

REVIEW

# Revealing the molecular signatures of host-pathogen interactions

Chiea-Chuen Khor<sup>1,2,3</sup> and Martin L Hibberd<sup>1,4\*</sup>

## Abstract

Advances in sequencing technology and genome-wide association studies are now revealing the complex interactions between hosts and pathogen through genomic variation signatures, which arise from evolutionary co-existence.

Infectious agents have a long history of causing disease in humans and are a major contributor of childhood mortality worldwide (Table 1). For many infectious agents, a complex interaction between the host and pathogen has developed over time as a result of an evolutionary co-existence and adaptation that is now reflected in both human and pathogen genomes. Initial studies have looked at these human-pathogen interactions on a broad scale, asking what components of the (typically human) genome are associated with disease susceptibility [1]. However, infectious diseases, unlike many other diseases, offer the chance to dissect this more precisely as the non-host or 'environmental' component (the pathogen) has a genome that can be assayed just as accurately as the human genome. Today, we have the appropriate tools and platforms in place to begin examining these adaptations from a whole-genome perspective, using samples directly from clinical studies. By directly observing this human-pathogen interaction *in natura*, we can avoid the requirement for artificial models, which tend to be tedious to construct and may not reconstitute human physiology accurately [2]. This generates a powerful approach, as it integrates two complementary (but previously thought to be distinct) areas of infectious disease research - host genetic susceptibility to infection outcomes and pathogen genetic variation leading to differential virulence and individual susceptibility.

On the basis of data obtained using biological candidate approaches, it has been assumed that broad-based responses mediated by B and T cells (both CD4<sup>+</sup> and CD8<sup>+</sup>) were the primary defenses against infectious agents and that these were sufficient for protection against most infecting pathogens [3]. Defects in Interleukin 2 (IL-2) signaling resulted in failed clonal expansion of both B and T cells, leading to severe combined immune deficiency [3]. Other examples include X-linked agammaglobulinemia - caused by mutations in Bruton's tyrosine kinase - where no mature antibody-producing B cells are produced, resulting in no antibody production [4]. However, in the majority of common infectious diseases studied so far from a genomic and genome-wide association perspective, the evidence has tended to point away from a picture of high-penetrance, generalized susceptibility caused by critical defects in single genes, and instead towards a more complex picture of multiple lower penetrance effects implicating individual genetic components for each invading pathogen. There are some notable exceptions to this, such as the shared human leukocyte antigen (HLA) associations (suggesting T-cell involvement) in leprosy susceptibility [5] and HIV viral set point [6]. As HLA is crucial for recognition of processed pathogen molecules and the initiation of the CD4<sup>+</sup> and CD8<sup>+</sup> T-cell response, it was unsurprising that some degree of shared susceptibility between pathogens is observed.

Recent evidence has suggested an important role for innate immune interactions in disease susceptibility [5,7] and specific evolution and adaptation of both host and pathogen genomes to the current state of mutual co-existence. Genome-wide association studies (GWASs) necessarily assay common genetic variation of mostly low penetrance, and this approach is now being used to study host genetic susceptibility to infectious diseases (Table 2). Some of these recent GWAS data on infectious diseases have indeed pointed clearly to surprisingly specific pathogen-receptor interactions rather than broad-based susceptibilities. Similar advancements have been made from the pathogen perspective, in which GWASs on the pathogen genome have also yielded unexpected insights. The example in *Plasmodium falciparum* malaria

\*Correspondence: hibberdml@gis.a-star.edu.sg

<sup>1</sup>Infectious Diseases, Genome Institute of Singapore, 60 Biopolis Street, #02-01 Genome Building, Singapore 138672

Full list of author information is available at the end of the article

**Table 1. Global causes of childhood deaths**

Cause	%
Pneumonia	18
Malaria	8
Sepsis	6
AIDS	2
Pertussis	2
Meningitis	2
Measles	1
Tetanus	1
Other infections	9
Preterm birth complications	12
Birth asphyxia	9
Congenital abnormalities	3
Diarrhea	15
Injury	3
Others	9

Infections contribute to 49% of childhood deaths globally. Causes that led to less than 1% of deaths are not presented. Adapted from [61].

showcases the adaptability of the pathogen genome when confronted by the human immune response and anti-malarial pharmacotherapy [8-10].

In contrast to highly penetrant Mendelian genetic defects causing broad-based susceptibilities with severe clinical outcomes, more common host genetic determinants of susceptibility of lower penetrance and modest effect size could be confined to specific pathogen species [11] or even serotypes (and possibly genotypes). A better understanding of host-pathogen interactions will allow more thorough dissection of the host immune response generated by specific pathogen invasion. Because pathogens adapt to survival pressures by evading drugs, vaccines, and the host immune response and have a higher intrinsic mutation rate than humans, this improved understanding will potentially have an impact on vaccine design and novel therapy discovery. In this review, we focus on three examples of this - dengue, Malaria, and meningococcal disease - to reflect the literature currently available.

### Host-pathogen interactions with dengue virus

Dengue fever is caused by infection from one of four serotypes of the dengue (DEN) virus (DEN-1 to DEN-4), which belong to the Flaviviridae family (other members of which include the yellow fever and Hepatitis C viruses) [12,13]. It is an acute systemic viral infection with a wide spectrum of disease manifestations, ranging from sub-clinical infection to severe and fatal disease. The commonest severe complications are a transient increase in vascular permeability and altered hemostasis, which

could lead to life-threatening hypovolemic circulatory shock (called dengue shock syndrome, DSS). It is the most common mosquito-borne infection after malaria, and its recent resurgence is largely attributed to a combination of factors, including exponential population growth, rapid urbanization, air travel and lack of proper vector control [14]. Previous studies have provided clues regarding the importance of pathogen genetic variation and severity of the infection. Amino acid variants determined from full-genome sequencing of the dengue virus revealed different potentials for causing severe dengue as opposed to uncomplicated dengue [15]. Current evidence also suggests that different serotypes of dengue virus are associated with more or less severe disease, following infections with DEN-2 [16-19].

From the human perspective, previous studies have pointed to several branches of the immune system in the pathogenesis of dengue [15-19], including the HLA system and dendritic cells, although there are few convincing genetic validation studies. However, a recent large-scale GWAS and replication study showed very strong statistical evidence of association between genetic variants at two distinct loci (*MICB* and *PLCE1*) and increased susceptibility to DSS (Table 2) [20]. *MICB* encodes MHC class I polypeptide-related sequence B, which is an inducible activating ligand for the NKG2D type II receptor on natural killer (NK) cells. NK cells are distinct from B or T cells and are crucial for the early response to viral infections and in shaping the subsequent adaptive immune response to viral infection [21-23]. Interestingly another recent GWAS found that a genetic variant in the closely related *MICA* gene is strongly associated with Hepatitis C virus-induced hepatocellular carcinoma, suggesting a pivotal role for MIC proteins in the pathogenesis of these Flaviviridae infections [24]. *PLCE1* encodes phospholipase C epsilon gamma, and missense mutations in *PLCE1* cause nephrotic syndrome [25], a kidney disorder in which dysfunction of the glomerulus basement membrane results in proteinuria and hypoproteinemia that, when severe, leads to reduced vascular oncotic pressure and edema. These elements of nephrotic syndrome have striking similarities with the hypovolemic shock in severe dengue and suggest an important role for *PLCE1* in maintaining normal vascular endothelial cell barrier function. These associations with *MICB* and *PLCE1* are not serotype specific, but instead are applicable across all four dengue virus serotypes [20].

The GWAS on severe dengue has revealed unexpected insights into its pathogenesis, as it points specifically to NK cell pathology, shared susceptibility pathways with shock, and possibly shared pathology with other Flaviviridae [20]. However, the locations of the causal genetic variants responsible for the observed associations are still unknown and await elucidation through fine-scale

**Table 2. Host susceptibility genes identified by GWAS studies**

Pathogen	Phenotype	Gene	N cases/controls	Effect size	P-value	References
Dengue virus	Severe dengue	<i>MICB</i>	3,742/4,952	1.34	$4.41 \times 10^{-11}$	[20]
	Severe dengue	<i>PLCE1</i>	3,742/4,952	0.80	$3.08 \times 10^{-10}$	[20]
<i>N. meningitidis</i> <sup>a</sup>	Meningococcal disease	<i>CFH</i>	1,443/6,079	0.64	$2.2 \times 10^{-11}$	[7]
<i>P. falciparum</i> <sup>a</sup>	Severe malaria	<i>ABO</i>	3,906/5,489	1.20	$2.0 \times 10^{-7}$	[49,62]
	Severe malaria	<i>HBB</i>	2,045/3,758	0.63	$3.7 \times 10^{-11}$	[62]
<i>M. leprae</i>	Leprosy	<i>NOD2</i>	3,960/7,180	1.59	$3.77 \times 10^{-40}$	[5]
		<i>RIPK2</i>	3,960/7,180	0.76	$1.38 \times 10^{-16}$	[5]
		<i>HLA-DR</i>	3,960/7,180	0.67	$5.35 \times 10^{-27}$	[5]
		<i>TNFSF15</i>	3,960/7,180	1.37	$3.39 \times 10^{-21}$	[5]
		<i>C13orf31-CCDC122</i>	3,960/7,180	1.68	$3.72 \times 10^{-54}$	[5]
<i>M. tuberculosis</i>	Tuberculosis	-	3,632/7,501	1.19	$6.8 \times 10^{-9}$	[63]
HIV	Viral set point <sup>a</sup>	<i>HLA-B</i>	486	9.6%	$9.36 \times 10^{-12}$	[6]
	Viral set point <sup>a</sup>	<i>HLA-C</i>	486	6.5%	$3.77 \times 10^{-9}$	[6]
HBV	Chronic infection	<i>HLA-DP</i>	2,086/4,301	1.75	$6.34 \times 10^{-39}$	[64,65]

<sup>a</sup>Analyzed as a quantitative trait.

mapping. What is notable is that these observations suggest that subtle genomic differences in members of one viral family affect related, yet distinct, components of the human immune response. But is this molecular pathway of susceptibility the only route to Flaviviridae infection or is this shared pathway the only one for a wide variety of viruses? The genetic variation at both *MICB* and *PLCE1* explained no more than 2.5% of the overall heritable variance in susceptibility to DSS, whereas very little evidence of disease association was observed elsewhere in the genome. Will a finer dissection of the infecting pathogen genome reveal more and discover some of the 'missing heritability'? Whole-genome sequencing of the dengue virus isolated from patients with clinical data has been initiated on a large scale [19,26-29], and identification of mutations in dengue virus genes and their systematic reconciliation with human genotype and clinical phenotype will be possible in the near future.

### The impact of host and pathogen genetic variations on individual susceptibility to meningococcal disease

*Neisseria meningitidis* is a Gram-negative, polysaccharide-encapsulated bacterium and is the cause of meningococcal meningitis and sepsis, which are potentially fatal infections without antibiotic treatment. Five out of the twelve identified *N. meningitidis* serotypes have been reported to cause epidemics (A, B, C, W135, and Y) [30]. Subgroup analysis with specific meningococcal serotypes has revealed differential virulence capabilities of each serotype, with serotypes B and C causing more adverse infection outcomes than the others [31-33]. Unlike for

meningococcal serotypes A, C, Y, and W135, a vaccine is yet to be developed for serotype B *N. meningitidis*, thus highlighting the need for improved understanding of the host-pathogen interactions involved. Earlier genetic studies examining human susceptibility to *N. meningitidis* sepsis were limited to known genes whose biological functions have been well characterized [34-36], and many were likely to be false-positive findings as they subsequently were not successfully validated by independent studies [37]. It is hoped that by better defining the critical elements at the host-pathogen interface using unbiased whole-genome approaches, more definitive answers on the susceptibility determinants of overall meningococcal infection and the extent of the immunogenic differences between serotype B *N. meningitidis* and the other serotypes will emerge.

A recent case-control GWAS study [7] on *N. meningitidis* sepsis (which included serotypes A, B, C, Y, and W-135) in a UK collection, with validation in sample collections from Austria, the Netherlands, and Spain, showed significant evidence of association between genetic markers in Factor H (*CFH*) and *CFH*-related genes (*CFHR3*, *CFHR1*) on chromosome 1 and decreased susceptibility to meningococcal disease (Table 2). *CFH* and *CFHR3* are atypical members of the complement family in that they are negative regulators of complement signaling, protecting host cells from destruction through complement activation in response to infection [38].

Previous human mutation studies had shown complement deficiency to be a susceptibility factor for many *Neisseria* species [39], so perhaps this study [38] may not seem surprising. However, when examined, a new mechanism was revealed. It is not the absence of complement

that is important to the occurrence of widespread disease, but rather the presence of the factor H complement regulator. This means that the addition of complement to patients who were deficient in complement pathway components might not prevent them from suffering repeated disease, whereas the removal of the complement regulator might, because CFH protects the bacteria through a complex interaction of the host protein and bacteria (Figure 1). This new mechanistic understanding of the disease may have other implications as well, because the associations observed with *CFH* and *CFHR* genetic variants were significantly stronger (per-allele increased risk of disease between 1.6-fold and 1.8-fold compared with baseline; *P*-value about  $10^{-5}$  to  $10^{-8}$ ) in collections that had a variety of *N. meningitidis* serogroup infections (enrolled before wide-spread serotype C conjugate vaccination), compared with those collections collected after widespread serotype C vaccination was adopted (Spain; per-allele increase in risk about 1.3- to 1.4-fold, *P*-value about  $10^{-2}$  to  $10^{-3}$ ). The substantially weaker association at the *CFH-CFHR3-CFHR1* locus in the predominantly serotype B sample collection implies that serotype B *N. meningitidis* has subtly distinct mechanisms of host interaction from those of other serotypes and may explain why serotype B *N. meningitidis* has been refractory to vaccine development attempts.

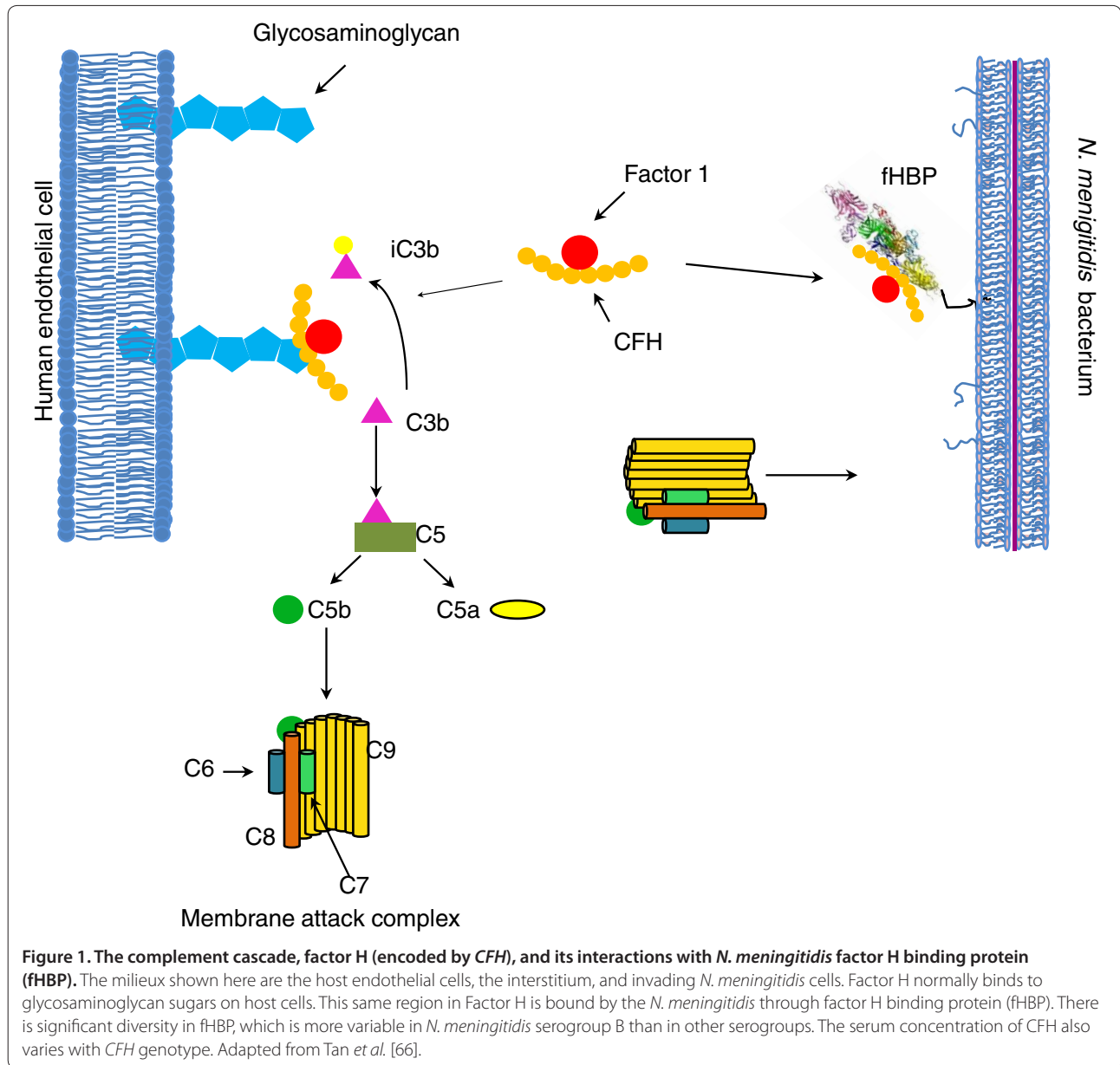
*N. meningitidis* defends itself from host complement-mediated killing through sequestration of human CFH by factor H binding protein (fHBP), a surface lipoprotein present on all strains of *N. meningitidis* [40]. Owing to its antigenic potential and ability to induce bactericidal antibodies, fHBP is one of the antigens incorporated in the recombinant vaccine for group B *N. meningitidis* in vaccine trials. This GWAS observation [7] thereby confirmed an interaction between *N. meningitidis* and human CFH, mediated in a complex manner through fHBP, which was not previously detected in mutation studies of patients suffering from recurrent meningococcal infections due to complement deficiency (Figure 1). As different allelic variants of CFH itself resulted in easily discernable differences in actual human susceptibility to meningococcal disease, therapeutic designs based on synergistic action targeting both fHBP and CFH remain a possibility. Such an approach could be more effective than manipulation of either molecule alone. The fHBP gene has recently been sequenced in various serotypes, revealing significantly higher sequence diversity in serotype-B than non-serotype-B *N. meningitidis*. This makes serotype-B *N. meningitidis* more challenging to target from the point of view of both the host and the vaccinologist [41]. This could in part explain the much weaker association signal seen in the GWAS and implies a correlation between increased fHBP

variability and virulence, which needs to be confirmed by future work.

### ***Plasmodium falciparum***

Malaria is a predominantly tropical disease vectored by the *Anopheles* mosquito, and up to 40% of the worldwide population is at risk of infection. Although infection by four different species of *Plasmodium* (*falciparum*, *malariae*, *ovale*, and *vivax*) is causative of malaria, *P. falciparum* is the predominant form globally. It also results in higher complication and mortality rates than the other three species [42]. Susceptibility to and severity as a result of *Plasmodium* infection are well studied examples of how host genetic factors affect the disease process. Natural genetic variation resulting in the sickle cell trait and Duffy blood group antigens are two clear examples whereby the human genome adapts specifically in response to an infectious agent. The parasite also exerts considerable selective pressure on the human genome [43], underlying yet again the importance of understanding the host-pathogen interface from both perspectives.

Although the immune responses to malaria have been extensively documented in both animal and human models [44,45], robust evidence on the specific response that is critical for protective immunity remains elusive [46,47]; often, the induction of a high degree of immunogenicity does not correlate with protective immunity [48]. GWASs have now made it possible to search systematically for strong association between functional variation in a given immune gene and susceptibility to (or severity of) infection to identify the specific immune responses responsible for protective immunity. All genetic studies for human susceptibility to malaria have yielded two consistent results: the sickle hemoglobin trait is associated with a 5- to 10-fold reduced susceptibility to severe malaria, and the ABO blood group is associated with a more modest reduction (a per-allele odds ratio of 1.2-fold increased risk for blood groups A, B, and AB compared with blood group O) [49]. Both findings have been confirmed by a subsequent GWAS on severe malaria in West Africa (Table 2), revealing that the major interaction point of interest from the host perspective is the red blood cell. Indeed, it is this interaction between *P. falciparum* and the red blood cell that gives rise to all the clinical symptoms of malaria, and normally functioning erythrocyte physiology has been shown to be crucial to parasite survival [50,51]. However, what is surprising is that genetic variation in all immune-related genes (such as the HLA, tumor necrosis factor (TNF) and lymphotoxin alpha (LTA) families), the very genes that have been thought to have critical roles in the immune response and clearance of *P. falciparum* infections [52,53] do not consistently show association with susceptibility in large collections of malaria patients.



The paucity of findings from the viewpoint of the host beyond the red blood cell suggests that for malaria, genetic variations in *P. falciparum* itself might account for substantially more disease variance. The genomic approach adopted from the perspective of *P. falciparum* has been two-pronged: using genome-wide association and detection of natural selection to identify molecules crucial at the host-pathogen interface. It has been observed that the strong selective pressure exerted by *P. falciparum* on the erythrocyte has led to increased incidence of otherwise rare hematological disorders (such as sickle-cell disease, hemoglobin C, and possibly thalassemia) [54,55]. A better understanding of the selective pressure that malaria has exerted on the immune

system would thus yield considerable insight into the process of malaria pathogenesis and severity. Indeed, surveys of the *P. falciparum* genome using sequencing and genotyping have revealed it to be highly variable. Genes encoding *Plasmodium* proteins interacting with the host immune system, such as *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), are often under balancing selection resulting from pressures from the host immune system on one hand and the need to maintain diversity on the other [8]. The sequencing of *P. falciparum* virulence genes that interact with the red blood cell (*var*/PfEMP1) and transporters (such as *pfcr* for chloroquine, *pfsurfin* for dihydroartemisinin and *dfhr* for cycloguanil) responsible for resistance to anti-malarial

drugs revealed greater than average genomic diversity and the presence of positive selection signatures in these *Plasmodium* genes, indicating the continued evolution of the parasite in response to survival pressures from both the human immune response and to medical intervention. Indeed, a GWAS of *P. falciparum* genetic variation using an array of over 17,500 markers obtained by parasite genome sequencing revealed very strong associations between genetic variations in many drug transporter genes, including the *P. falciparum* chloroquine transporter gene (PfCRT), *P. falciparum* dihydrofolate reductase (PfDHFR) and resistance to anti-malarial chemotherapy [9]. Strong evidence of association with drug resistance and positive selection was also observed at *pfSURFIN*, a gene that had until then not been implicated in drug resistance. Curiously, PfSURFIN was reported to be co-transported with PfEMP1 to the infected surface of the red blood cell and is thought to be part of a protein complex involved in binding or transport of chemical compounds. The extension of this *P. falciparum* GWAS to severe malaria susceptibility in humans (as opposed to uncomplicated infection) will be extremely informative [56].

### Future challenges

The outcome of a specific episode of infection is strongly suspected to depend in part on specific interactions between host and pathogen genotypes [57]. The cumulative effect size from the contribution from the host and pathogen genomes is likely to be larger than each genome alone, and clinical outcomes of infection are unlikely to be explained by a reductionist approach that studies disparate individual components [58]. Even in less complex organisms than humans, host resistance to pathogens has been shown to vary dramatically across different combinations of host and pathogen genotypes [59], with an even more complicated picture if the host is susceptible to infection by multiple different strains of the same pathogen species. To this end, GWASs and pathogen sequencing have revealed fresh and often unexpected insights for host-pathogen interactions in revealing hitherto unsuspected molecules (Table 2). In cases of both dengue and meningococcal disease, GWASs revealed surprising interacting host molecules (the NK cell pathway for dengue and complement inhibition by factor H for meningococcal disease) that are important in disease pathogenesis. In addition, the GWAS on meningococcal disease revealed that pathogen genetic variation (in terms of different infecting strains) also contributes substantially to overall inter-individual susceptibility to disease - and this may also be case with dengue and other diseases.

Although the GWAS effort from the human perspective has not produced the novel insights for malaria seen for

other infectious diseases, it is clear that this is more than compensated for by approaches from the parasite perspective, which have already begun to bear fruit. Var genes identified from analysis of the sequenced *Plasmodium* genome are already turning up potential antigenic sites of interaction with the human genome, with the potential for being vaccine targets [8], and combined selection and association analysis have revealed PfCRT, PfSURFIN, PfMDR and *P. falciparum* apical membrane antigen 1 (PfAMA-1) as crucial molecules interacting with both anti-malarial medications and the host immune system, respectively, as well as *PFE1445c*, which encodes a *Plasmodium* conserved protein [9]. So, in this case, more detailed work stemming from analysis of *P. falciparum* resequencing data could reveal crucial interaction points between host and parasite.

The new challenges in defining host-pathogen interactions are starting to become computational, because to comprehensively analyze and integrate whole genome data for both humans and pathogens, complex bioinformatics toolsets are required, as the order of human-pathogen genotype-phenotype combinations increases exponentially. The search is further complicated by the well known observations that pathogen genomes are not as conserved as the human genome; for example, the per-nucleotide mutation rate is  $5 \times 10^{-4}$  per generation [8] for the *P. falciparum* genome, compared with about  $2.5 \times 10^{-8}$  for humans [60].

To aid in these challenges, high-throughput analysis pipelines can now be used for alignment and mapping of pathogen genomes generated by next-generation sequencing experiments, thus allowing a comprehensive catalog documenting all pathogen genetic polymorphisms (not unlike the human HapMap project). This effort will enable in-depth characterization of pathogen phylogeny and measurement of correlations between pathogen genetic mutations and virulence, together with key human phenotypes of the infection (such as cytokine profile and disease severity). In most cases, the integration of genomic data obtained from GWASs and resequencing efforts from both the host and pathogen should reveal a complete catalog of the inter-individual susceptibility to infection.

### Author details

<sup>1</sup>Infectious Diseases, Genome Institute of Singapore, 60 Biopolis Street, #02-01 Genome Building, Singapore 138672. <sup>2</sup>NUS-GIS Centre for Molecular Epidemiology, 60 Biopolis Street, #02-01 Genome Building, Singapore 138672. <sup>3</sup>Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, NUHS Tower Block, Level 12, 1E Kent Ridge Road, Singapore 119228. <sup>4</sup>Department of Epidemiology and Public Health, Yong Loo Lin School of Medicine, National University of Singapore, 16 Medical Drive (MD 3), Singapore 117597.

## References

- Newport MJ, Huxley CM, Huston S, Hawrylowicz CM, Oostra BA, Williamson R, Levin M: **A mutation in the interferon-gamma-receptor gene and susceptibility to mycobacterial infection.** *N Engl J Med* 1996, **335**:1941-1949.
- Coers J, Starnbach MN, Howard JC: **Modeling infectious disease in mice: co-adaptation and the role of host-specific IFN $\gamma$  responses.** *PLoS Pathog* 2009, **5**:e1000333.
- Puck JM, Deschênes SM, Porter JC, Dutra AS, Brown CJ, Willard HF, Henthorn PS: **The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X-linked severe combined immunodeficiency, SCID1.** *Hum Mol Genet* 1993, **2**:1099-1104.
- Vetrie D, Vorechovský I, Sideras P, Holland J, Davies A, Flinter F, Hammarström L, Kinnon C, Levinsky R, Bobrow M: **The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases.** *Nature* 1993, **361**:226-233.
- Zhang FR, Huang W, Chen SM, Sun LD, Liu H, Li Y, Cui Y, Yan XX, Yang HT, Yang RD, Chu TS, Zhang C, Zhang L, Han JW, Yu GQ, Quan C, Yu YX, Zhang Z, Shi BQ, Zhang LH, Cheng H, Wang CY, Lin Y, Zheng HF, Fu XA, Zuo XB, Wang Q, Long H, Sun YP, Cheng YL, et al.: **Genomewide association study of leprosy.** *N Engl J Med* 2009, **361**:2609-2618.
- Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB: **A whole-genome association study of major determinants for host control of HIV-1.** *Science* 2007, **317**:944-947.
- Davila S, Wright VJ, Khor CC, Sim KS, Binder A, Breunis WB, Inwald D, Nadel S, Betts H, Carrol ED, de Groot R, Hermans PW, Hazelzet J, Emonts M, Lim CC, Kuijpers TW, Martinon-Torres F, Salas A, Zenz W, Levin M, Hibberd ML: **Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease.** *Nat Genet* 2010, **42**:772-776.
- Mu J, Awadalla P, Duan J, McGee KM, Keebler J, Seydel K, McVean GA, Su XZ: **Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome.** *Nat Genet* 2007, **39**:126-130.
- Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, Chotivanich K, Wilairatana P, Krudsood S, White NJ, Udomsangpetch R, Cui L, Ho M, Ou F, Li H, Song J, Li G, Wang X, Seila S, Sokunthea S, Socheat D, Sturdevant DE, Porcella SF, Fairhurst RM, Welles TE, Awadalla P, Su XZ: ***Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs.** *Nat Genet* 2010, **42**:268-271.
- Van Tyne D, Park DJ, Schaffner SF, Neafsey DE, Angelino E, Cortese JF, Barnes KG, Rosen DM, Lukens AK, Daniels RF, Milner DA, Johnson CA, Shlyakhter I, Grossman SR, Becker JS, Yamins D, Karlsson EK, Ndiaye D, Sarr O, Mboup S, Happi C, Furlotte NA, Eskin E, Kang HM, Hartl DL, Birren BW, Wiegand RC, Lander ES, Wirth DF, Volkman SK, Sabeti PC: **Identification and functional validation of the novel antimalarial resistance locus PF10\_0355 in *Plasmodium falciparum*.** *PLoS Genet* 2011, **7**:e1001383.
- Khor CC, Hibberd ML: **Shared pathways to infectious disease susceptibility?** *Genome Med* 2010, **2**:52.
- Barros MC, Galasso TG, Chaib AJ, Degallier N, Nagata T, Ribeiro BM: **Yellow fever virus envelope protein expressed in insect cells is capable of syncytium formation in lepidopteran cells and could be used for immunodetection of YFV in human sera.** *Virology* 2011, **8**:261.
- Parkinson T, Pryde DC: **Small molecule drug discovery for Dengue and West Nile viruses: applying experience from hepatitis C virus.** *Future Med Chem* 2010, **2**:1181-1203.
- Gubler DJ: **Dengue and dengue hemorrhagic fever.** *Clin Microbiol Rev* 1998, **11**:480-496.
- Leitmeyer KC, Vaughn DW, Watts DM, Salas R, Villalobos I, de Chacon C, Ramos R, Rico-Hesse: **Dengue virus structural differences that correlate with pathogenesis.** *J Virol* 1999, **73**:4738-4747.
- Fox A, Le NM, Simmons CP, Wolbers M, Wertheim HF, Pham TK, Tran TH, Trinh TM, Nguyen TL, Nguyen VT, Nguyen DH, Farrar J, Horby P, Taylor WR, Nguyen VK: **Immunological and viral determinants of dengue severity in hospitalized adults in Ha Noi, Viet Nam.** *PLoS Negl Trop Dis* 2011, **5**:e967.
- Vaughn DW, Green S, Kalayanarooj S, Innis BL, Nimmannitya S, Suntayakorn S, Endy TP, Raengsakulrach B, Rothman AL, Ennis FA, Nisalak A: **Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity.** *J Infect Dis* 2000, **181**:2-9.
- Rico-Hesse R, Harrison LM, Salas RA, Tovar D, Nisalak A, Ramos C, Boshell J, de Mesa MT, Nogueira RM, da Rosa AT: **Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas.** *Virology* 1997, **230**:244-251.
- Schreiber MJ, Holmes EC, Ong SH, Soh HS, Liu W, Tanner L, Aw PP, Tan HC, Ng LC, Leo YS, Low JG, Ong A, Ooi EE, Vasudevan SG, Hibberd ML: **Genomic epidemiology of a dengue virus epidemic in urban Singapore.** *J Virol* 2009, **83**:4163-4173.
- Khor C-C, Bich Chau TN, Pang J, Davila S, Long HT, RTH, Dunstan SJ, Wills B, Farrar J, Tram TV, Gan TT, Nguyen Binh NT, Tri LT, Lien LB, Tuan NM, Hong Tham NT, Lanh MN, Nguyet NM, Hieu NT, Chau NVV, Thuy TT, Tan DEK, Sakuntabhai A, Teo Y-Y, Hibberd ML and Simmons CP: **Genome-wide association study identifies susceptibility loci for dengue shock syndrome at MICB and PLCE1.** *Nature Genet* 2011, In press
- Steinle A, Li P, Morris DL, Groh V, Lanier LL, Strong RK, Spies T: **Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family.** *Immunogenetics* 2011, **53**:279-287.
- González S, López-Soto A, Suarez-Alvarez B, López-Vázquez A, López-Larrea C: **NKG2D ligands: key targets of the immune response.** *Trends Immunol* 2008, **29**:397-403.
- Andoniou CE, Andrews DM, Degli-Esposti MA: **Natural killer cells in viral infection: more than just killers.** *Immunol Rev* 2006, **214**:239-250.
- Kumar V, Kato N, Urabe Y, Takahashi A, Muroyama R, Hosono N, Otsuka M, Tateishi R, Omata M, Nakagawa H, Koike K, Kamatani N, Kubo M, Nakamura Y, Matsuda K: **Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma.** *Nat Genet* 2011, **43**:455-458.
- Hinkes B, Wiggins RC, Gbadegesin R, Vlangos CN, Seelow D, Nürnberg G, Garg P, Verma R, Chaib H, Hoskins BE, Ashraf S, Becker C, Hennies HC, Goyal M, Wharram BL, Schachter AD, Mudumana S, Drummond I, Kerjaschki D, Waldherr R, Dietrich A, Ozaltin F, Bakkaloglu A, Cleper R, Basel-Vanagaite L, Pohl M, Griebel M, Tsygin AN, Soylyu A, Müller D, et al.: **Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible.** *Nat Genet* 2006, **38**:1397-1405.
- Vu TT, Holmes EC, Duong V, Nguyen TQ, Tran TH, Quail M, Churcher C, Parkhill J, Cardosa J, Farrar J, Wills B, Lennon NJ, Birren BW, Buchy P, Henn MR, Simmons CP: **Emergence of the Asian 1 genotype of dengue virus serotype 2 in Viet Nam: in vivo fitness advantage and lineage replacement in South-East Asia.** *PLoS Negl Trop Dis* 2010, **4**:e757.
- Raghwan J, Rambaut A, Holmes EC, Hang VT, Hien TT, Farrar J, Wills B, Lennon NJ, Birren BW, Henn MR, Simmons CP: **Endemic dengue associated with the co-circulation of multiple viral lineages and localized density-dependent transmission.** *PLoS Pathog* 2011, **7**:e1002064.
- Christenbury JG, Aw PP, Ong SH, Schreiber MJ, Chow A, Gubler DJ, Vasudevan SG, Ooi EE, Hibberd ML: **A method for full genome sequencing of all four serotypes of the dengue virus.** *J Virol Methods* 2010, **169**:202-206.
- Hoang LT, Lynn DJ, Henn M, Birren BW, Lennon NJ, Le PT, Duong KT, Nguyen TT, Mai LN, Farrar JJ, Hibberd ML, Simmons CP: **The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults.** *J Virol* 2010, **84**:12982-12994.
- Pace D: **Quadrivalent meningococcal ACYW-135 glycoconjugate vaccine for broader protection from infancy.** *Expert Rev Vaccines* 2009, **8**:529-542.
- Darton T, Guiver M, Naylor S, Jack DL, Kaczmarek EB, Borrow R, Read RC: **Severity of meningococcal disease associated with genomic bacterial load.** *Clin Infect Dis* 2009, **48**:587-594.
- Spanjaard L, Bol P, de Marie S, Zanen HC: **Association of meningococcal serotypes with the course of disease: serotypes 2a and 2b in the Netherlands, 1959-1981.** *J Infect Dis* 1987, **155**:277-282.
- Trotter CL, Fox AJ, Ramsay ME, Sadler F, Gray SJ, Mallard R, Kaczmarek EB: **Fatal outcome from meningococcal disease – an association with meningococcal phenotype but not with reduced susceptibility to benzylpenicillin.** *J Med Microbiol* 2002, **51**:855-860.
- Hibberd ML, Sumiya M, Summerfield JA, Booy R, Levin M: **Association of variants of the gene for mannose-binding lectin with susceptibility to meningococcal disease.** Meningococcal Research Group. *Lancet* 1999, **353**:1049-1053.
- Lorenz E, Mira JP, Frees KL, Schwartz DA: **Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock.** *Arch Intern Med* 2002, **162**:1028-1032.
- Fijen CA, Kuijper EJ, Hannema AJ, Sjöholm AG, van Putten JP: **Complement deficiencies in patients over ten years old with meningococcal disease**

- due to uncommon serogroups. *Lancet* 1989, **2**:585-588.
37. Wright V, Hibberd M, Levin M: **Genetic polymorphisms in host response to meningococcal infection: the role of susceptibility and severity genes.** *Vaccine* 2009, **27 Suppl 2**:B90-B102.
38. Fritsche LG, Lauer N, Hartmann A, Stippa S, Keilhauer CN, Oppermann M, Pandey MK, Köhl J, Zipfel PF, Weber BH, Skerka C: **An imbalance of human complement regulatory proteins CFHR1, CFHR3 and factor H influences risk for age-related macular degeneration (AMD).** *Hum Mol Genet* 2010, **19**:4694-4704.
39. Fijen CA, Kuijper EJ, te Bulte MT, Daha MR, Dankert J: **Assessment of complement deficiency in patients with meningococcal disease in The Netherlands.** *Clin Infect Dis* 1999, **28**:98-105.
40. Schneider MC, Prosser BE, Caesar JJ, Kugelberg E, Li S, Zhang Q, Quoraishi S, Lovett JE, Deane JE, Sim RB, Roversi P, Johnson S, Tang CM, Lea SM: ***Neisseria meningitidis* recruits factor H using protein mimicry of host carbohydrates.** *Nature* 2009, **458**:890-893.
41. Mothibeli KM, du Plessis M, von Gottberg A, Murphy E, Hoiseth SK, Zlotnick G, Klugman KP: **Distribution of factor H binding protein beyond serogroup B: variation among five serogroups of invasive *Neisseria meningitidis* in South Africa.** *Vaccine* 2011, **29**:2187-2192.
42. Perlmann P, Troye-Blomberg M: **Malaria blood-stage infection and its control by the immune system.** *Folia Biol (Praha)* 2000, **46**:210-218.
43. Sarkar S, Biswas NK, Dey B, Mukhopadhyay D, Majumder PP: **A large, systematic molecular-genetic study of G6PD in Indian populations identifies a new non-synonymous variant and supports recent positive selection.** *Infect Genet Evol* 2010, **10**:1228-1236.
44. Artavanis-Tsakonas K, Riley EM: **Innate immune response to malaria: rapid induction of IFN-gamma from human NK cells by live *Plasmodium falciparum*-infected erythrocytes.** *J Immunol* 2002, **169**:2956-2963.
45. Ing R, Gros P, Stevenson MM: **Interleukin-15 enhances innate and adaptive immune responses to blood-stage malaria infection in mice.** *Infect Immun* 2005, **73**:3172-3177.
46. Porter DW, Thompson FM, Berthoud TK, Hutchings CL, Andrews L, Biswas S, Poulton I, Prieur E, Correa S, Rowland R, Lang T, Williams J, Gilbert SC, Sinden RE, Todryk S, Hill AV: **A human Phase I/IIa malaria challenge trial of a polyprotein malaria vaccine.** *Vaccine* 2011. doi:10.1016/j.vaccine.2011.03.083.
47. Schneider J, Gilbert SC, Blanchard TJ, Hanke T, Robson KJ, Hannan CM, Becker M, Sinden R, Smith GL, Hill AV: **Enhanced immunogenicity for CD8+ T cell induction and complete protective efficacy of malaria DNA vaccination by boosting with modified vaccinia virus Ankara.** *Nat Med* 1998, **4**:397-402.
48. Hill AV, Reyes-Sandoval A, O'Hara G, Ewer K, Lawrie A, Goodman A, Nicosia A, Folgori A, Colloca S, Cortese R, Gilbert SC, Draper SJ: **Prime-boost vectored malaria vaccines: progress and prospects.** *Hum Vaccin* 2010, **6**:78-83.
49. Fry AE, Griffiths MJ, Auburn S, Diakite M, Forton JT, Green A, Richardson A, Wilson J, Jallow M, Sisay-Joof F, Pinder M, Peshu N, Williams TN, Marsh K, Molyneux ME, Taylor TE, Rockett KA, Kwiatkowski DP: **Common variation in the ABO glycosyltransferase is associated with susceptibility to severe *Plasmodium falciparum* malaria.** *Hum Mol Genet* 2008, **17**:567-576.
50. Cooke BM, Mohandas N, Coppel RL: **The malaria-infected red blood cell: structural and functional changes.** *Adv Parasitol* 2001, **50**:1-86.
51. Kirk K: **Membrane transport in the malaria-infected erythrocyte.** *Physiol Rev* 2001, **81**:495-537.
52. Good MF, Engwerda C: **Defying malaria: arming T cells to halt malaria.** *Nat Med* 2011, **17**:49-51.
53. Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, Brewster D, McMichael AJ, Greenwood BM: **Common west African HLA antigens are associated with protection from severe malaria.** *Nature* 1991, **352**:595-600.
54. Williams TN, Mwangi TW, Wambua S, Peto TE, Weatherall DJ, Gupta S, Recker M, Penman BS, Uyoga S, Macharia A, Mwacharo JK, Snow RW, Marsh K: **Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait.** *Nat Genet* 2005, **37**:1253-1257.
55. Modiano D, Luoni G, Sirima BS, Simporé J, Verra F, Konaté A, Rastrelli E, Olivieri A, Calissano C, Paganotti GM, D'Urbanò L, Sanou I, Sawadogo A, Modiano G, Coluzzi M: **Haemoglobin C protects against clinical *Plasmodium falciparum* malaria.** *Nature* 2001, **414**:305-308.
56. Winter G, Kawai S, Haeggström M, Kaneko O, von Euler A, Kawazu S, Palm D, Fernandez V, Wahlgren M: **SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes.** *J Exp Med* 2005, **201**:1853-1863.
57. Lambrechts L, Fellous S, Koella JC: **Coevolutionary interactions between host and parasite genotypes.** *Trends Parasitol* 2006, **22**:1216.
58. Zak DE, Aderem A: **Systems biology of innate immunity.** *Immunol Rev* 2009, **227**:264-282.
59. Wilfert L, Schmid-Hempel P: **The genetic architecture of susceptibility to parasites.** *BMC Evol Biol* 2008, **8**:187.
60. Nachman MW, Crowell SL: **Estimate of the mutation rate per nucleotide in humans.** *Genetics* 2000, **156**:297-304.
61. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C: **Global, regional, and national causes of child mortality in 2008: a systematic analysis.** *Lancet* 2010, **375**:1969-1987.
62. Jallow M, Teo YY, Small KS, Rockett KA, Deloukas P, Clark TG, Kivinen K, Bojang KA, Conway DJ, Pinder M, Sirugo G, Sisay-Joof F, Usen S, Auburn S, Bumpstead SJ, Campino S, Coffey A, Dunham A, Fry AE, Green A, Gwilliam R, Hunt SE, Inouye M, Jeffreys AE, Mendy A, Palotie A, Potter S, Ragoussis J, Rogers J, Rowlands K, et al.: **Genome-wide and fine-resolution association analysis of malaria in West Africa.** *Nat Genet* 2009, **41**:657-665.
63. Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, Osei I, Gyapong J, Sirugo G, Sisay-Joof F, Enimil A, Chinbuah MA, Floyd S, Warndorff DK, Sichali L, Malema S, Crampin AC, Ngwira B, Teo YY, Small K, Rockett K, Kwiatkowski D, Fine PE, Hill PC, Newport M, Lienhardt C, Adegboola RA, Corrah T, Ziegler A, Morris AP, Meyer CG, Horstmann RD, Hill AV: **Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2.** *Nat Genet* 2010, **42**:739-741.
64. Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, Kubo M, Tsunoda T, Kamatani N, Kumada H, Puseenam A, Sura T, Daigo Y, Chayama K, Chantratita W, Nakamura Y, Matsuda K: **A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians.** *Nat Genet* 2009, **41**:591-595.
65. Mbarek H, Ochi H, Urabe Y, Kumar V, Kubo M, Hosono N, Takahashi A, Kamatani Y, Miki D, Abe H, Tsunoda T, Kamatani N, Chayama K, Nakamura Y, Matsuda K: **A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population.** *Hum Mol Genet* 2011, **20**:3884-92.
66. Tan LK, Carlone GM, Borrow R: **Advances in the development of vaccines against *Neisseria meningitidis*.** *N Engl J Med* 2010, **362**:1511-1520.

doi:10.1186/gb-2011-12-10-229

Cite this article as: Khor C-C, Hibberd ML: **Revealing the molecular signatures of host-pathogen interactions.** *Genome Biology* 2011, **12**:229.