Characterizing Vaccine Responses Using Host Genomic and Transcriptomic Analysis

Daniel O’Connor and Andrew J. Pollard
Department of Paediatrics, University of Oxford, and the National Institute for Health Research Oxford Biomedical Research Centre, United Kingdom

Vaccines have had a profound influence on human health with no other health intervention rivaling their impact on the morbidity and mortality associated with infectious disease. However, the magnitude and persistence of vaccine immunity varies considerably between individuals, a phenomenon that is not well understood. Recent studies have used contemporary technologies to correlate variation in the genome and transcriptome to immunological measures of vaccine responsiveness. These approaches have provided fresh insight into the intrinsic factors determining the potency and duration of vaccine-induced immunity. The fundamental challenge will be to translate these findings into innovative and pragmatic strategies to develop new and more effective vaccines.

Keywords. vaccine immunology; immunogenetics; heritability; genomics; transcriptomics.

Since Edward Jenner’s epochal demonstration in 1796 that inoculation with the vaccinia virus induced resistance to the related and highly virulent variola virus, billions of doses of vaccines have been administered. Vaccines have had a profound influence on health particularly in childhood, averting an estimated 2.5 million child deaths per year [1]. However, interindividual variation in vaccine responses remains both perplexing and problematic, with a considerable proportion not achieving the putative level of protection following routine immunizations (eg, 10%–15% of adults fail to respond to 3 doses of hepatitis B virus [HBV] vaccine) [2].

Heterogeneous vaccine responses in infancy are even more noteworthy as immune responses are generally poorer in the young and immunity can wane rapidly [3]. A number of factors including age, sex, nutrition, ultraviolet light exposure, smoking, and infectious disease have been shown to influence vaccine responses [4]. Importantly, a large proportion of the variation seen in responses is thought to be attributable to factors present in the vaccinee’s genome. Recent advances in the “omics,” particularly genomics and transcriptomics, present highly amenable avenues to explore and dissect the mechanisms involved in complex biological systems, such as vaccine responses (Figure 1).

HERITABILITY OF VACCINE RESPONSES

A number of studies in twins have evaluated the role of genetic variation on vaccine responses, showing these to be 39%–89% heritable (Table 1) [5–8]. Heritability can be assessed by comparing phenotypical concordance between monozygotic (MZ) and dizygotic (DZ) twins as both twin pairs are likely to be exposed to similar environmental factors but share 100% and 50% of their segregating DNA sequence, respectively. An early study of Danish adults following a single dose of 23-valent pneumococcal polysaccharide vaccine found greater correlation in the antibody rise 4 weeks postvaccination in MZ than DZ twin pairs for all 8 pneumococcal antigens assessed. Interestingly, the heritability of responses to pneumococcal polysaccharides appeared to be serotype dependent, with only serotypes 1,
2, 4, and 19F reaching statistical significance [8]. Another twin study in German adults also found greater correlation in vaccine-specific antibody concentrations 4 weeks after 3 doses of the combined HBV vaccine and hepatitis A virus (HAV) vaccine in MZ than DZ twin pairs [5].

A prospective study of a birth cohort of 207 Gambian twin pairs assessed the heritability of both humoral and cellular responses to a number of vaccine antigens [6, 7]. Significantly greater correlation within MZ twin pairs was observed for all specific antibody responses (hepatitis B, polio, tetanus, diphtheria, and Haemophilus influenzae type b [Hib]), as well as for total serum immunoglobulin G levels [6, 7]. Furthermore, interferon-γ release following stimulation with mycobacterial purified protein derivative (PPD), killed Mycobacterium tuberculosis, pertactin,

![Figure 1](https://academic.oup.com/cid/article-abstract/57/6/860/328274)

Figure 1. Overview of molecular approaches to describe and predict vaccine responses.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Vaccine</th>
<th>Study Population</th>
<th>Response Measured</th>
<th>Heritability</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5]</td>
<td>Inactivated HAV + recombinant HBsAg</td>
<td>192 monozygotic and 190 dizygotic German adult twins aged 18–65 y</td>
<td>Anti-HAV antibody</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Anti-HBsAg antibody</td>
<td>61%</td>
</tr>
<tr>
<td>[8]</td>
<td>Pneumococcal vaccination</td>
<td>48 monozygotic and 36 dizygotic Caucasian adult twins aged 21–65 y</td>
<td>IgG and IgG2, serotype-specific antibodies</td>
<td>Yes, varies between serotypes</td>
</tr>
<tr>
<td>[7]</td>
<td>Hib conjugate vaccine</td>
<td>86 monozygotic and 294 dizygotic Gambian twins aged 5 mo</td>
<td>Anti-PRP IgG</td>
<td>51%</td>
</tr>
<tr>
<td>[6]</td>
<td>Recombinant HBsAg</td>
<td>96 monozygotic and 318 dizygotic Gambian twins aged 5 mo</td>
<td>Anti-HBsAg antibody</td>
<td>77%</td>
</tr>
<tr>
<td>[6]</td>
<td>Oral polio vaccine</td>
<td>Anti-poliovirus antibody</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>[6]</td>
<td>DTP</td>
<td>Anti-tetanus toxoid antibody concentrations</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td>[6]</td>
<td>BCG, 0.05 mL</td>
<td>Anti-diphtheria toxoid antibody concentrations</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytokine responses to PPD</td>
<td>46%</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BCG, Bacillus Calmette-Guérin; DTP, diphtheria-tetanus-pertussis; HAV, hepatitis A virus; HBsAg, hepatitis B surface antigen; Hib, Haemophilus influenzae type b; IgG, immunoglobulin G; PPD, purified protein derivative; PRP, purified polyribosylribitol phosphate.
and filamentous hemagglutinin, as well as interleukin 13 release following PPD, mycobacterial 65-kDa heat shock protein, pertussis toxin, and tetanus toxoid (TT), were all also shown to be heritable [6].

Heritability measures are context and population specific; therefore, caution should be taken when interpreting and comparing reported heritability measures. For example, factors that influence overall phenotypic variance, such as measurement assay variability and population heterogeneity, will affect heritability calculations. Studies of vaccine responses early in life may be less confounded by the accumulation of environmental factors influencing immune responses, potentially accounting for the greater heritability exhibited in specific antibody to hepatitis B vaccine in infants (77%) than in adults (61%) [5, 6]. Consistent with this hypothesis, the infant study also showed greater heritability for vaccines given at birth (BCG and oral polio vaccine [OPV]) than even those administered at 2 months of age (diphtheria toxoid [DT] and TT) [6]. However, substantial disparity is observed even in heritability of responses to vaccines given concomitantly, implying the relative genetic contribution may also be vaccine dependent. Interestingly, antibody responses to protein toxoids, TT and DT, correlate and show similar heritability, potentially indicating shared genetic determinants [6]; in contrast, antibody responses to hepatitis B vaccine and OPV show no correlation and disparate heritability, suggesting that distinct genetic factors are involved [6].

**GENETIC DETERMINANTS OF VACCINE RESPONSES**

Despite the evident heritability of vaccine responses, the particular factors involved remain largely uncharacterized. Several candidate gene studies and, more recently, 2 genome-wide association studies (GWASs) have investigated the effect of genetic variation on immunological and physiological responses to vaccination.

**Genome-wide Association Studies**

A GWAS of hepatitis B vaccine responses in 3644 healthy Indonesians found 47 single-nucleotide polymorphisms (SNPs) representing 3 independent signals across the human leukocyte antigen (HLA) region that associated with vaccine-specific antibody concentrations ($P \leq 1 \times 10^{-7}$) [9]. The strongest of these independently segregating signals was rs3135363 ($P = 6.53 \times 10^{-22}$), a SNP tagging a haplotype block that includes HLA-DR, while the 2 other index SNPs were in the 3’ untranslated region (UTR) of HLA-DPB1 and in a HLA class III gene-rich region [9]. Intriguingly, the HLA-DPB1 tagging SNP rs9277535 has previously been associated with chronicity of hepatitis B infection in several Asian subpopulations [10]. Clearance of hepatitis B virus coincides with the production of anti–hepatitis B surface antigen (HBsAg) antibodies; however, chronically infected individuals fail to produce these antibodies despite detectable HBsAg in their circulation. The antigenic component of hepatitis B vaccine is HBsAg. Therefore, association concordance between the hepatitis B vaccine response and HBV clearance studies suggests that individuals are predisposed to chronic infection by a genetically determined inability to produce anti-HBsAg antibodies, and that part of this genetic influence lies proximal to the HLA-DPB1 locus [9, 10].

Another published vaccine GWAS assessed antibody levels in a multiethnic cohort of 1014 healthy adults following pustule-forming smallpox vaccination [11]. Three distinct ethnic groups were included in the analysis: whites (n = 580), African Americans (n = 217), and Hispanics (n = 217). Only a single SNP within the promoter region of a pseudogene surpassed the authors’ proposed level of significance ($P < 5 \times 10^{-7}$) in the white cohort, with an additional 10 SNPs, including one within the intronic region of the B-cell receptor gene, reaching $P < 1 \times 10^{-5}$ [11].

GWAS provides a genome-wide assessment of many hundreds of thousands of genetic variants and has the benefit of not requiring exhaustive prior knowledge of the genes involved in a particular trait. However, it is important to consider the implications of this expansive approach, particularly the inherent need for stringent statistical corrections. Using Bonferroni correction for approximately 1 million independent variants in the European American genome gives a threshold for statistical significance of $P < 5 \times 10^{-8}$. Such a conservative threshold makes it necessary to work with very large sample sizes to provide statistical power and may result in the exclusion of many markers of small to moderate effect.

**Candidate Gene Studies**

Until recently, the candidate gene approach had been the mainstay of investigations into the genetic determinants of vaccine responses. Selecting genes with a putative role in immunological responses has the clear disadvantage of dependency on a priori knowledge. Conversely, candidate gene studies require less statistical corrections than genome-wide approaches and can therefore be more powerful and conducive to the moderate sample size setting. A variety of genes have been investigated, but for simplicity these can be separated into 3 broad themes: antigen processing and presentation, innate recognition, and cellular signaling.

**Antigen Processing and Presentation**

HLA molecules facilitate the presentation of peptide antigens to $\alpha\beta$ T-cell receptors. HLA genes are highly polymorphic, introducing a diverse ability to bind and present peptide within the population. Comparisons of responses within HLA-identical and HLA-disparate DZ twins have estimated that 40% of the genetic variation influencing specific antibody responses to hepatitis B vaccine is attributable to the HLA genes alone [5].
HLA polymorphisms have also been associated with responses to measles, rubella, and anthrax vaccines [12–14].

Innate Recognition Receptors

Innate recognition receptors, such as Toll-like receptors (TLRs), instigate immunological responses to antigens based upon the recognition of conserved microbial motifs and/or intrinsic host cell “danger” signals. Polymorphisms in a number of the innate receptor pathways, including TLRs, have been associated with variable vaccine responses [15–17]. In particular, SNPs within the TLR3 gene locus have been associated with measles-specific antibody persistence following measles-mumps-rubella vaccination in adolescents [15]. Furthermore, TLR3 SNPs have been associated with the persistence of serogroup C meningococcal (MenC)–specific antibody concentrations in children and adolescents following MenC-conjugate vaccination [16]. In addition, a SNP within TIRAP, (rs1893352) an essential adaptor molecule that binds to both TLR2 and TLR4, has been associated with Hib vaccine failure, with the recessive homozygote genotype being 7.6 times more common in the bacteremic study cohort than in healthy controls (P = 1.3 × 10−6) [18].

Cellular Signaling

Cytokines are small intercellular signaling proteins with pleiotropic roles in humoral and cellular immunity. Polymorphisms in numerous cytokine genes have been associated with responses to vaccination, with some cytokine genes being implicated in responses to a variety of vaccines. For example, polymorphisms within IL-10 have been associated with concentrations of specific antibody to diphtheria toxoid, hepatitis A vaccine, hepatitis B vaccine, and pneumococcal serotype 6B capsular polysaccharide [19–21]. The minor allele in the promoter region on IL-10 (rs1800896) has been associated with higher antibody responses to both diphtheria and hepatitis A vaccines, while another IL-10 promoter SNP (rs1554286) has been associated with Hib vaccine failure [18–20].

Vaccine-Associated Adverse Events

Vaccine-associated adverse events are important in the public acceptance of vaccines. Although severe reactions are very rare, both the public and the media often have an inflated perceived risk. On the other hand, mild reactions such as pain, swelling, and fevers are relatively common and may also be important in the public perception of vaccine safety. The determinants of vaccine-associated adverse reactions are not well understood and there are currently few data evaluating the genetic involvement. However, SNPs in MTHFR (methyleneetetralydrofolate reductase) and IRF-1 (interferon regulatory factor 1) have been associated with fever, skin eruptions, and/or lymphadenopathy following vaccinia vaccine [22].

Replication and Functional Validation

There are an accumulating number of genetic variants being implicated in vaccine responses; however, few have been assigned any functional consequence. The major “A” allele at rs9277535 in the 3′ UTR of HLA-DPB1, associated with clearance of HBV and hepatitis B vaccine antibody responses, has recently been linked with the expression of HLA-DPB1 messenger RNA in both liver and lymphoblastoid cell lines [9, 10, 23]. Taken together, these data suggest that variations within the 3′ UTR of HLA-DPB1 influence responses to HBV and hepatitis B vaccine via HLA-DPB1 expression. Another variant with a strong case for a functional implication is the nonsynonymous MAL/TIRAP variant Ser180Leu, associated with Hib vaccine failure, which lies within the proposed functional interface with MyD88 (myeloid differentiation primary response gene 88) and may result in a partial loss of function [18, 24]. More studies providing functional validation for associated genetic loci are needed to enable robust inferences to be formed about the genetic determinants of vaccine responses.

In the absence of functional validation, replication in an independent population is critical to corroborate “true” signals and eliminate false positives. Replication is particularly important as associations seldom surpass conservative thresholds for statistical significance, due to corrections for multiple testing and inadequate sample sizes. Tiered study designs, with an initial “discovery” cohort genotyped broadly that is used to produce a smaller subset of variants to be followed up in a “replication” cohort, have been shown to be both powerful and cost-effective [25]. For example, the GWAS of hepatitis B vaccine responses adopted this tiered approach, genotyping 455 508 SNPs in a discovery set of 1683 individuals and pursuing a subset of 1706 SNPs in a replication cohort of 1931 individuals, resulting in 47 significant SNPs after corrections for multiple testing [9]. The association between HLA-DQA1*0201 and both measles- and rubella-specific antibody responses have also been replicated in independent populations [13, 14].

VACCINE TRANSCRIPTOMICS

Contemporary transcriptomic analyses have revolutionized the way in which complex biological systems can be interrogated. Transcriptomic analyses are being utilized to describe and predict responses to existing vaccine, as well as in early vaccine development and preclinical research (Tables 2 and 3) [26–35].

Use of Transcriptomic Analyses in Clinical Trials

Transcriptomic analyses can be used in vaccine clinical trials to (1) characterize gene expression changes evoked following vaccination, (2) correlate transcriptional changes to measures of immunogenicity and/or tolerability, and (3) describe novel biomarkers of vaccine immunogenicity or efficacy. A seminal study in this area used microarray transcriptional profiling to
<table>
<thead>
<tr>
<th>Reference</th>
<th>Vaccine</th>
<th>Summary of Study Details</th>
<th>Summary of Key Findings</th>
</tr>
</thead>
</table>
| [30]      | Vaccinia (live-attenuated) or YF (live-attenuated) | 44 healthy adults (24 receiving vaccinia and 20 receiving YF vaccine)  
Sampling prior to and 4–7 d following vaccination  
Direct transcriptomic analysis of PBMCs  
Agilent human 1 cDNA microarrays (Vaccinia study) and Human 1A Oligo Microarray Kit (V2) (YF study) | Vaccinia-induced 475 and YF-induced 615 differentially expressed genes  
80 genes were differentially expressed following both vaccines  
Vaccinia responses were characterized by repression of B- and T-cell associated genes and the induction of monocyte/macrophage-associated genes  
YF responses predominantly induced the downregulation of genes associated with protein biosynthesis |
| [28]      | Influenza (inactivated) or 23-valent pneumococcal (plain polysaccharide) | 46 healthy adults  
Sampling at hours 0, 1.5, 3, 6, 9, 12, 15, 24, 36, and 48 and days 3, 7, 10, 14, and 21  
Direct transcriptomic analysis of whole blood  
Illumina Human HT-12v3 Expression BeadChips | Modular analysis described 8 gene sets that were consistently modulated following either influenza or 23-valent pneumococcal vaccination  
The majority of significant modules were vaccine specific with the 2 vaccines eliciting distinct innate immune responses  
Nine modules were uniquely responsive to influenza, with 3 of the 4 upregulated modules being associated with viral responses  
Five of the 6 modules that were uniquely responsive to 23-valent pneumococcal vaccine were associated with inflammation |
| [26]      | YF (live-attenuated) | 25 healthy adults  
Sampling at days 0, 1, 3, 7, 21 postvaccination  
Direct transcriptomic analysis of PBMCs  
Affymetrix Human U133 Plus 2.0 Array | 65 DEGs in the first week following vaccination  
Gene sets at day 3 and 7 predicted specific CD8+ and antibody responses  
The expression of TNFRSF17 at day 7 was highly predictive of antibody responses |
| [29]      | Influenza vaccine (inactivated) | 119 healthy men  
Sampling at days 0, 1, 3, and 14 postvaccination  
Direct transcriptomic analysis of whole blood  
Illumina Human HT-12v3 Expression BeadChips | Early expression of several genes encoding proteins interacting with the innate pathogen sensors TLR7 and TLR8, postvaccination  
Maximum change in gene expression occurred around day 1 but that the peak number of genes with maximum expression values occurs at day 14  
The expression of TNFSF13B the protein product of which is a ligand for the protein encoded by TNFRSF17 shown to be correlated with antibody responses |
| [27]      | Influenza (live-attenuated or inactivated) | 56 healthy adults (28 receiving each vaccine)  
Sampling at days 0, 3, and 7 postvaccination  
Direct transcriptomic analysis of PBMCs and sorted cell populations  
Affymetrix Human U133 Plus 2.0 Array | Disparate profiles evoked with live-attenuated influenza and TIV  
DEGs following TIV were predominantly from the mDCs and B cells, whereas pDC generated the highest number of DEGs following LAIV  
TNFRSF17 was shown to be predictive of antibody responses |
| [31]      | Malaria (protein subunit) | 39 healthy adult volunteers  
Sampling prevaccination, 1, 3, and 14 d postvaccination, as well as 5 d postchallenge  
Direct transcriptomic analysis of PBMCs  
Affymetrix Human U133 A 2.0 and U133 Plus 2.0 arrays | 63 inflammatory response genes differentially expressed 1 and 3 d postvaccination  
RTS,S vaccinees differentially expressed >10 times as many genes as unvaccinated individuals when challenged with malaria  
A 393 gene classifier set 5 days after challenge was shown to correlate with parasitemia following challenge |
describe 97 differentially expressed genes (DEGs) in peripheral blood mononuclear cells (PBMCs) in the week following live-attenuated yellow fever vaccine (YF-17D) [26]. Sixty-five of these DEGs were consistently modulated in an independent replication cohort, with gene ontology analysis showing enrichment of genes related to immune response, innate immune response, and response to virus [26]. This study was also able to show a correlation between the expression of particular genes and specific CD8+ T-cell and antibody responses, most intriguingly demonstrating that these correlated genes could predict immunogenicity in an independent replication cohort with a high degree of accuracy [26]. Of particular note, a gene set containing the TNFRSF17 gene, which encodes a receptor for the B-cell growth factor BLYS-BAFF, was highly predictive of antibody responses [26].

A further study of the PBMCs in the week following influenza vaccination highlighted disparities between gene profiles evoked following the live-attenuated influenza vaccine (LAIV) given intranasally and the intramuscularly administered trivalent inactivated influenza vaccine (TIV) [27]. Although a core set, including genes involved in inflammatory and antimicrobial responses, was altered similarly, the majority of the DEGs were vaccine specific. Recipients of the LAIV differentially expressed several interferon-related genes at day 3, potentially reflecting the replication competency of live-attenuated viral vaccines, which is consistent with observations following YF-17D [26, 27]. Furthermore, the analysis of sorted cells implied that DEGs following TIV predominantly originated from the myeloid dendritic cells and B cells, whereas plasmacytoid dendritic cells contributed the greatest DEGs following LAIV [27]. These data imply that the 2 vaccine strategies, which differ both in composition and route of administration, evoke very divergent molecular responses that are propagated by different cell types. This study then correlated gene expression with fold-rises in hemagglutination inhibition (HAI) antibody titers, showing the expression of a gene set containing TNFRSF17 that was highly predictive of antibody responses [27]. Furthermore, a recent study has found that the expression of a module of plasmablast-related genes, including TNFRSF17, correlated with HAI titers and serotype-specific antibody concentrations following TIV and 23-valent pneumococcal vaccine, respectively [28]. A further study found that the expression of the B-cell growth factor BLYS-BAFF (the ligand for TNFRSF17) correlated with post-TIV HAI responses [29]. Transcriptional studies have also alluded to novel roles for genes with previously unappreciated contributions to immune regulation. For example, CaMKIV expression was shown to inversely correlate with HAI responses following TIV, a finding that was subsequently functionally substantiated by the demonstration that CaMKIV-deficient mice show enhanced TIV antibody responses [27].

Gene expression analysis has also been undertaken in the investigation of vaccines for which the correlates of protection have yet to be determined, such as malaria and tularemia [31, 32]. Sixty-three inflammatory response genes were shown to be differentially expressed in PMBCs a day after the novel adjuvanted RTS,S malaria vaccine [31]. When these vaccinees were subsequently challenged with Plasmodium falciparum, 6219 genes were upregulated and 5670 downregulated, compared to just 461 upregulated and 626 downregulated genes in vaccine-naive individuals [31]. Despite the increased breadth of responses in vaccine recipients, two-thirds were not protected; however, a 393 gene classifier set 5 days after challenge accurately predicted parasitemia following infectious challenge [31].

**Use of Transcriptomic Analyses in Preclinical Studies**

Preclinical vaccine studies have also made use of transcriptomic approaches, particularly to gain insight into the workings of vaccine adjuvants [33–35]. A study investigating the effects of intramuscular injection with MF59 oil-in-water emulsion, nonmethylated CpG oligonucleotide (CpG), or aluminium hydroxide (alum) in mice found that these adjuvants modulated a core set...
of 168 genes in muscle specimens. This core set was enriched for genes involved in cytokine–cytokine receptor interaction, host–pathogen interaction, and defense immunity protein activity [33]. MF59 induced nearly 3 times as many genes as the other adjuvants, preferentially inducing the transcription of genes involved in complement activation, prostaglandin synthesis, interleukin 1 signaling, and leukocyte transendothelial migration [33].

The authors of this study then went on to assess gene expression following influenza subunit and TT vaccines, adjuvanted with either MF59, alum, CpG, resiquimod (R848), or Pam3CSK4 [34]. MF59 was the strongest modulator of transcriptional events at the injection site, particularly genes involved in leukocyte transendothelial migration and cytokine binding, implying that this adjuvant may be a strong recruiter of cells from the bloodstream [34]. However, R848 was shown to be the strongest modulator of gene expression in the draining lymph nodes, which was characterized by enrichment of type I and II interferon genes as well as T-helper 1 cytokine genes, suggesting that this adjuvant is a potent activator of lymphoid tissue [34].

**GENOMIC AND TRANSCRIPTOMIC INDICATORS OF VACCINE EFFICACY**

Hitherto, only 2 published studies have attempted to relate genomic or transcriptomic analyses directly to vaccine efficacy. The first described an association between variations in MAL/TIRAP and interleukin 10 with Hib vaccine failure, showing individuals homozygote for the MAL/TIRAP Ser180Leu polymorphism to be 5.6 times more at risk of invasive, nonmeningitis Hib disease [18]. Vaccine failures may be a particularly informative cohort to study in great depth, as it is logical to believe that individuals with “extreme phenotypes” also harbor genetic variants with larger effects [36]. A second study described gene expression profiles after the novel RTS,S malaria vaccine that associated with protection following infectious challenge [31]. This study showed that peripheral blood from protected individuals 14 days postvaccination was enriched for genes in the proteasome degradation pathway, involved in the efficient processing of the MHC peptides [31].

---

**Table 3. Summary of Preclinical Studies Utilizing Genome-wide Transcriptomic Analysis to Assess Candidate Vaccines**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Assessment</th>
<th>Study Details Summary</th>
<th>Summary of Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>[33]</td>
<td>Mouse (BALB/c)</td>
<td>Intramuscular vaccine adjuvant: MF59 oil-in-water emulsion, CpG or Alum</td>
<td>Mice 6–8 wk of age, 3 mice per vaccine and time point; Sampling 3, 6, 12, 24, 48, and 96 h postvaccination; Whole homogenized quadriceps muscle; Agilent 44K Whole Mouse Genome Microarray</td>
<td>Core set of 168 induced by all 3 adjuvants, characterized by the upregulation of cytokine, chemokine, and adhesion molecule genes; Most early innate genes were significantly regulated at the 6 h time point; Genes involved in antigen processing and presentation induced 6–12 h after MF59 and CpG, whereas 1–2 d after alum; MF59 was the most potent activator of gene expression, inducing nearly 3 times as many genes as the other adjuvants</td>
</tr>
<tr>
<td>[34]</td>
<td>Mouse (BALB/c)</td>
<td>Intramuscular influenza vaccine (inactivated) and tetanus toxoid vaccine alone, or adjuvanted with MF59, CpG, R848, or Pam3CSK4</td>
<td>Mice 6–8 wk of age, 4 mice per vaccine group; Sampling 6 h post-vaccination; Whole homogenized quadriceps muscle and draining lymph nodes; Agilent 44K Whole Mouse Genome Microarray</td>
<td>1410 genes differentially expressed in muscle following 1 or more of these interventions; MF59 was the strongest modulator of transcriptional events at the injection site, whereas R848 was strongest modulator of gene expression in the draining lymph nodes</td>
</tr>
<tr>
<td>[35]</td>
<td>Mouse (BALB/c or C57BL/6)</td>
<td>Intraperitoneally vaccine adjuvant: MPL liposomes or Alum</td>
<td>Mice &lt;6 mo of age, 6 mice per vaccine group; Sampling 26 h postvaccination; Peritoneal exudate cells; Affymetrix GeneChip mouse genome 430A 2.0 arrays</td>
<td>Both adjuvants upregulated a common set of 415 genes; MPL liposomes upregulated more than twice as many genes as alum; MPLs result in the induction of a wider range of NK and T-cell-attracting CXC chemokine ligands, which appeared to potentiate a greater influx of neutrophils, monocytes/macrophages, and activated NK cells</td>
</tr>
</tbody>
</table>

Abbreviations: Alum, aluminium hydroxide; CpG, nonmethylated CpG oligonucleotide; MPL, monophosphoryl lipid A; NK, natural killer.
Although this finding certainly requires further functional and mechanistic investigations, it highlights the potential use of transcriptomic analyses to elucidate novel correlates of protection.

CURRENT GENOMIC LIMITATIONS

Genetic association analyses are complicated by the vastness of genomic variation at the population level, with approximately 4 million bases of the human genome varying between 2 randomly chosen individuals [37]. This complexity is further compounded by the possibility of multiple variants having only modest individual but considerable synergistic and/or epistatic influences. Alternatively, rare variants with strong effect sizes may be critical, but these are not well covered by current DNA microarrays. Moreover, the heritability seen in vaccine responses may not be entirely accounted for by variation in the genomic DNA sequence. Empirical observations have shown the heritability of several complex traits to be more dynamic and context dependent than would be expected by the transmission of stable DNA sequence, which has led to speculation that heritable epigenetic modifications may also be important [38].

Technical and particularly biological replication is necessary to validate transcriptomic datasets, as these data are especially vulnerable to artefact. Peripheral blood presents an easily accessible source of circulating immune cells and is consequently the most utilized tissue in vaccine-related transcriptomic studies. However, a multitude of temporal and physiological factors, as well as technical aspects of sampling, can affect transcript measurement [39]. Sample preparation and cellular composition can strongly influence transcript levels; for example, whole blood differs drastically from PBMC samples [39]. Whole blood samples may decrease the detection sensitivity of transcripts originating from cells other than reticulocytes; conversely, the manipulation of blood samples to ascertain PBMCs induces artefactual changes indicative of prolonged handling [39]. Furthermore, a multitude of genetic and environmental factors may impact on both gene expression and vaccine responses; therefore, it is likely to be advantageous to assess populations uniform for age, sex, ethnicity, and vaccination history. Pathway-based methodologies, using a priori knowledge to aid in the analyses, have also been highlighted as less prone to spurious results [40].

To date, innate responses to vaccination have been best characterized at the transcriptional level, as these responses correspond to large fold-changes in gene expression in the first hours and days postvaccination [26, 27, 29, 31]. On the other hand, adaptive immune responses may result in broad but subtler global changes in the peripheral blood transcriptome that have not yet been well characterized [29]. Further investigations are necessary to describe genes involved in the adaptive immune response to vaccination, although these may require large numbers of participants and/or the enrichment of antigen-specific cells to increase power to detect more subtle differential expression.

Microarrays have become the routine platform for the assessment of transcriptomic changes following vaccination. However, microarrays have a number of intrinsic limitations, including constraint to known coding genes, cross-hybridization of similar sequences, variable hybridization of polymorphic sequences, and polymerase chain reaction amplification bias as well as a limited dynamic range due to high backgrounds and signal saturation.

GENOMIC FORECAST

Ever-declining costs, along with rapidly advancing analytical capabilities, mean that next-generation sequencing technologies will become the mainstay of future genomic studies into vaccine responses. Whole-genome sequencing will enable the capture of common and rare variants, facilitating the assessment of their relative contribution on vaccine responses. Large sample sizes will be needed to power single rare variant analysis in “healthy” vaccines. However, until sequencing costs become inexpensive enough, individuals at one or both ends of the phenotypical distribution are likely to be targeted. In vaccine studies, these extremes may include vaccine failure or persistent nonresponders; these individuals may be enriched for functional rare variants with large effect sizes, such that sufficient observation will be made even with modest sample sizes.

RNA sequencing will also transform the characterization and quantification of the transcriptomic changes induced by vaccination. Next-generation sequencing has several advantages over microarrays, such as the identification of transcripts not previously annotated and a superior dynamic range. Furthermore, sequencing enables the assessment of alternative splice variants, allelic expression, noncoding RNAs, and small RNA species.

CONCLUSIONS

Variations at both the genomic and transcriptomic level have been shown to correlate with immunological measures of vaccine responsiveness. Genomic analyses have been shown to be highly amenable to all stages of vaccine development, from the preclinical models through to the evaluation of licensed vaccines. Understanding vaccinee factors that influence immune response has vast translational implications, and may ultimately lead to directed and rational development of new and more efficacious vaccines and vaccine adjuvants with better immunogenicity and safety profiles.

Notes

Financial support. D. O. received funding from the European Union’s seventh Framework program under EC-GA no. 279185 (EUCLIDS) during the production of this manuscript. Potential conflicts of interest. A. J. P. conducts clinical trials of vaccines for University of Oxford, which receives grant funding from vaccine manufacturers.
References


