Host Biomarkers and Paediatric Infectious Diseases: From Molecular Profiles to Clinical Application

H.K. Brand, P.W.M. Hermans, and R. de Groot

1 Introduction

Infectious diseases are an important cause of death among children under the age of 5 (Stein et al., 2004). Most of these deaths are caused by preventable or curable infections. Limited access to medical care, antibiotics, and vaccinations remains a major problem in developing countries. But infectious diseases also continue to be an important public health issue in developed countries. With the help of modern technologies, some infections have been effectively controlled; however, new diseases such as SARS and West Nile virus infections are constantly emerging. In addition, other diseases such as malaria, tuberculosis, and bacterial pneumonia are increasingly resistant to antimicrobial treatment.

The physician who manages pediatric patients with infectious diseases is confronted with several related challenges. First, one should establish a specific diagnosis, preferably early in the course of disease. Despite improvements in culture and non-culture diagnostics, in many cases, the causative micro-organism remains unknown. Consequent delays in initiation of appropriate treatment can contribute to the emergence of antibiotic resistance.

A second challenge is to identify those patients most likely to develop severe disease. To date, physicians have little information on prognosis and likely disease outcome in the individual patient. It would be extremely useful to be able to identify patients at risk of more severe disease (e.g., secondary bacterial infection during viral respiratory tract infection), as such prediction could inform management decisions.

The third associated challenge is to select the most appropriate treatment strategy for an individual patient. While some patients require intensive support, others will recover without additional medication or supportive care. To date, few tools are available to monitor the course of disease after initiation of medical treatment.

R. de Groot (✉)
Department of Pediatrics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
e-mail: r.degroot@cukz.umcn.nl

A. Finn et al. (eds.), *Hot Topics in Infection and Immunity in Children VI*, Advances in Experimental Medicine and Biology 659, DOI 10.1007/978-1-4419-0981-7_2,

Biomarkers have been used for years to help clinical decision-making. C-reactive protein (CRP) is probably the best known marker used to monitor infection. Although useful, it does not reliably distinguish viral from bacterial infections. More recently-developed markers such as procalcitonin seem promising, with detectable rises early in the course of infection and high negative predictive value as seen in children with fever of unknown origin (van Rossum et al., 2004; Galetto-Lacour et al., 2003; Herd, 2007). However, this marker has also insufficient power to discriminate between viral and bacterial infections. Additionally, these conventional biomarkers for infectious diseases do not provide microorganism-specific prognostic information.

The completion of the Human Genome Project and the introduction of powerful DNA microarray chips and proteomic technologies in the mid-1990s have created the opportunity to identify genes and proteins that may serve as biomarkers in infectious diseases. The identification of biomarkers may enable the development of exciting potential clinical applications in which genes and proteins that are differentially expressed in healthy and infected individuals can be investigated (The International HapMap Project, 2003). These approaches may provide detailed insight into the pathogenesis of disease, host pathogen interactions, and disease-specific expression patterns. In addition, diagnostic and prognostic biomarkers and markers that monitor disease or response to therapy may be developed using these technologies.

This chapter describes recent advances in genomics and proteomics in the field of biomarkers for infectious diseases and summarizes current clinical applications and future perspectives.

2 Molecular Profiling: Current Techniques

2.1 Genomics

Single nucleotide polymorphisms (SNPs), single base pair changes at specific spots in the genome, are the most common type of genetic variation. The human genome carries over 10 million nucleotides that vary in at least 1% of the population. Currently, approximately 6 million nucleotides have been validated and this number is still growing (Bryant et al., 2004; Walker and Siminovitch, 2007). Although SNPs are the changes most frequently explored using high throughput technologies, other genetic variations are also common in the human genome and may influence the individual's susceptibility to disease. These include variations in gene copy number, repeating sequence motifs, insertions, and deletions (Crawford et al., 2005).

The completion of the human genome map in combination with the development of microarray-based comparative genomic hybridization and genome-wide SNP platforms have permitted the screening of the entire genome to identify genetic loci linked to certain diseases, susceptibility to disease or response to therapy (Feuk et al., 2006). These genome wide association studies allow identification of genetic risk factors for a wide variety of common and more complex diseases by measuring hundreds of thousands of genetic variants simultaneously.

Using SNP platforms, MalariaGen (2009), a genomic epidemiological network, and the Wellcome Trust Case-Control Consortium (WTCCC, 2006) have performed a large scale study on disease susceptibility. Both consortia have analyzed up to 500,000 SNPs in thousands of African individuals diagnosed with malaria or tuberculosis between 2006 and 2008. In addition, the WTCCC will include approximately 2,000 cases and 3,000 controls for 8 other diseases, which makes it one of the biggest projects aimed at the identification of genetic variations that may predispose a patient to disease (Genome-wide Association Study, 2007). These genome-wide analyses may contribute to the identification of individuals at risk and produce more effective prevention strategies and individual treatment strategies.

2.2 Transcriptomics

Where genomics provides information on genetic susceptibility to certain infectious diseases, transcriptomics provides information on the activity of genes at a certain moment under certain conditions. It is the study of the complete set of RNA transcripts produced by the genome. During all biological processes, part of the genome is specifically transcribed into messenger RNA (mRNA, transcriptome) and translated to proteins (proteome). The transcriptome can be analyzed using gene expression microarrays. These chips contain either the whole genome or a subset of specific genes. mRNA is extracted from experimental samples, reverse transcribed and labeled with fluorescent dyes. The extracted labeled cDNA is then hybridized with the microarray and the fluorescence of the array is determined using an array scanner. Following image analysis, the data are subjected to bioinformatics processes to identify statistically significant changes in gene expression between different samples. The technique can be used to characterize gene expression in both pathogen and host, providing detailed insight into host-pathogen interactions during infection (Liu et al., 2006). To study them on a molecular level, Kawada et al. generated gene-expression profiles in peripheral blood mononuclear cells (PBMCs) isolated from children with influenza virus infection. Many genes associated with the immune response such as interferon regulated genes appeared to be strongly upregulated during influenza infection. In addition, they compared gene expression profiles of influenza-infected children with and without convulsions. They found that transcription levels of pro-inflammatory cytokine genes in patients with a febrile convulsion were not significantly different from those in patients without febrile convulsion. This kind of approach may help to clarify the pathogenesis of influenza and its neurological complications (Kawada et al., 2006).

2.3 Proteomics

While gene expression profiles may not completely correlate with intracellular protein content, proteomics can provide insight into the structure and dynamics of the end product, proteins. Proteomics is the study of the proteome, the complete set of proteins, their modifications, interactions, and localization. Proteomic technologies

enable detailed analyses of protein expression and evaluation of post-translational modification and protein stability and turnover that cannot be assessed by genomic and transcriptomic profiling alone. For many years, two-dimensional gel electrophoresis has been the standard technology to isolate specific proteins and allow protein identification by subsequent mass spectrometry. During the past decade, mass spectrometry has improved and now enables the analysis of protein expression, structure, and function without the need for labor intensive and time consuming electrophoresis (Graves and Haystead, 2002; Patterson and Aebersold, 2003). In addition, several mass spectrometry-based approaches have been developed that allow the relative or absolute quantification of proteins.

In an attempt to identify biomarkers and to develop screening tools with high sensitivity and specificity, proteomics technology has been applied to analyze biofluids such as serum, saliva, or urine. A commonly used technique is a surface enhanced laser desorption/ionization time-of-flight (SELDI-TOF), which is a proteomics technique that allows the identification of large numbers of proteins in a short period of time. This technique provides a specific mass spectral profile from each analyzed protein sample. By comparing profiles from affected individuals to those derived from healthy controls, a specific protein signature is obtained (Coombes et al., 2005; Hodgetts et al., 2007). Agranoff and colleagues have used it to characterise distinct profiles for several microbial infections and to investigate serum responsiveness to *M. tuberculosis* identifying serum biomarkers from patients with advanced tuberculosis. SELDI-TOF has the potential to identify tuberculosis at an early stage, assisting early diagnosis and therapy, which is important for favorable outcome (Agranoff et al., 2005).

3 Clinical Applications in Pediatric Infectious Diseases

The recent advances in proteomic and genomic technologies have allowed the identification of genes and proteins that may serve as biomarkers for the diagnosis and monitoring of infectious diseases. Application of knowledge from these technologies into clinical practice is still at an early stage. In the following sections, we will first discuss current literature on the contribution of the aforementioned technologies for the determination of disease pathogenesis, for the susceptibility to infection, and for diagnostics. In the next section, we will discuss clinical applications of proteomic and transcriptomic technologies, and, in the last section, we will focus on their future perspectives.

3.1 What Has Been Studied up to now?

With increasing use of genomic and proteomic technologies, more insight has been obtained into host-pathogen interactions and pathogenesis of infectious diseases. The response of the host to pathogens is reflected in changes in gene expression

and can be measured by microarray based gene expression technologies. Likewise, transcriptional profiling studies have proven to be a powerful approach for analyzing and understanding host–pathogen interactions. Based upon the host response to various pathogens, Jenner et al. have identified common and more specific gene expression patterns. They collected and systematically compared transcriptional profiling datasets from 32 published microarray-based in vitro studies which collectively examined 77 different host-pathogen interactions. In response to this wide variety of pathogens, they identified a cluster of 511 genes that share a common response upon infection with a pathogen. According to the localization of the cell in which they function, these genes have been clustered into different functional groups in order to provide an overview of cellular physiology involved in the common host response. Moreover, analyzing different transcriptional profiling studies also revealed pathogen specific gene expression patterns. Several host genes of the aforementioned common host response were found to be downregulated in the presence of pathogens or specific pathogen proteins. These transcriptional profiling studies has provided insight into how micro-organisms alter host gene expression patterns to subvert the immune responses. For example, transcriptional profiling has identified that viruses such as herpes simplex virus (HSV)-1, human cytomegalovirus, and human papillomavirus-31 are partially or completely able to inhibit the induction of Interferon stimulated genes, which have a central role in the defense against viruses (Jenner and Young, 2005).

Streptococcus pneumoniae and influenza virus are the most common causes of pneumonia. Consequently, they contribute to substantial morbidity and mortality worldwide. It has been known for years that influenza infection predisposes to secondary bacterial infection, mainly caused by *S. pneumoniae* and *S. aureus*. The catastrophic influenza A pandemic in 1918 in which approximately 40–50 million persons were killed, may have involved synergy between influenza and pneumococcal infections. Gene expression profiling is a powerful tool to explore the molecular mechanism of synergy between pathogens. By host gene expression analyses of the lungs in mice, Rousseau et al. differentiated pneumococcal infections from influenza. Rosseau et al., have also identified common gene expression patterns in infectious disease as well as unique pathogen-specific gene expression signatures that may help clarify the mechanisms behind the synergy between influenza virus and *S. pneumoniae* (Rosseau et al., 2007). In response to influenza infection, Tong et al. performed gene expression analyses of middle ear epithelial cells (Tong et al., 2004). Tong et al. suggest that increased expression of inflammatory mediator genes such as IP-10 and CXCL11 could lead to a shift in *S. pneumoniae* adherence by activation of host epithelial and endothelial cells, providing a favorable environment in the middle ear cavity for a secondary bacterial infection with *S. pneumoniae*.

The response upon exposure to pathogens varies widely between individuals. Some people are more susceptible to a certain infection than others. This differential susceptibility is partly caused by genetic variations between individuals that may predispose either to development of disease or to a more severe course. Although the large variation in clinical responses among individuals, also depends upon environmental and microbial factors. The classic example of host genetic susceptibility

is the resistance of heterozygous hemoglobin S individuals to malaria infection (caused by *Plasmodium falciparum*) (Allison, 1954). Other approaches used to elucidate genetic and environmental effects on infections include studies in identical and non-identical twins (Hill, 1998; Jepson et al., 2001; Jepson et al., 1995) and comparisons of risk in adopted children and their biological and adoptive parents. One such study suggested that adopted children with a biological parent who died early of an infectious disease had a higher risk of mortality from similar infections while the death of an adoptive parent due to infection had no influence on the risk of disease in the children (Sorensen et al., 1988).

Recent advances in microarray technologies have enabled genome wide searches for genes influencing susceptibility to infectious diseases. Analysis of genetic susceptibility aims to link the genetic code (genotype) to the risk of a certain disease state (phenotype). Given the large number of human genes, many with unknown function, genome wide studies have the advantage that previously unconsidered genes can be identified and provide more sensitivity for the detection of subtle genetic effects and gene recruitment in affected individuals. However, many reported genetic associations have not been replicated in subsequent studies, and, for secure results, large numbers of affected and unaffected individuals are required. Furthermore, because of the complexity of data analysis, microarray technologies are time and labor intensive (Xavier and Rioux, 2008; Cooke and Hill, 2001; Hirschhorn and Daly, 2005). Nevertheless, several large scale population based studies have been performed and support the role of genome wide searches in the identification of genes influencing disease susceptibility (WTCCC, 2006; O'Brien and Nelson, 2004; An et al., 2007; Hill, 2006). For example, a genetic association study performed by O'Brien et al. has led to the identification of various genetic factors that affect HIV-1/AIDS. Genetic association analysis of several large cohorts of HIV infected individuals resulted in the identification of 14 AIDS restriction genes, which are human genes with polymorphic variants that influence the outcome of HIV-1 exposure or infection. The study of O'Brien et al., illustrates the discovery of previously unknown genes involved in susceptibility to infection using SNP haplotype-based association studies in clinically well-described epidemiological cohorts (O'Brien and Nelson, 2004).

Another potential application of microarray technologies is diagnosis of infection both by direct and indirect identification of pathogens. For example, microarrays composed of DNA sequences of various pathogens allow the identification of many organisms in a single test. Wang et al. developed an array composed of all fully sequenced reference viral genomes that allows the detection of approximately 1000 viruses (Wang et al., 2002). Moreover, by sequencing hybridized material of unknown pathogens, this array permits identification of new viruses, and, in 2003, it proved successful in the global effort to identify the novel corona virus associated with severe acute respiratory syndrome (SARS) (Wang et al., 2002; Ksiazek et al., 2003).

In contrast to direct identification, infections can also be characterised indirectly through specific host responses. An advantage of such pathogen-specific molecular signatures in the host is that they may be present at various stages of infection,

even when the pathogen is not detectable using standard or direct diagnostic tests. Ramilo et al. (2007), used gene expression analysis to diagnose different pathogen fingerprints in pediatric patients with respiratory infections caused by influenza A virus and Gram-negative (*E.coli*) or Gram-positive (*Staphylococcus aureus* and *Streptococcus pneumoniae*) bacteria. Classifier genes, which discriminate influenza A from bacterial infections (*S. pneumoniae and S. aureus*) and *E. coli* from *S. aureus* infections, were identified and validated. Another example of "–omic" technology use in diagnosing infection is the development of a protein based signature for diagnosing trypanosomiasis or sleeping sickness, which affects half a million people yearly in sub-Saharan Africa. Trypanosomiasis, if left untreated, is a debilitating disease with a lethaloutcome; it was successfully controlled in the past, but, since the 1970s, has re-emerged as an epidemic of immense proportions causing huge, yet widely underestimated morbidity and mortality of up to 50,000 cases every year (Stich et al., 2002). Establishing the diagnosis remains complicated, as current diagnostic tests lack the sensitivity to detect low parasite loads in peripheral blood. Papadopoulos et al. analyzed serum samples from patients and controls using SELDI –TOF mass spectrometry and identified distinct serum proteomic signatures in both groups (Papadopoulos et al., 2004). After depleting serum samples of antibody components, the authors identified two prominent protein peaks at 23/24 and 47 kDa in patients. These proteomic signatures may provide the basis for new diagnostic tests and alternative methods to monitor the host response to treatment. Moreover, additional characterization of these differentially expressed proteins may allow the development of simpler, cheaper antibody based tests (Papadopoulos et al., 2004; Agranoff et al., 2005).

3.2 Current and Potential Clinical Applications of "–omic" Technologies

More than 50% of all children admitted to the hospital have fever or other non-specific symptoms related to infection (Schaad, 1997). Although not necessarily suffering from bacterial infection, a significant proportion of these children will receive antibiotics. Although clinical history, physical examination and conventional diagnostic investigations (e.g., x-rays, blood tests) may point to an extent towards cause, pathogen identification remains difficult or even impossible. During episodes of acute fever, clinicians prefer to rely on cultures taken from the site of infection. However, such cultures often cannot be obtained at the right time or from the relevant site (e.g., middle ear or lungs) and results are not available for at least 24 h after sampling so that pathogens often remain undetected. Furthermore, contamination and colonization, particularly in upper airway samples, can obscure results. Gene expression profiles may identify bacterial pathogens and discriminate between bacterial infections, infections caused by other pathogens, and non-infectious causes of fever (like auto-inflammatory diseases). Using microarrays, organisms can be identified either directly or indirectly through their effects on host gene expression.

Tang et al. showed that gene expression profiling of neutrophils can distinguish sepsis from non-infectious inflammation (e.g., Systematic Inflammatory Response Syndrome, SIRS) in intensive care patients. They performed microarrays on a cohort of septic (N=71) and non-septic (N=23) patients and identified 50 classifier genes differentially expressed between the two groups, which are involved in inflammatory responses, immune regulation, and mitochondrial function. Broadly genes involved in the upregulation of immune responses were expressed less in patients with sepsis than in control patients, whereas genes involved in down regulation were expressed more, suggesting that sepsis may have an inhibitory effect on immune regulation. Pathway analyses support the finding that the immune regulation is inhibited during sepsis by showing that genes involved in the NF-B pathway were expressed less in patients with sepsis, whereas the inhibitory gene *NFKBIA* was expressed more. A prediction model for disease severity was developed from these data and validated in a second more heterogeneous patient group (Tang et al., 2007). A major advantage of gene expression profiling is that it may enable the development of less expensive diagnostic tools such as quantitative real time-polymerase chain reaction (RT-PCR) detection and quantification of specific DNA sequences in septic patients instead of entire gene expression profiles.

Children with auto-inflammatory diseases often present with non-specific systemic symptoms like rash and fever, which precede more specific symptoms like arthritis. Diagnosis of inflammatory diseases is often difficult due to the presentation with non-specific symptoms and the low incidence of these disease. Patients are often empirically treated for more likely causes of symptoms, including infections. Such delay in diagnosis and initiation of appropriate treatment is suboptimal for the child and may also result in misuse of antibiotics, contributing to emergence of antibiotic resistance. To differentiate patients with auto-inflammatory diseases (e.g., systemic onset juvenile idiopathic arthritis) from patients with acute viral and bacterial infections Allantez et al. analyzed leukocyte gene expression profiles of different blood leukocyte subpopulations that were obtained from these patients. Based on their results, a specific blood signature was developed that enabled differentiation between infection and other febrile inflammatory diseases (Allantaz et al., 2007).

Along with permitting characterization of infections when pathogens are not directly detectable, measurement of specific gene expression by the host can potentially permit distinction between colonization and infection with pathogenic micro-organisms. For example, bacterial secondary infections in children with primary viral lower respiratory tract infections are often diagnosed based on cultures from upper airway samples. The question remains whether the detected organism is really the cause of infection or whether it is just a contaminant from the upper airway (Jacobs and Dagan, 2004). The development of a diagnostic test based on gene expression patterns in the host may provide more specific information. In the future, the development of such diagnostic tools may help the clinician choose an effective treatment strategy and reduce inappropriate antibiotic use.

A diagnostic delay in infectious diseases can lead to delayed initiation of therapy, severe complications, and long term consequences that may include death. Such a

delay in diagnosis may be prevented by the development of diagnostic biomarkers. Encephalitis, for example, is a complex, severe, neurological syndrome associated with significant morbidity and mortality. It is characterized by inflammation of the brain parenchyma, and children often present with drowsiness, fever, headache, seizures, or focal neurological signs. A delay in treatment may lead to irreversible brain damage. Diagnosis is often presumptive and based on clinical characteristics or increased serological antibody titers. Unfortunately, the causative pathogen is often hard to detect in the central nervous system itself. The final diagnosis is sometimes based on the detection of pathogens in cultures from other sites such as the respiratory tract (Glaser et al., 2006). Indirectly diagnosing encephalitis based on respiratory samples is rather inaccurate and demonstrates the need for new and better diagnostic tools. New microarray and proteomics technologies may contribute to improved diagnosis and treatment. Microarrays have been developed to simultaneously identify different viral and bacterial pathogens in cerebrospinal fluid (CSF) (Boriskin et al., 2004; Ben et al., 2008). To our knowledge, gene expression studies for differentiating pathogens based on the host response have not yet been performed. The differentiation of pathogens based on the host response may provide better insight in pathogenesis and disease specific profiles in blood or cerebrospinal fluid; it may also prove useful in diagnosing encephalitis. The identification of biomarkers related to clinical profiles or recognition of subgroups in encephalitis may help predict outcome and provide insight into the efficacy of therapy.

Another infection in which diagnostic difficulties often lead to a delay in appropriate therapy is infective endocarditis. The clinical diagnosis of infective endocarditis may be difficult, as fever can be the only symptom. Rapid diagnosis followed by appropriate treatment is of critical importance for survival. However, in 3–31% of patients, causative pathogens remain undetected. Fenollar et al. analyzed serum samples from 88 patients with a clinical suspicion of endocarditis. They identified 66 different protein peaks in patients with confirmed endocarditis as compared to those in whom the diagnosis was excluded (Fenollar et al., 2006). From these 66, they developed a diagnostic assay based on 7 protein peaks. Despite this limited number of differentially expressed proteins, the test was still able to classify the majority (88%) of patients correctly.

3.3 Future Perspectives for Biomarker Development

Proteomic and genomic technologies have been shown to contribute to improved insight into disease pathogenesis and may be useful in diagnosing infections and providing information about disease susceptibility. However, clinical application of these technologies has not yet been developed. Future research should focus on the validation of previously identified biomarkers as well as the development of new diagnostic assays.

The relationship between gene expression profiles and disease outcome is another interesting field of research. Prognostic biomarkers could be import in infectious diseases, helping predict likely disease course and a selection of patients

most likely to benefit from treatment. At present, their clinical use is limited to the field of cancer research where several studies suggest that molecular classification of tumors, based on gene expression, may identify distinct prognostic subtypes. Alizedah et al. have detected two subtypes of diffuse B-cell lymphoma with different survival patterns (Alizadeh et al., 2000). The Mamma print, a test developed by the Dutch Cancer Institute, identifies different breast cancer subtypes by analysis of expression profiles involving 70 genes indicative of poor prognosis (van't Veer et al., 2002; van't Veer and Bernards, 2008). These studies are based on hierarchical clustering of subgroups with similar gene expression profiles. Hierarchical clustering methods in gene expression analyses might prove useful in pediatric infectious diseases; although to date, few studies have been done. Chaussabel et al. have generated a potentially useful framework for the visualization and functional interpretation of microarray based disease-specific transcriptional signatures (Chaussabel et al., 2008), and, in addition, to the identification of biomarkers for monitoring inflammatory disease activity (in SLE) showed its potential value in the evaluation of disease progression and thus prognosis (Chaussabel et al., 2008).

In our department, we are currently conducting a clinical study to identify classifier genes that predict outcome in children suffering from lower respiratory tract viral infections (VIRGO study). Using microarray analyses of blood leukocytes and respiratory samples, we aim to identify biomarkers that distinguish children with a relatively mild course of disease from those who will deteriorate and require supplemental oxygen or mechanical ventilation. In the early phase of infection, these biomarkers may help decide whether a child needs to be hospitalized.

Another potential application is to guide treatment by allowing therapy to be tailored both to the specific pathogen, including its antimicrobial resistance properties and to host characteristics, including the immune response, which leads to more focused drug use and improved outcome. An early example of genotype-guided, individualized treatment strategy is the adjustment of isoniazid dosing regimen in adults with tuberculosis. *N*-acetyltransferase type 2 (NAT2) plays an important role in isoniazid metabolism and genetic polymorphisms of this enzyme can alter the response to the drug. Determining the NAT2 genotype prior to isoniazid administration can predict pharmacokinetic variability and therapeutic response (Kinzig-Schippers et al., 2005).

Individualized treatment strategies will be extremely helpful in treating tuberculosis in children. At present, children with tuberculosis are treated with up to four tuberculostatic agents for two months followed by a two drug regimen during a 4 month continuation phase. To date, there is no laboratory tool to monitor response to therapy. Moreover, difficulties in identification of *M. tuberculosis* from, for example, induced sputum or gastric lavage, contribute to diagnostic and therapeutic uncertainty (Newton et al., 2008). Consequently, non-specific clinical features such as symptom improvement, weight gain, and radiological features of chest disease are used as markers for therapeutic response (Donald and Schaff, 2007). The identification of biomarkers for tuberculosis disease activity may provide more specific monitoring strategies leading to more focused prescribing, fewer adverse effects, and less multidrug resistance.

4 Conclusions

Diagnostic uncertainty in infectious disease may lead to delayed diagnosis and inappropriate use of antibiotics. The development of diagnostic biomarkers for infectious diseases may result in more rapid diagnosis, more reliable discrimination between infection and non-infectious diseases, more improved management, better course and outcomes, and less inappropriate use of antibiotics.

Microarrays and proteomic technologies are beginning to contribute to improved understanding of the pathogenesis of a wide variety of infectious diseases and have great prospects for the future. These technologies can be targeted both at direct pathogen detection and at characterization of the host response, which may also assist in diagnosis and disease monitoring as well as predicting the individual's susceptibility to disease, response to medical therapy, and overall prognosis. Although promising, the clinical application of these technologies in infectious diseases is limited at present. Current research focuses on sophisticated highly specialized techniques, but future work will need to be directed at clinical validation studies to collect data on clinical applicability, accuracy and cost effectiveness. Translating biomarker research into clinically useful tests will be difficult and time and labor intensive. The ultimate goal is to develop clinically relevant, cheap, rapid diagnostic and prognostic biomarker tests which use biological samples that are easy to obtain from the patient and which generate reliable and easily interpreted results.

References

Agranoff, D., Stich, A., Abel, P., & Krishna, S. (2005). Proteomic fingerprinting for the diagnosis of human African trypanosomiasis. *Trends Parasitol* 21(4), 154–157.

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A. et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403(6769), 503–511.

Allantaz, F., Chaussabel, D., Stichweh, D., Bennett, L., Allman, W. et al. (2007). Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. *J Exp Med* 204(9), 2131–2144.

Allison, A.C. (1954). Protection afforded by sickle-cell trait against subtertian malareal infection. *Br Med J* 6, 1(4857), 290–294.

An, P., Wang, L.H., Hutcheson-Dilks, H., Nelson, G., Donfield, S., Goedert, J.J., & Rinaldo, C.R. et al. (2007). Regulatory polymorphisms in the cyclophilin A gene, PPIA, accelerate progression to AIDS. *PLoS Pathog* 3(6), e88.

Ben, R.J., Kung, S., Chang, F.Y., Lu, J.J., Feng, N.H., & Hsieh, Y.D. (2008). Rapid diagnosis of bacterial meningitis using a microarray. *J Formos Med Assoc* 107(6), 448–453.

Boriskin, Y.S., Rice, P.S., Stabler, R.A., Hinds, J.A., Ghusein, H., Vass, K. et al. (2004). DNA microarrays for virus detection in cases of central nervous system infection. *J Clin Microbiol* 42(12), 5811–5818.

Bryant, P.A., Venter, D., Robins-Browne, R., & Curtis, N. (2004). Chips with everything: DNA microarrays in infectious diseases. *Lancet Infect Dis* 4(2), 100–111.

Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N. et al. (2008). A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 29(1), 150–164.

Cooke, G.S., & Hill, AV. (2001). Genetics of susceptibility to human infectious disease. *Nat Rev Genet* 2(12), 967–977.

Coombes, K.R., Morris, J.S., Hu, J., Edmonson, S.R., & Baggerly, K.A. (2005). Serum proteomics profiling – a young technology begins to mature. *Nat Biotechnol* 23(3), 291–292.

Crawford, D.C., Akey, D.T., & Nickerson, D.A. (2005). The patterns of natural variation in human genes. *Annu Rev Genomics Hum Genet* (6), 287–313.

Donald, P.R., & Schaaf, H.S. (2007). Old and new drugs for the treatment of tuberculosis in children. *Paediatr Respir Rev* 8(2), 134–141.

Feuk, L., Marshall, C.R., Wintle, R.F., & Scherer, S.W. (2006). Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum Mol Genet* 15(Spec No 1): R57–R66.

Fenollar, F., Goncalves, A., Esterni, B., Azza, S., Habib, G., Borg, J.P., & Raoult, D. (2006). A serum protein signature with high diagnostic value in bacterial endocarditis: results from a study based on surface-enhanced laser desorption/ionization time-of-flight mass spectrometry. *J Infect Dis* 194(10), 1356–1366.

Galetto-Lacour, A., Zamora, S.A., & Gervaix, A. (2003). Bedside procalcitonin and C-reactive protein tests in children with fever without localizing signs of infection seen in a referral center. *Pediatrics* 112(5), 1054–1060.

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. (2007). *Nature* 7, 447(7145), 661–678.

Glaser, C.A., Honarmand, S., Anderson, L.J., Schnurr, D.P., Forghani, B., Cossen, C.K. et al. (2006). Beyond viruses: clinical profiles and etiologies associated with encephalitis. *Clin Infect Dis* 43(12), 1565–1577.

Graves, P.R., & Haystead, T.A. (2002). Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* 66(1), 39–66.

Herd, D. (2007)In children under age three does procalcitonin help exclude serious bacterial infection in fever without focus? *Arch Dis Child* 92(4), 362–364.

Hill, A.V. (1998). The immunogenetics of human infectious diseases. *Annu Rev Immunol* (16), 593–617.

Hill, A.V. (2006). Aspects of genetic susceptibility to human infectious diseases. *Annu Rev Genet* (40), 469–486.

Hirschhorn, J.N., & Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6(2), 95–108.

Hodgetts, A., Levin, M., Kroll, J.S., & Langford, P.R. (2007). Biomarker discovery in infectious diseases using SELDI. *Future Microbiol* (2), 35–49.

The International HapMap Project. (2003). *Nature* 18, 426(6968), 789–796.

Jacobs, M.R., & Dagan, R. (2004). Antimicrobial resistance among pediatric respiratory tract infections: clinical challenges. *Semin Pediatr Infect Dis* 15(1), 5–20.

Jenner, R.G., & Young, R.A. (2005). Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 3(4), 281–294.

Jepson, A.P., Banya, W.A., Sisay-Joof, F., Hassan-King, M., Bennett, S., & Whittle, H.C. (1995). Genetic regulation of fever in Plasmodium falciparum malaria in Gambian twin children. *J Infect Dis* 172(1), 316–319.

Jepson, A., Fowler, A., Banya, W., Singh, M., Bennett, S., Whittle, H., & Hill, A.V. (2001). Genetic regulation of acquired immune responses to antigens of Mycobacterium tuberculosis: a study of twins in West Africa. *Infect Immun* 69(6), 3989–3994.

Kawada, J., Kimura, H., Kamachi, Y., Nishikawa, K., Taniguchi, M., Nagaoka, K., & Kurahashi, H. et al. (2006). Analysis of gene-expression profiles by oligonucleotide microarray in children with influenza. *J Gen Virol* 87(Pt6), 1677–1683.

Kinzig-Schippers, M., Tomalik-Scharte, D., Jetter, A., Scheidel, B., Jakob, V., Rodamer, M. et al. (2005). Should we use N-acetyltransferase type 2 genotyping to personalize isoniazid doses? *Antimicrob Agents Chemother* 49(5), 1733–1738.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S. et al. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* 348(20), 1953–1966.

Liu, M., Popper, S.J., Rubins, K.H., & Relman, D.A. (2006). Early days: genomics and human responses to infection. *Curr Opin Microbiol* 9(3), 312–319.

MalariaGen. (2009). Retrieved from http://malariagen net/access

Newton, S.M., Brent, A.J., Anderson, S., Whittaker, E., & Kampmann, B. (2008). Paediatric tuberculosis. *Lancet Infect Dis* 8(8), 498–510.

O'Brien, S.J., & Nelson, G.W. (2004). Human genes that limit AIDS. *Nat Genet* 36(6), 565–574.

Papadopoulos, M.C., Abel, P.M., Agranoff, D., Stich, A., Tarelli, E., Bell, B.A. et al. (2004). A novel and accurate diagnostic test for human African trypanosomiasis. *Lancet* 363(9418), 1358–1363.

Patterson, S.D., & Aebersold, R.H. (2003). Proteomics: the first decade and beyond. *Nat Genet* 33(Suppl), 311–323.

Ramilo, O., Allman, W., Chung, W., Mejias, A., Ardura, M., Glaser, C. et al. (2007). Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109(5), 2066–2077.

Rosseau, S., Hocke, A., Mollenkopf, H., Schmeck, B., Suttorp, N., Kaufmann, S.H., & Zerrahn, J. (2007). Comparative transcriptional profiling of the lung reveals shared and distinct features of Streptococcus pneumoniae and influenza A virus infection. *Immunology* 120(3), 380–391.

Schaad, U.B. (1997). Toward an integrated program for patient care in pediatric infections. *Pediatr Infect Dis J* 16(3 Suppl), S34–S38.

Sorensen, T.I., Nielsen, G.G., Andersen, P.K., & Teasdale, T.W. (1988). Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med* 318(12), 727–732.

Stein, C.E., Inoue, M., & Fat, D.M. (2004). The global mortality of infectious and parasitic diseases in children. *Semin Pediatr Infect Dis* 15(3), 125–129.

Stich, A., Abel, P.M., & Krishna, S. (2002). Human African trypanosomiasis. *BMJ*, 325(7357), 203–206.

Tang, B.M., McLean, A.S., Dawes, I.W., Huang, S.J., & Lin, R.C. (2007). The use of gene-expression profiling to identify candidate genes in human sepsis. *Am J Respir Crit Care Med* 176(7), 676–684.

Tong, H.H., Long, J.P., Li, D., & DeMaria, T.F. (2004). Alteration of gene expression in human middle ear epithelial cells induced by influenza A virus and its implication for the pathogenesis of otitis media. *Microb Pathog* 37(4), 193–204.

van Rossum, A.M., Wulkan, R.W., & Oudesluys-Murphy, A.M. (2004). Procalcitonin as an early marker of infection in neonates and children. *Lancet Infect Dis* 4(10), 620–630.

van't Veer, L.J., & Bernards, R. (2008). Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 452(7187), 564–570.

van't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L. et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415(6871), 530–536.

Walker, E.J., & Siminovitch, K.A. (2007). Primer: genomic and proteomic tools for the molecular dissection of disease. *Nat Clin Pract Rheumatol* 3(10), 580–589.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., & DeRisi, J.L. (2002). Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci USA* 26, 99(24), 15687–15692.

Wellcome Trust Case-Control Consortium (WTCCC) (2006). Retrieved from http://ccc.sanger.ac.uk/info/overview.shtml

Xavier, R.J., & Rioux, J.D. (2008). Genome-wide association studies: a new window into immune-mediated diseases. *Nat Rev Immunol* 8(8), 631–643.