2000; 5:153–162. [PubMed: 10678177]
36. Filler SG. Candida-host cell receptor-ligand interactions. Curr.Opin.Microbiol. 2006; 9:333–339. [PubMed: 16837237]
37. Kitten T, Barbour AG. Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia hermsii*. Proceedings of the National Academy of Sciences USA. 1990; 87:6077–6081.
38. Barbour AG, Burman N, Carter CJ, Kitten T, Bergstrom S. Variable antigen genes of the relapsing fever agent Borrelia hermsii are activated by promoter addition. Mol Microbiol. 1991; 5:489–493. [PubMed: 2041480]
39. Zhang JR, Norris SJ. Genetic variation of the Borrelia burgdorferi gene vlsE involves cassette-specific, segmental gene conversion. Infect Immun. 1998; 66:3698–3704. [PubMed: 9673251]
40. Zhang JR, Norris SJ. Kinetics and in vivo induction of genetic variation of vlsE in Borrelia burgdorferi. Infect Immun. 1998; 66:3689–3697. [PubMed: 9673250]
41. Criss AK, Kline KA, Seifert HS. The frequency and rate of pilin antigenic variation in *Neisseria gonorrhoeae*. Mol.Microbiol. 2005; 58:510–519. [PubMed: 16194236]
42. Serkin CD, Seifert HS. Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. J.Bacteriol. 1998; 180:1955–1958. [PubMed: 9537399]
43. Haas R, Meyer TF. The repertoire of silent pilus genes in *Neisseria gonorrhoeae*: evidence for gene conversion. Cell. 1986; 44:107–115. [PubMed: 2866848]
44. Bernards A, et al. Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. Cell. 1981; 27:497–505. [PubMed: 6101223] [This paper 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 AF, Lundgren A, Yi J, Rurangirwa FR, Palmer GH. Antigenic variation of *Anaplasma marginale* by expression of MSP2 mosaics. Infect Immun. 2000; 68:6133–6138. [PubMed: 11035716]
46. Brayton KA, Palmer GH, Lundgren A, Yi J, Barbet AF. Antigenic variation of *Anaplasma marginale msp2* occurs by combinatorial gene conversion. Mol Microbiol. 2002; 43:1151–1159. [PubMed: 11918803]
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. 2004; 52:1579–1596. [PubMed: 15186410]
48. Sunkin SM, Stringer JR. Residence at the expression site is necessary and sufficient for the transcription of surface antigen genes of *Pneumocystis carinii*. Mol.Microbiol. 1997; 25:147–160. [PubMed: 11902717]
49. Stringer JR. Antigenic variation in *Pneumocystis*. J.Eukaryot.Microbiol. 2007; 54:8–13. [PubMed: 17300510]
50. al Khedery B, Allred DR. Antigenic variation in *Babesia bovis* occurs through segmental gene conversion of the *ves* multigene family, within a bidirectional locus of active transcription. Mol.Microbiol. 2006; 59:402–414. [PubMed: 16390438] [This paper describes the identification

of an expression site for the primary surface antigen of *Babesia bovis*, including the possibility that the genes encoding the two subunits are expressed from a single, bidirectional promoter.]

51. Kraemer SM, et al. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. BMC.Genomics. 2007; 8:45. [PubMed: 17286864]
52. Frank M, et al. Frequent recombination events generate diversity within the multi-copy variant antigen gene families of *Plasmodium falciparum*. Int J Parasitol. 2008; 38:1099–1109. [PubMed: 18395207]
53. 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. 2007; 66:55–73. [PubMed: 17880423] [This paper is the first 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. 2007; 66:220–236. [PubMed: 17784912]
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. 2003; 71:3821–3830. [PubMed: 12819065]
56. Glew MD, et al. Characterization of a multigene family undergoing high-frequency DNA rearrangements and coding for abundant variable surface proteins in *Mycoplasma agalactiae*. Infect.Immun. 2000; 68:4539–4548. [PubMed: 10899853]
57. Lysnyansky I, Rosengarten R, Yogev D. Phenotypic switching of variable surface lipoproteins in *Mycoplasma bovis* involves high-frequency chromosomal rearrangements. J.Bacteriol. 1996; 178:5395–5401. [PubMed: 8808927]
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. 2001; 183:5698–5708. [PubMed: 11544233]
59. Neyrolles O, et al. Antigenic characterization and cytolocalization of P35, the major *Mycoplasma penetrans* antigen. Microbiology. 1999; 145(Pt 2):343–355. [PubMed: 10075417]
60. Bhugra B, Voelker LL, Zou N, Yu H, Dybvig K. Mechanism of antigenic variation in *Mycoplasma pulmonis*: interwoven, site-specific DNA inversions. Mol.Microbiol. 1995; 18:703–714. [PubMed: 8817492]
61. Sitaraman R, Denison AM, Dybvig K. A unique, bifunctional site-specific DNA recombinase from *Mycoplasma pulmonis*. Mol.Microbiol. 2002; 46:1033–1040. [PubMed: 12421309]
62. Roske K, et al. Phase variation among major surface antigens of *Mycoplasma penetrans*. Infect.Immun. 2001; 69:7642–7651. [PubMed: 11705944]
63. Goldmit M, Bergman Y. Monoallelic gene expression: a repertoire of recurrent themes. Immunol.Rev. 2004; 200:197–214. [PubMed: 15242406]
64. Casadesus J, Low D. Epigenetic gene regulation in the bacterial world. Microbiol.Mol.Biol.Rev. 2006; 70:830–856. [PubMed: 16959970]
65. Hakimi MA, Deitsch KW. Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. Curr.Opin.Microbiol. 2007; 10:357–362. [PubMed: 17719264]
66. Cross GAM. Antigenic variation in trypanosomes: secrets surface slowly. Bioessays. 1996; 18:283–291. [PubMed: 8967896]
67. Greaves DR, Borst P. *Trypanosoma brucei* variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion. J.Mol.Biol. 1987; 197:471–483. [PubMed: 2831370]
68. Hughes K, et al. A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. EMBO J. 2007; 26:2400–2410. [PubMed: 17431399]
69. Figueiredo LM, Janzen CJ, Cross GA. A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS.Biol. 2008; 6:e161. [PubMed: 18597556]
70. Gommers-Ampt JH, et al. β-D-Glucosyl-Hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*. Cell. 1993; 75:1129–1136. [PubMed: 8261512]

71. Van Leeuwen F, et al. Localization of the modified base J in telomeric VSG gene expression sites of *Trypanosoma brucei*. Genes Dev. 1997; 11:3232–3241. [PubMed: 9389654]
72. Vanhamme L, et al. Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. Mol.Microbiol. 2000; 36:328–340. [PubMed: 10792720]
73. Gunzl A, et al. RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. Eukaryot.Cell. 2003; 2:542–551. [PubMed: 12796299]
74. Navarro M, Gull K. A pol I transcriptional body associated with VSG monoallelic expression in *Trypanosoma brucei*. Nature. 2001; 414:759–763. [PubMed: 11742402] [This paper provides the first evidence for a specific subnuclear expression site where active transcription of the gene encoding the primary cell surface antigen takes place. This is 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. 1998; 17:5418–5426. [PubMed: 9736619]
76. Freitas-Junior LH, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell. 2005; 121:25–36. [PubMed: 15820676]
77. Chookajorn T, et al. Epigenetic memory at malaria virulence genes. Proc.Natl.Acad.Sci.U.S.A. 2007; 104:899–902. [PubMed: 17209011]
78. Lopez-Rubio JJ, et al. 5′ flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol.Microbiol. 2007; 66:1296–1305. [PubMed: 18028313]
79. Duraisingh MT, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. Cell. 2005; 121:13–24. [PubMed: 15820675] [This paper and that published by reference 72 (Freitas-Junior et al) provide the first description of chromatin modifications associated with mutually exclusive *var* gene expression in malaria parasites.]
80. Ralph SA, Scheidig-Benatar C, Scherf A. Antigenic variation in *Plasmodium falciparum* is associated with movement of *var* loci between subnuclear locations. Proc.Natl.Acad.Sci.U.S.A. 2005; 102:5414–5419. [PubMed: 15797990]
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. 2007; 63:1237–1247. [PubMed: 17257309]
82. Dzikowski R, et al. Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO Rep. 2007; 8:959–965. [PubMed: 17762879]
83. Dzikowski R, Deitsch KW. Active transcription is required for maintenance of epigenetic memory in the malaria parasite *Plasmodium falciparum*. J Mol Biol. 2008; 382:288–297. [PubMed: 18656490]
84. Kulakova L, Singer SM, Conrad J, Nash TE. Epigenetic mechanisms are involved in the control of *Giardia lamblia* antigenic variation. Mol.Microbiol. 2006; 61:1533–1542. [PubMed: 16968226]
85. De Las PA, et al. Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to. Genes Dev. 2003; 17:2245–2258. [PubMed: 12952896]
86. Stoenner HG, Dodd T, Larsen C. Antigenic variation of *Borrelia hermsii*. Journal of Experimental Medicine. 1982; 156:1297–1311. [PubMed: 7130900]
87. Dai Q, et al. Antigenic variation by *Borrelia hermsii* occurs through recombination between extragenic repetitive elements on linear plasmids. Mol.Microbiol. 2006; 60:1329–1343. [PubMed: 16796672]
88. Barbour AG, Dai Q, Restrepo BI, Stoenner HG, Frank SA. Pathogen escape from host immunity by a genome program for antigenic variation. Proc.Natl.Acad.Sci.U.S.A. 2006; 103:18290–18295. [PubMed: 17101971]

89. Marcello L, Barry JD. From silent genes to noisy populations-dialogue between the genotype and phenotypes of antigenic variation. J.Eukaryot.Microbiol. 2007; 54:14–17. [PubMed: 17300511]
90. Frank M, Dzikowski R, Amulic B, Deitsch K. Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol.Microbiol. 2007; 64:1486–1498. [PubMed: 17555435]
91. Horrocks P, Pinches R, Christodoulou Z, Kyes SA, Newbold CI. Variable *var* transition rates underlie antigenic variation in malaria. Proc.Natl.Acad.Sci.U.S.A. 2004; 101:11129–11134. [PubMed: 15256597]
92. Jensen AT, et al. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. J.Exp.Med. 2004; 199:1179–1190. [PubMed: 15123742]
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. 2006; 74:3904–3911. [PubMed: 16790763]
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. 2006; 193:1567–1574. [PubMed: 16652286]
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. Molecular Microbiology. 2003; 49:179–191. [PubMed: 12823820]
96. Cadavid D, Thomas DD, Crawley R, Barbour AG. Variability of a bacterial surface protein and disease expression in a possible mouse model of systemic Lyme borreliosis. J.Exp.Med. 1994; 179:631–642. [PubMed: 8294872]
97. Alugupalli KR, et al. The resolution of relapsing fever borreliosis requires IgM and is concurrent with expansion of B1b lymphocytes. J Immunol. 2003; 170:3819–3827. [PubMed: 12646649]
98. Fikrig E, et al. Sera from patients with chronic Lyme disease protect mice from Lyme borreliosis. J Infect Dis. 1994; 169:568–574. [PubMed: 8158028]
99. Barthold SW, Bockenstedt LK. Passive immunizing activity of sera from mice infected with *Borrelia burgdorferi*. Infect Immun. 1993; 61:4696–4702. [PubMed: 8406868]
100. Johnson RC, Kodner C, Russell M. Passive immunization of hamsters against experimental infection with the Lyme disease spirochete. Infect Immun. 1986; 53:713–714. [PubMed: 3744563]
101. Lukehart SA, Miller JN. Demonstration of the in vitro phagocytosis of *Treponema pallidum* by rabbit peritoneal macrophages. J Immunol. 1978; 121:2014–2024. [PubMed: 361893]
102. Morgan CA, Molini BJ, Lukehart SA, Van Voorhis WC. Segregation of B and T cell epitopes of *Treponema pallidum* repeat protein K to variable and conserved regions during experimental syphilis infection. J Immunol. 2002; 169:952–957. [PubMed: 12097401]
103. LaFond RE, Molini BJ, Van Voorhis WC, Lukehart SA. Antigenic variation of TprK V regions abrogates specific antibody binding in syphilis. Infect Immun. 2006; 74:6244–6251. [PubMed: 16923793]
104. McGregor IA. Studies in the acquisition of immunity of *Plasmodium falciparum* infections in Africa. Trans R Soc Trop Med Hyg. 1964; 58:80–92. [PubMed: 14117559]
105. Baruch DI, Gormley JA, Ma C, Howard RJ, Pasloske BL. *Plasmodium falciparum* erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. Proceedings of the National Academy of Sciences USA. 1996; 93:3497–3502.
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. 1995; 182:409–418. [PubMed: 7629503]
107. Badell E, et al. Human malaria in immunocompromised mice: an in vivo model to study defense mechanisms against *Plasmodium falciparum*. J Exp.Med. 2000; 192:1653–1660. [PubMed: 11104807]

108. O'Connor RM, Allred DR. Selection of *Babesia bovis*-infected erythrocytes for adhesion to endothelial cells coselects for altered variant erythrocyte surface antigen isoforms. J Immunol. 2000; 164:2037–2045. [PubMed: 10657656]

109. Baker-Zander SA, Shaffer JM, Lukehart SA. 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. 1993; 6:273–279. [PubMed: 8499892]

110. Lukehart SA, Shaffer JM, Baker-Zander SA. A subpopulation of *Treponema pallidum* is resistant to phagocytosis: possible mechanism of persistence. J Infect Dis. 1992; 166:1449–1453. [PubMed: 1431264]

111. Sell S, Gamboa D, Baker-Zander SA, Lukehart SA, Miller JN. Host response to *Treponema pallidum* in intradermally-infected rabbits: evidence for persistence of infection at local and distant sites. J Invest Dermatol. 1980; 75:470–475. [PubMed: 7003026]

112. Lukehart SA, Baker-Zander SA, Lloyd RM, Sell S. Characterization of lymphocyte responsiveness in early experimental syphilis. II. Nature of cellular infiltration and *Treponema pallidum* distribution in testicular lesions. J Immunol. 1980; 124:461–467. [PubMed: 6153103]

113. Baker-Zander SA, Lukehart SA. Macrophage-mediated killing of opsonized *Treponema pallidum*. J Infect Dis. 1992; 165:69–74. [PubMed: 1727898]

114. Futse JE, Brayton KA, Dark MJ, Knowles DP Jr. Palmer GH. Superinfection as a driver of genomic diversification in antigenically variant pathogens. Proc Natl Acad Sci U S A. 2008; 105:2123–2127. [PubMed: 18252822]

115. Montgomery J, et al. Differential *var* gene expression in the organs of patients dying of *falciparum* malaria. Mol.Microbiol. 2007; 65:959–967. [PubMed: 17617167]

116. Miller LH, Good MF, Milon G. Malaria Pathogenesis. Science. 1994; 264:1878–1883. [PubMed: 8009217]

117. Barry AE, et al. Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*. PLoS.Pathog. 2007; 3:e34. [PubMed: 17367208]

118. Anderson RM, May RM. Population biology of infectious diseases: Part I. Nature. 1979; 280:361–367. [PubMed: 460412]

The evolutionary success of an infectious agent is dependent upon several factors, only some of which are determined by the pathogen. This concept has been elegantly expressed by Anderson & May[118] who quantified the success of an infectious agent according to the equation $R_o = \beta cD$, in which R_o is the average number of new infections caused by a single infected person. If this number is >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 ability to disarm the effect of acquired immunity by changing epitopes. Successful evasion of host immunity results in increased value for D, the duration of the infectious state.

Online Summary

❖ Mammals have evolved an elaborate, multifaceted immune system to respond to the ever-present threat of infection by pathogenic microorganisms. Bacterial, protozoan and fungal pathogens have responded by evolving equally elaborate systems to avoid destruction by their hosts. This process of co-evolution has resulted in the development of complex genetic systems underlying antigenic variation by numerous pathogenic microorganisms.

❖ The process of antigenic variation is focused at the host/pathogen interface, and in particular at the cell surface of the infectious organisms. Molecules displayed on the pathogen cell surface often mediate adhesion within specific niches and are frequently virulence determinants.

❖ Some systems of antigenic variation involve activation and silencing of genes that encode molecules exposed to the immune system of the infected host. In its simplest form, this entails changes in the expression of genes that are regulated individually, an ON/OFF process referred to as phase variation.

❖ In other organisms, a single expression site is present for a key protein, with multiple silent gene copies or cassettes existing elsewhere in the genome. The sequence of the expressed gene changes by gene conversion (or duplicative transposition) of large or small DNA sequences from the silent pseudogenes into the expression site.

❖ In more sophisticated systems, the pathogen has evolved large, multi-copy gene families with each copy encoding a different form of the surface antigen. In these organisms, each individual gene has all of the elements necessary for expression, and each undergoes silencing and activation as described above; however, an additional layer of regulation exists to ensure that only a single gene is active at any given time. Thus gene silencing and activation within the family is coordinated and strictly mutually exclusive.

❖ While many of the genetic systems underlying antigenic variation, for instance slipped-strand mispairing or gene conversion, involve alterations to the genome, in several organisms changes in gene expression do not involve any alterations in the primary DNA sequence. These systems instead rely on “epigenetic” modifications to control gene activation and silencing, the hallmarks of which include histone modifications, the use of modified nucleotides, changes in chromatin structure, and nuclear organization.

❖ In a few cases, the order in which specific antigen variants are expressed over the course of an infection is determined by the sequence of the encoding genes. This can help to extend the length of an infection or the infectious stage, thus increasing the likelihood of transmission to a new host.

❖ Antigenic variation also enhances the ability of a pathogen to infect a host that has resolved (or been cured of) prior infection (i.e., re-infection) or who is persistently infected with the same organism (superinfection). This both expands the population of susceptible hosts and permits genetic exchange between organisms.

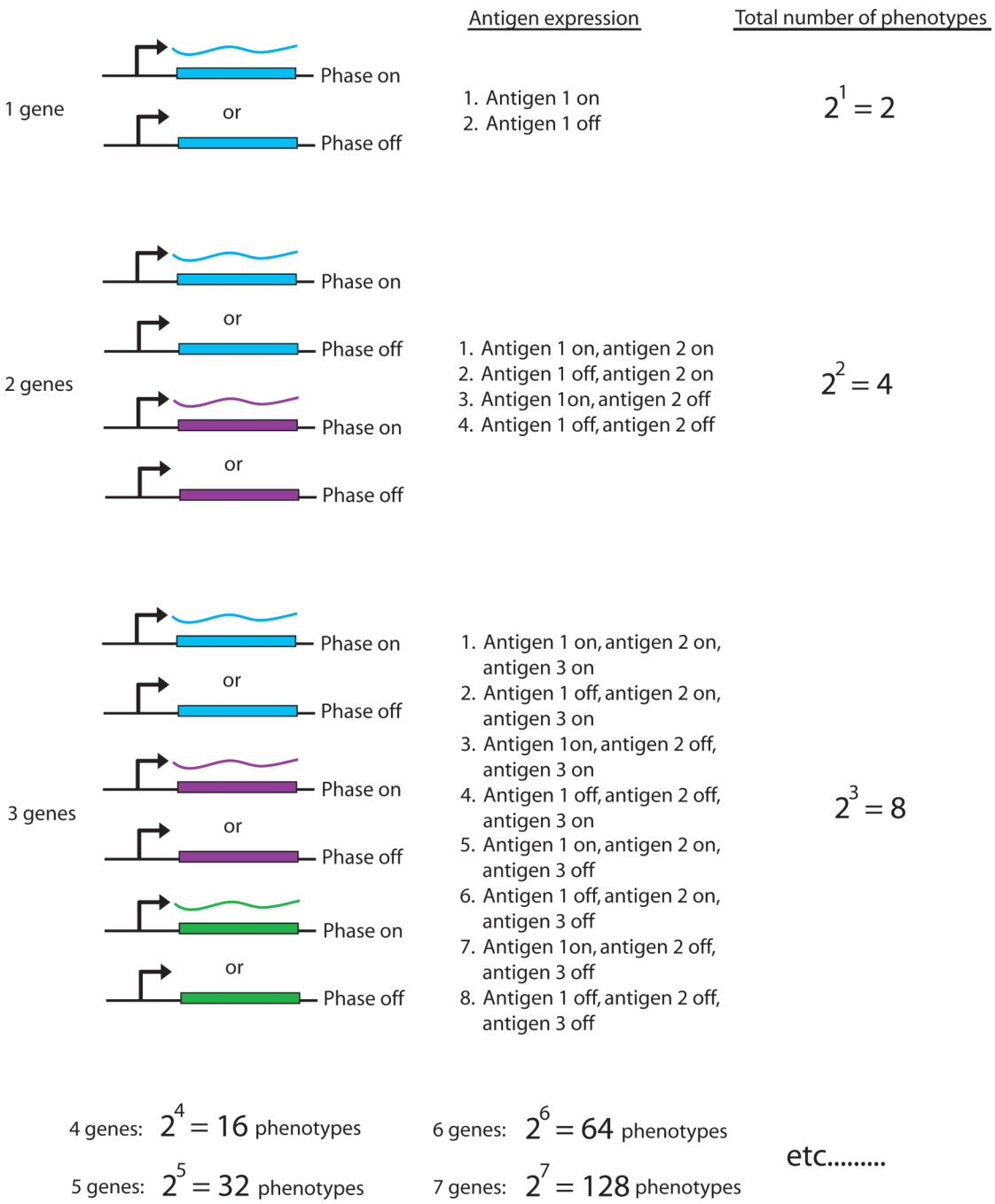

Figure 1.

Example of how increasing numbers of phase variant genes can contribute to large numbers of phenotypes. Observe how the total number of phenotypes can be calculated as 2^n, where N represents the number of independently regulated genes.

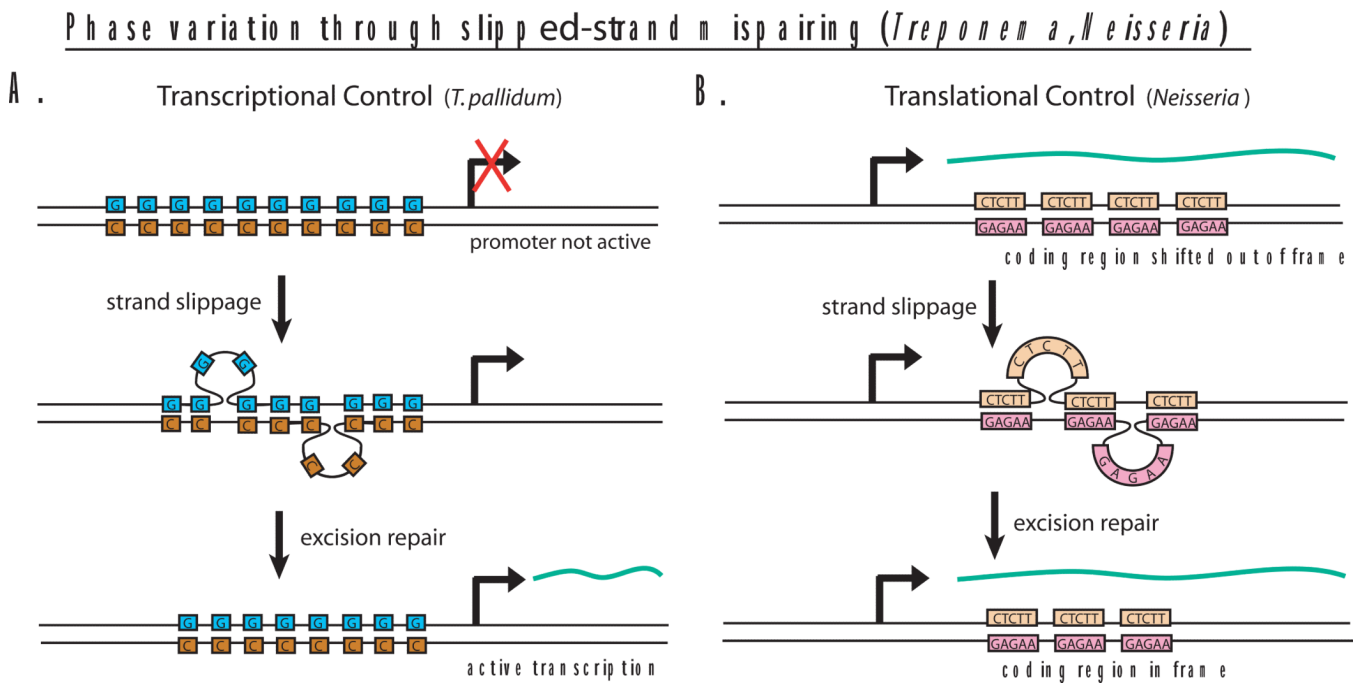

Figure 2.

Schematic diagram showing 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* (left) 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 within the open reading frame can regulated Opa expression in Neisseria at the level of translation.

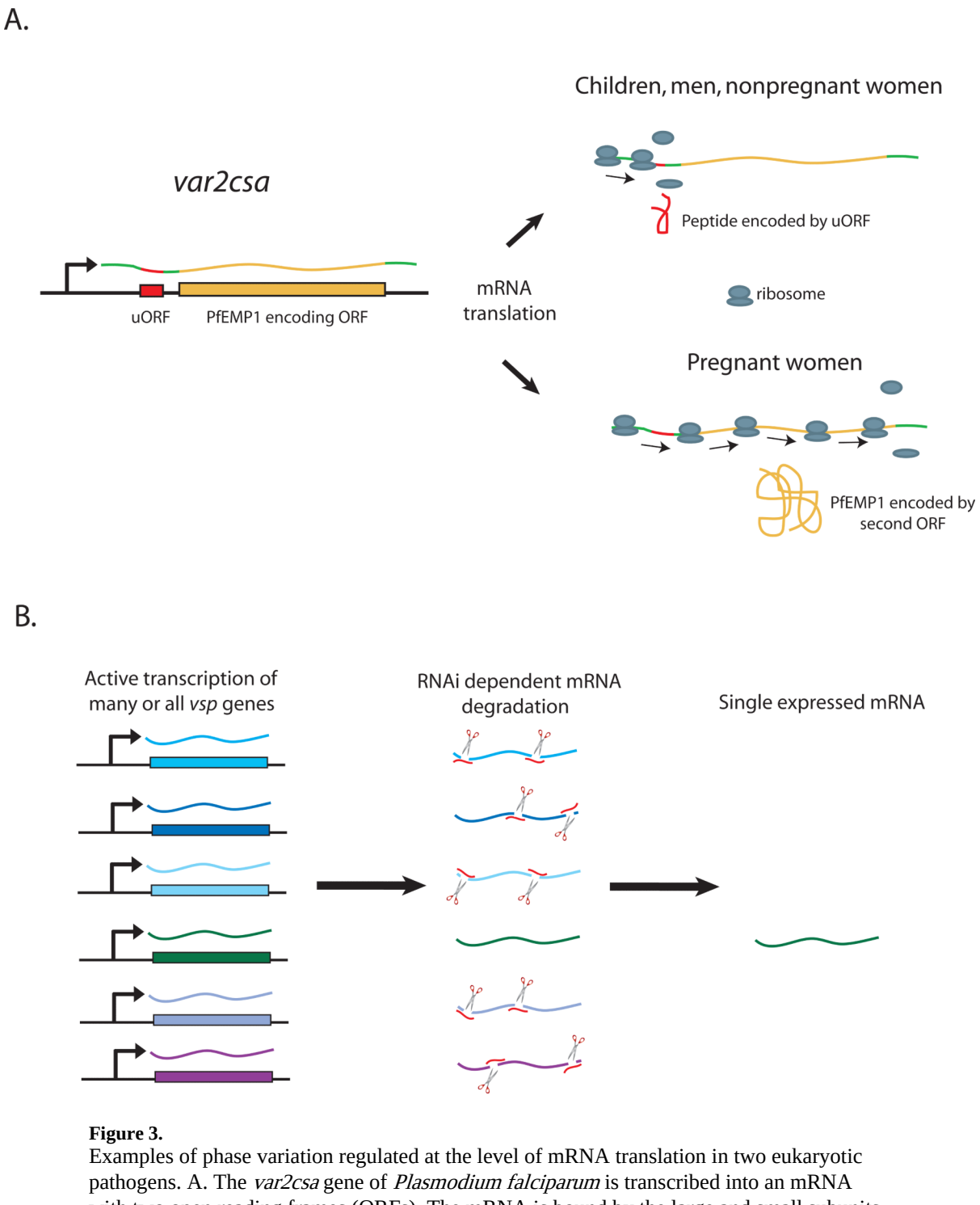

Figure 3.

Examples of phase variation regulated at the level of mRNA translation in two eukaryotic pathogens. 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 (gray circles), which then moves along the transcript (arrows) until it initiates translation at a start codon. The first ORF (uORF) is translated when the parasites infect children, men or nonpregnant women, however the ribosome is thought to dissociate from the mRNA prior to reaching the second ORF, thus preventing expression of the encoded form of PfEMP1. The second ORF is only translated when parasites infect pregnant women and the encoded form of PfEMP1 only functions in the presence of a placenta. B.

Mutually exclusive expression of the *vsp* gene of *Giardia lamblia* is achieved through the RNAi pathway. Small RNAs (red strands) target the RNAi degradation machinery (scissors) to mRNAs from all but a single *vsp* gene. Only mRNA from one gene escapes degradation and is translated into protein, although the mechanism that enables this escape is unknown.

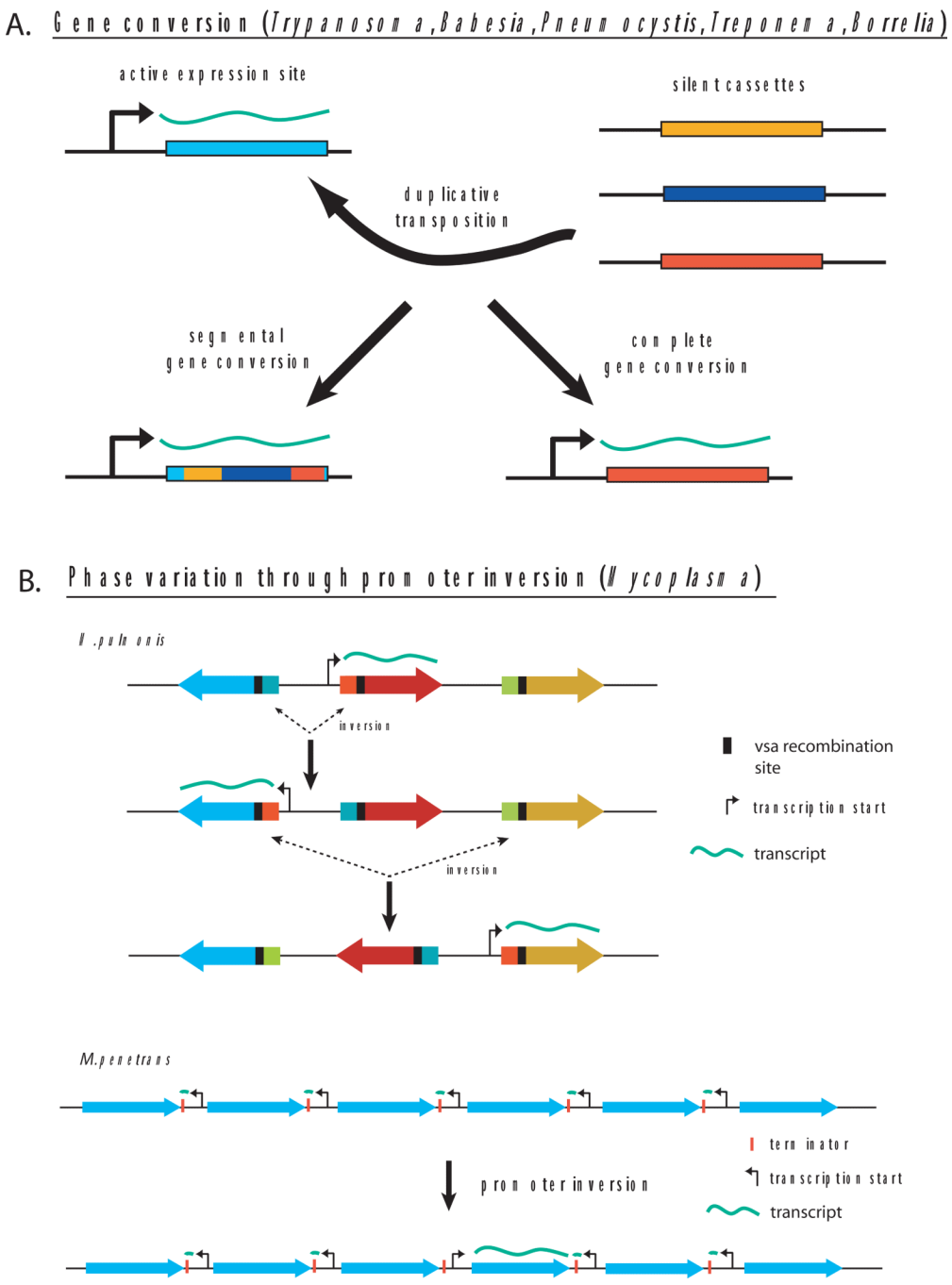

Figure 4.

Antigenic variation through mechanisms that rely on DNA recombination. A. Gene conversion resulting in the duplication of a segment of a silent gene into an active expression site. The duplicated segment can include an entire open reading frame or small portions, creating a chimeric gene as illustrated here. B. Promoter inversion. As shown for *M. pulmonis*, inversion of a single promoter leads to expression of one *vsa* gene at a time (top panel). Inversion occurs by recombination between regions of sequence similarity (black boxes). In the *mpl* genes of *M. penetrans* (bottom panel), tandem arrays of genes are each regulated by separate promoters. In one orientation, a promoter drives the expression of a short transcript (ended by a transcriptional terminator, red box) that does not encode

protein. When inverted, the promoter drives transcription through the *mpl* open reading frame, leading to expression of the encoded protein.

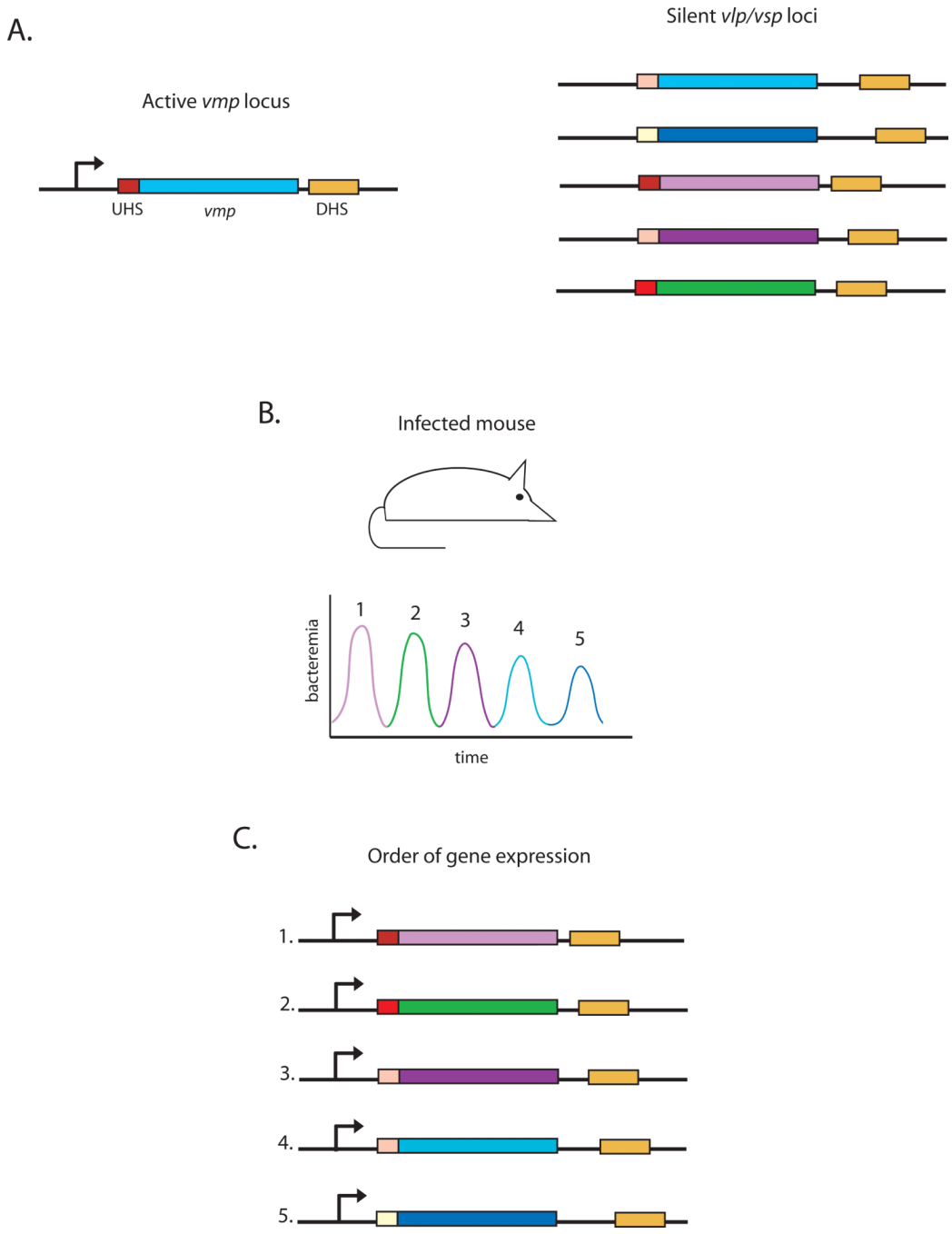

Figure 5.

A loose order of expression can be programmed into large multi-copy gene families. *vmp* gene expression in Borrelia is an example. A. The active *vmp* gene contains two elements that contribute to expression switching called the UHS and the DHS. These elements are also found within the silent genes, with varying degrees of sequence identity (displayed here as differing degrees of dark red color). 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 degree of sequence identity within the UHS and the distance between the DHS and the coding region.

Table 1

Examples of multi-copy sequence repertoires encoding variant surface antigens.

Species	Gene	Copy number	Encoded surface antigen
Bacterial Systems			
Borrelia hermsii	*vsp, vlp*	~30, ~30	variable short protein (VSP), variable long protein (VLP)
Neisseria meningitidis	*pil, opa*	~19, 4-5	type IV pili, opacity proteins
Neisseria gonorrhoeae	*pil, opa*	~19, 11	type IV pili, opacity proteins
Treponema pallidum	*tpr*	12	*T. pallidum* repeat proteins A-L, some are predicted to be surface exposed
Protozoan Systems			
Trypanosoma brucei	*vsg*	~1000	variant surface glycoprotein (VSG)
Babesia bovis	*vesa, vesβ*	~130-160	variant expressed surface antigen (VESA)
Giardia lamblia	*vsp*	~150	variant surface protein (VSP)
Plasmodium falciparum	*var*	~60	*Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1)
Fungal Systems			
Pneumocystis carnii	*msg, msr, prt*	~85 of each	major surface glycoprotein (MSG), MSG related protein (MSR), protease 1 (PRT1)
Candida albicans	*als*	9	agglutinin like sequence (ALS)
Candida glabrata	*epa*	17-23	epithelial adhesin (EPA)