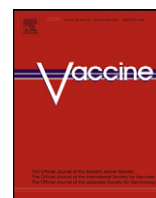




Contents lists available at [SciVerse ScienceDirect](#)

Vaccine

journal homepage: www.elsevier.com/locate/vaccine



A genomics-based approach to assessment of vaccine safety and immunogenicity in children

Olivia J. White^{a,1}, Katherine L. McKenna^{a,1}, Anthony Bosco^a, Anita H.J van den Biggelaar^a, Peter Richmond^b, Patrick G. Holt^{a,*}

^a Telethon Institute for Child Health Research, Centre for Child Health Research, The University of Western Australia, Perth, Australia

^b School of Paediatrics and Child Health, University of Western Australia, Princess Margaret Hospital, Western Australia, Australia

ARTICLE INFO

Article history:

Received 15 September 2011

Received in revised form

10 November 2011

Accepted 26 December 2011

Available online xxx

Keywords:

Immunisation

Infants

Safety

Cellular immunity

Atopy

Genomics

ABSTRACT

Immune responses to vaccines in infants and young children are typically Th2-biased, giving rise to concerns regarding potential atopy-like side effects, and antagonism of Th1-associated sterilising immunity. Conventional immunological methodology has limited capacity to effectively address these problems because of the inherent complexity of the immune responses involved. In the present study, we sought to develop an unbiased systems biology approach to elucidate superficially similar Th2-associated responses to paediatric vaccines and allergens, and to differentiate between them *via* gene coexpression network analysis. We demonstrate below that in immune responses to the diphtheria/acellular pertussis/tetanus and pneumococcal polysaccharide conjugate vaccines, potentially antagonistic Th1-/IFN-associated and Th2-associated gene networks coexist in an apparent state of dynamic equilibrium, whereas in Th2-dominant allergen-specific responses of atopics the Th1 and IFN networks are respectively disrupted and downregulated. Capacity to detect and interpret these covert differences between responses to vaccines and allergens relies on the use of sophisticated algorithms that underpin coexpression network analysis, which identify genes that function co-ordinately in complex pathways. This methodology has significant potential to identify covert interactions between inflammatory pathways triggered by vaccination, and as such may be a useful tool in prediction of vaccine safety/efficacy.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Early postnatal life is the time of greatest risk for morbidity and mortality resulting from infectious diseases, due principally to the functionally immature state of the immune system during this period which includes an overall Th2 “bias” relative to that seen at later ages [1]. There is broad agreement that the eventual solution to this problem involves the earliest possible use of an ever-widening range of vaccines during early life, increasingly in multivalent form to minimise the number of vaccinations required to provide optimal coverage. This is placing increasing pressure on vaccine producers and researchers alike to accelerate the pace of vaccine development and testing, but without compromising safety. This is a significant issue because the technology to create highly immunogenic vaccine preparations is rapidly

outstripping capacity to objectively assess their effects using conventional methodology. The problem will grow over time because future infant vaccine development is predicted to move towards the use of adjuvants for the selective stimulation of aspects of the immune response that are developmentally compromised in the target age group, and this will in turn increase the need for accurate immunological profiling prior to release.

Immunological profiling of vaccines in human infants has traditionally relied principally on measurement of humoral immune responses, and despite a consensus that vaccine side effects are more likely to involve aberrant cellular immune responses (CMI) there are no statutory requirements for provision of CMI data as part of either safety or immunogenicity testing. Nevertheless, CMI assessment is increasingly being used within the vaccine research community [2–7], in particular to address mechanistic issues related to potential vaccine side effects. However the laborious nature of conventional *in vitro* immunological assay systems, including those based on multiplex approaches, severely limits the number of parameters that can be practically assessed in such studies; this is not ideal given the necessity to measure a range of targets exemplary of different aspects of the vaccine recipients’ adaptive responses in parallel, if a comprehensive picture of safety/immunogenicity is to be achieved. Moreover the initial

* Corresponding author at: Division of Cell Biology, Telethon Institute for Child Health Research, PO Box 855, West Perth, WA 6872, Australia. Tel.: +61 8 9489 7838; fax: +61 8 9489 7707.

E-mail addresses: mstubbbs@icmr.uwa.edu.au, patrick@icmr.uwa.edu.au (P.G. Holt).

¹ Authors contributed equally to this work.

choice of immunological targets to be measured immediately biases the overall assessment and in particular limits possibilities for uncovering novel/unexpected vaccine effects (positive or negative) that may fall outside current paradigms.

In the present study we assess an alternative approach which addresses these limitations. Our starting point are recent findings which indicate that specific immune responses involve the coordinate activation of not dozens but hundreds–thousands of genes, which can be detected with high precision and in an unbiased fashion by genome wide expression profiling of mRNA from T-memory cells stimulated with appropriate antigens. It is recognised that the genes that are differentially regulated in these responses function within highly reproducible “networks” which can be detected within expression data sets *via* the use of appropriate algorithms [8], and the functionality of these networks can be deduced *via* the attendant use of appropriate bioinformatics tools [9–11]. Employing these approaches we have recently mapped the Th2-associated gene network that underpins the Th-memory response of human atopics to allergen and in doing so discovered a range of novel genes not previously associated with the allergic phenotype [12,13]. We have now applied these techniques to profiling of Th-memory generated in response to acellular pertussis (DTaP) and pneumococcal polysaccharide (PCV) vaccines, both of which are known to induce strong Th2- in addition to Th1-polarized immunity [4,6,14–17] and in the case of DTaP is associated with substantial production of Th2-dependent IgE against antigens in the vaccine [4,16]. The apparent close similarities between responses to certain vaccines and those to allergens have fuelled safety concerns which remain unresolved. We present new findings below based on genome wide profiling of DTaP and PCV vaccine responses in parallel with responses of atopic children to house dust mite allergen. We confirm the participation within the anti-vaccine responses of the full spectrum of Th2-associated genes which are networked comparably to what is observed in atopic responses. However we demonstrate that the Th2 component of these anti-vaccine responses is counterbalanced by parallel networks encoding pathways associated with the expression of Th2-antagonistic and antimicrobial immunity which are either not present or actively downregulated within the Th2-polarised responses associated with allergic inflammation.

2. Materials and methods

2.1. Study populations

Three paediatric populations were utilized for studies in this paper which were approved by our institutional human ethics committee, and informed consent was obtained from the parents and from teenage volunteers. The first study population involved 24 normal subjects that were recruited from the community to a study of DTaP vaccination conducted at the Vaccine Trials Group (Perth, Western Australia) at age 4–6 years when presenting for a 4th dose of DTaP-IPV. These subjects had received 3 doses of DTaP-IPV (GSK) in infancy as per the Australian National Immunisation Program [18]. Samples were collected immediately prior to the administration of the scheduled DTaP-IPV booster at 4–6 years. The second study population comprised 24 healthy 14 year-olds from an unselected community cohort who were identified as sensitized to house dust mite allergen (HDM) on the basis of a positive skin test reaction to HDM extract (wheal ≥ 3 mm) and the presence of HDM-specific serum IgE (≥ 0.7 kU/L). The third study involved assessment of specific T-memory responses at 9 months of age to a 7-valent PCV (Prevenar; Pfizer) in 50 infants from Papua New Guinea. The primary data from this study and details regarding immunisation schedule were recently published by our group [17], and a microarray data set from this study (available in the Gene Expression Omnibus repository [www.ncbi.nlm.nih.gov/projects/geo/] under

the accession number GSE25263) was used for independent validation of findings relating to the pertussis vaccine.

2.2. Cell culture

For vaccine studies cryobanked PBMC from 24 subjects were thawed and cultured in RPMI 1640 supplemented with 5% heat-inactivated human AB serum (DynaL Biotech) in the presence or absence of acellular pertussis vaccine antigen (aPVA, a combination of 1 μ g/ml of each of pertussis toxoid [PT], filamentous haemagglutinin [FHA], and pertactin [PRN], all generously supplied by GSK, Rixensart, Belgium). For atopy-related studies PBMC from 24 sensitized children were cultured in AIM-V medium (Invitrogen) in the presence or absence of 10 μ g/ml HDM extract (*Dermatophagoides pteronyssinus*: CSL). A culture period of 24 h was employed to minimise changes in gene expression profiles associated with *in vitro* differentiation which can occur during more prolonged culture; previous studies indicate that this is an appropriate time point to capture both early and late phase genes associated with Th-memory activation [12]. CD4⁺ T cells from antigen stimulated cultures were enriched to ~94% homogeneity by positive selection employing immunomagnetic bead separation (DynaL) prior to mRNA analyses (see below).

2.3. Microarray methodologies and data pre-processing

Total RNA was extracted from CD4⁺ T cells using TRIzol (Invitrogen) followed by RNeasy (Qiagen). Approximately 100 ng total RNA was labelled and hybridized to Human Gene 1.0 ST microarrays (Affymetrix), using standardised protocols and reagents from Affymetrix. Individual arrays were used for unstimulated and stimulated samples from each subject. The microarray data were pre-processed in Expression Console software (Affymetrix) employing the probe logarithmic intensity error algorithm (PLIER, parameters: PM-GCBG background subtraction, quantile normalisation, iterPLIER summarization). These data were further analysed in the R language and environment for statistical computing (www.r-project.org). Variance stabilization was performed by adding the small constant 16 to all the data points, followed by $\log_2$ transformation.

2.4. Identification of pertussis and HDM response genes

Genes differentially expressed between aPVA-stimulated cells and unstimulated control cells, and also between HDM-stimulated cells and unstimulated control cells, were identified employing the moderated paired *t*-statistic [19]. This analysis was performed using the Bioconductor package “limma” (Linear Models for Microarray Analysis) in which a linear model is fit to the data and empirical Bayes used to compute a moderated *t*-statistic. This method moderates the standard errors of the estimated log-fold changes of gene expression towards the pooled variance across all genes, resulting in more stable inference and improved power, especially for experiments with small numbers of microarrays [19]. To account for multiple testing, the *p* values derived from the moderated paired *t*-statistics were adjusted employing the false discovery rate (FDR) method [20]. HDM-responsive genes were defined as genes modulated by >1.2-fold after HDM stimulation and with an FDR-adjusted *q* value <0.05. Stimulation with pertussis antigen generated a weaker response and consequently genes were considered pertussis-responsive with an average fold change > 1.2 and/or FDR-adjusted *q* value <0.05. The fold change and test statistic results were graphically represented by volcano plot [21], a scatter plot of the change in gene expression between stimulated and unstimulated samples (*x*-axis; $\log_2$ fold change)

verses the significance of differential expression (y-axis; negative $\log_{10}$ -transformed p values).

2.5. Weighted gene coexpression network analysis

Weighted gene co-expression network analysis (WGCNA; [8]) was performed on stimulation-responsive genes. The correlation matrix was obtained by calculating the absolute Pearson correlations for all pair-wise gene expression values across microarray samples. This matrix was transformed into a weighted adjacency (or connectivity) matrix by raising the absolute value of each correlation to a power β . This soft thresholding approach emphasizes high correlations at the expense of low correlations on an exponential scale. The power β was chosen using the scale-free criterion of Zhang and Horvath [8]. The topological overlap (TO) matrix, which takes into account third party genes, was then generated and average linkage hierarchical clustering was performed using the topological overlap dissimilarity measure ($1 - TO$). A dynamic tree cutting algorithm, which cuts branches according to shape [22], was used to define modules. Intramodular connectivity (k_{in}) was calculated by summing the connection strengths for each gene with all other genes within a module [23]. The k_{in} values were scaled to lie between 0 and 1. Module eigengenes can be used to summarize the expression profiles of the module genes, and were calculated using the first principal component of the expression values across microarray samples for each module [24]. Another measure of connectivity is the module eigengene-based connectivity (k_{ME}), which is determined by correlating the expression profile of a gene with the module eigengene [25]. The module eigengene-based connectivity measure highly correlates with intramodular connectivity [23], however, k_{ME} can be used to calculate the correlation of the module eigengene to any gene on the microarray. A pairwise Fisher exact test was used to determine significant overlap between modules across networks [26].

2.6. Gene ontology analysis

The module gene lists were analysed with EASE (Expression Analysis Systematic Explorer) software implemented in DAVID (Database for Annotation, Visualization and Integrated Discovery; <http://david.abcc.ncifcrf.gov/tools.jsp>) [9,11] to functionally categorize genes. The data were analysed for themes using the gene ontology terms molecular function and biological process, as well as Biocarta and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The most significantly over-represented terms were identified using EASE scores, a modification of the Fisher exact test which calculates a p value for the extent of enrichment of a given functional term. Terms with an EASE score of less than 0.05 were considered significant.

The module gene lists were also analysed through the use of Ingenuity Pathways Analysis (IPA; Ingenuity Systems, www.ingenuity.com). Genes were categorized based on biological functions and a right-tailed Fisher's exact test, followed by Benjamini–Hochberg multiple testing correction was performed to determine the significance of any over-represented functions.

2.7. Module visualization

Modules were graphically depicted using the program VisANT [10]. Module depictions represent only the top 350 within-module gene–gene connections with the strongest topological overlap. Genes with more than 28 connections were considered “hub” genes and enlarged in the depiction. The “elegant relaxing” layout algorithm was used to visualize the module structures, which were then manually adjusted for clarity.

3. Results

3.1. Mixed Th1/Th2 memory responses to pertussis in DTaP-vaccinated children

Fig. 1 illustrates the gene expression profile of the CD4⁺ T-cell memory response to pertussis in preschool children vaccinated during infancy. To identify pertussis-responsive genes, the expression intensity in antigen-stimulated cells and unstimulated cells was compared using a moderated paired t -statistic [19]. The results are represented as a volcano plot (Fig. 1), a scatter plot depicting the fold change (x-axis) and the corresponding significance of these observations (y-axis) for each probe. Each dot within the plot designates a single gene, and the location of a subset of relevant upregulated genes of interest is illustrated in the upper right quadrant. As shown these include the classical Th1 genes (IFNG, IL12A, IL12B) and Th2 genes (IL4R, IL5, IL13) (average fold change > 1.2 and/or FDR-adjusted q value < 0.05; moderated paired t -statistic). In addition, many novel genes previously reported by our group to be upregulated in a Th2 atopic response were also upregulated following pertussis antigen stimulation, including DACT1, MAL, NDFIP2, RAB27B, DPP4 and CISH [12,13]. The mixed Th1/Th2 nature of this response was confirmed at the protein level employing culture supernatants collected at a later time point (Supplementary Fig. 1), and this pattern is consistent with that observed in responses to the PCV vaccine [17] which is used below for cross-validation.

3.2. Use of network analysis for identification of Th1- and Th2-associated gene modules in the pertussis response

To further characterise the pertussis response, we performed network analysis which provides a systems level understanding of the response based on the correlations in gene expression patterns. Genes which are highly correlated may represent sets of co-ordinately regulated genes involved in common biological pathways (reviewed in [27]).

A co-expression network was constructed employing the WGCNA algorithm using the 1572 pertussis-responsive genes (average absolute fold change > 1.2 and/or FDR adjusted q value < 0.05; moderated paired t -statistic) identified in pertussis-stimulated CD4⁺ T cells from vaccinated subjects (average absolute fold change > 1.2 and/or FDR adjusted q value < 0.05; moderated paired t -statistic). The scale-free criterion described by Zhang and Horvath [8] was used to choose the power ($\beta = 7$), resulting in a 1572×1572 weighted adjacency matrix containing pairwise connection strengths. Genes were clustered based on the topological overlap measure, and those with similar expression patterns were grouped together to form a network which contained 12 gene co-expression modules.

To determine how related the modules are, principle component analysis (PCA) was performed. PCA reduces the dimensionality of the data while retaining as much of the variability as possible [24]. A representative module expression profile, referred to as the module eigengene (ME), was determined for each module by taking the first principle component of the expression values across samples. The module eigengene represents the main trend of gene expression within the module, and modules with genes sharing similar expression profiles will also have similar MEs. Average linkage hierarchical clustering of the module eigengenes identified modules which cluster together and are therefore likely to be functionally related [28] (Supplementary Fig. 2). The focus of this analysis was Th1- and Th2-associated immune responses to the vaccine, and hence we next sought to identify the network modules in which the majority of Th1- and Th2-associated genes were expressed. To achieve this, each module of the pertussis network was biologically categorized by performing functional enrichment analysis

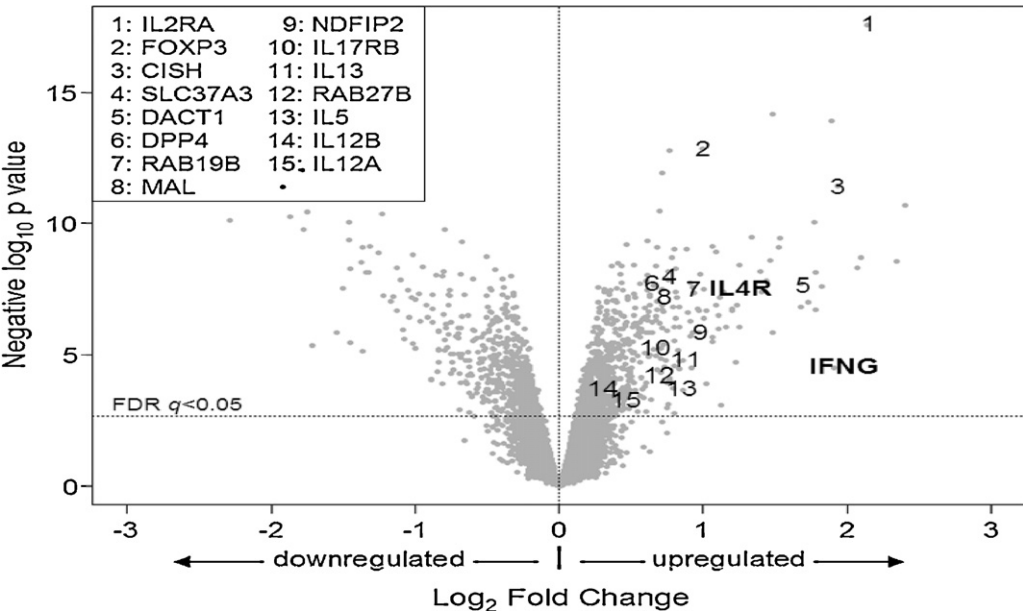


Fig. 1. Microarray analysis of CD4⁺ T cell responses to acellular pertussis vaccine antigen at the gene expression level. Volcano plot of the pertussis response, with differential expression magnitude plotted on the x-axis (stimulated/unstimulated on the log₂ scale (log₂ fold change)) and the significance of differential expression (negative log₁₀-transformed p values) plotted on the y-axis (n = 24 subjects). Each dot represents a gene. Numbers represent genes indicated in the legend. All genes above the horizontal line have an FDR-adjusted q value < 0.05. Positive and negative values on the x-axis indicate upregulated and downregulated expression, respectively.

using EASE (Expression Analysis Systematic Explorer; Table 1). This revealed a subcluster of three modules (circled in Supplementary Fig. 2) which on the basis of their localization at the lowest level within the dendrogram represent the most strongly linked modules within the overall pertussis network, and which contained the bulk of the Th1/Th2-associated genes of interest. This subcluster comprised module F which is enriched for Th2 genes including IL4R, IL5, IL13, LIF and CISH (full list in Supplementary Table I), module G enriched for Th1-differentiation-associated genes including IFNG, IL12A, IL12B, IL12RB1, IL12RB2 and IRF1 (Supplementary Table II), and module H which is enriched for genes downstream of Types I and III IFNs (Supplementary Table III) that are associated with antimicrobial immunity [29].

3.3. Comparison to atopy-associated Th2 responses

The presence of a strong Th2 signature in the pertussis-specific vaccine response raises questions of potential similarity between the latter and potentially pathogenic Th2 responses to allergens. We accordingly sought to directly compare these two responses

employing the same methodology. To analyse atopy-associated Th2 responses we followed the protocol developed in our earlier studies on reactivity to HDM allergen focusing upon the response profile of CD4⁺ T-cells at 24 h [12,13]. HDM-responsive genes were identified in a panel of 24 sensitized children by moderated paired t-statistic (average absolute fold change > 1.2 and FDR-adjusted q value < 0.05). Overlap between the atopic response and the pertussis vaccine response is shown in the venn diagram in Supplementary Fig. 3. A total of 790 genes were common to both responses including the core Th2 genes IL4R, IL5, IL13, IL17RB, DACT1, MAL, NDFIP2, RAB27B, DPP4 and CISH. The core Th1 genes IFNG, IL12A, IL12B, IL12RB1 and STAT1 were restricted to the panel of 782 genes that were unique to the pertussis response. The presence of IL-12A/IL-12B in module G suggests that the enriched CD4⁺ T-cell preparations used here contain contaminating dendritic cells, likely resulting from the activation and ensuing clustering of these cell types in the initial antigen-stimulated PBMC cultures; it should be noted that the IL-12 signal was not validated by qPCR.

An atopy coexpression network was constructed in a similar manner to the pertussis network using the 24 atopic subjects

Table 1
Functional annotation associated with the modules from the pertussis network.

Module	Ontology term ^a	EASE score ^b	Direction of expression ^c
A	TCR signalling	2.97 × 10 ⁻⁴	↓
B	Antigen processing and presentation	2.71 × 10 ⁻⁶	↓
C	Defense response	1.23 × 10 ⁻¹⁶	↓
D	ATP binding	1.24 × 10 ⁻³	↑
E	Positive regulation of transcription	1.42 × 10 ⁻³	↑
F	Jak-STAT signalling pathway	1.52 × 10⁻¹²	↑
G	T cell differentiation, Th1 differentiation	5.80 × 10⁻⁵	↑
H	Response to virus, RIG-I-like receptor signalling pathway	5.17 × 10⁻⁹	↑
I	Regulation of cytokine production	1.88 × 10 ⁻⁷	↑
J	Cytokine activity, chemokine activity	1.33 × 10 ⁻²⁷	↑
K	Nucleoside monophosphate catabolic process	2.37 × 10 ⁻²	↑
L	Cell activation	3.97 × 10 ⁻⁴	↓

^a EASE categories include GO, KEGG and BIOCARTEA categories. The most characteristic category for each module is shown with its corresponding EASE score. In some cases the second highest term is also shown.
^b The EASE score is a modified Fisher exact probability p value; the smaller the EASE score the more significant the ontology term.
^c The overall direction of expression of genes from each module following antigen stimulation is indicated.

and the 1450 HDM-responsive genes detected (average absolute fold change > 1.2 and FDR adjusted q value < 0.05; moderated paired t -statistic). Soft thresholding ($\beta = 8$) generated a weighted 1450×1450 adjacency matrix and WGCNA identified 10 gene co-expression modules. To determine whether any modules were preserved across the pertussis and atopy networks, the genes in each module from the two networks were compared. Strikingly, two modules in the atopic response displayed significantly more genes in common with corresponding pertussis modules than expected by chance. These comprised a “Th2” module in which 49% of the genes overlapped with the pertussis module F (Fisher t -test $p = 5.09\text{e}^{-21}$), and a corresponding “antimicrobial” module overlapping to a similar degree with pertussis module H (Fisher t -test $p = 1.20\text{e}^{-22}$). Furthermore, by EASE analysis these two module pairs contained predicted overlapping gene ontology terms, notably JAK-STAT signalling and response to virus and RIG-I-like receptor signalling pathway.

As noted in [Supplementary Fig. 3](#), both the pertussis and atopy responses each contained many genes unique to their respective responses. Interestingly, many of the genes unique to the pertussis response formed a distinct module (module G; [Supplementary Table II](#)) within the pertussis network, suggesting that this module is uniquely expressed and co-regulated in this response (Fisher t -test $p = 1.92\text{e}^{-20}$). As shown in [Table 1](#), functional categorization of module G suggests it is involved in Th1 differentiation.

The three modules of interest (F, G, H) in the pertussis network were further characterised using the more powerful IPA system. Cellular growth and proliferation was the most significant category for the module F with p -values ranging between 8.18e^{-13} and 6.42e^{-04} . Subcategories included proliferation of lymphocytes ($p = 8.18\text{e}^{-13}$), growth of T lymphocytes ($p = 8.37\text{e}^{-08}$), as well as other subcategories such as differentiation of T lymphocytes ($p = 1.86\text{e}^{-06}$) and Th2 immune response ($p = 8.09\text{e}^{-06}$). For module G, cell-mediated immune response was a highly significant category with p -values ranging between 1.14e^{-05} and 2.86e^{-02} . This category contained several subcategories, including differentiation of Th1 cells ($p = 1.14\text{e}^{-05}$), polarization of Th1 cells ($p = 1.61\text{e}^{-05}$) and development of helper T lymphocytes ($p = 1.61\text{e}^{-05}$). The top most significant functional categories for module H included antimicrobial response ($p = 4.30\text{e}^{-09}$ to 5.49e^{-02}) and inflammatory response ($p = 4.30\text{e}^{-09}$ to 9.48e^{-02}). These results confirm the presence of a Th2 module, a Th1 module and an antimicrobial module of co-regulated genes within the pertussis response. Comparative analyses of the atopy response employing the IPA system revealed strong module preservation and overlapping ontology terms (not shown) indicating the operation of similar pathways related to Th2 differentiation and antimicrobial defense, whereas the coexpression module comprising genes involved in Th1 differentiation was unique to the pertussis response.

3.4. Expression of module genes in the pertussis and atopy responses

The expression profiles of the genes from each of these modules were then compared in the pertussis and atopy responses. In [Fig. 2](#), volcano plots are shown for the pertussis response (top panel) and the atopy response (bottom panel), and the genes from each of the respective modules of interest are highlighted in black. All of the genes from both the Th2 pertussis module and the Th2 atopy module are positioned on the right-hand side of respective volcano plots, designating upregulation ([Fig. 2](#), first panel). Similarly, all the genes from the Th1 pertussis module (black inverted triangles) are uniformly upregulated following pertussis stimulation. Interestingly, the genes from the antimicrobial module (black circles) are upregulated in the pertussis response, but the corresponding genes

are downregulated in response to HDM ([Fig. 2](#), last panel). As noted in [Fig. 2](#) no contiguous Th1 coexpression module was present within the atopic response. However, ~35 of the relevant genes including Th1 signature genes IL12B2 and IRF1 are nevertheless significantly upregulated within the atopic response (marked in [Supplementary Table 2](#)), but as illustrated in the topological overlap matrix plot in [Fig. 3](#), unlike the situation within the pertussis response, they are not coordinately expressed as a discrete module but instead are randomly distributed throughout the overall atopic network.

3.5. Validation studies in independent study populations

Independent validation of the expression patterns of the genes in the Th1, Th2 and antimicrobial modules was also obtained employing publicly available data pertaining to Th-memory responses to the PCV vaccine from a separate study population of infants, recently reported by our group [17]. Despite differences in methodology, age of the study groups and sample size, the expression patterns for the same set of genes within these PCV memory response modules (illustrated in [Supplementary Fig. 4](#)) closely mirror those shown for the pertussis vaccine response in [Fig. 2](#). In an additional study, confirmation of the striking inverse expression pattern observed for the antimicrobial module in the atopic response (i.e. downregulation as opposed to the upregulation observed in the vaccine responses) was obtained in a separate set of samples from HDM sensitized subjects, employing culture conditions identical to those used for the vaccine response analyses ([Supplementary Fig. 5](#)).

3.6. Conservation of Th2 modular architecture is in the pertussis and atopy responses

Genes with strong correlations (high connectivity) within a co-expression network module are known as “hub” genes, and are more likely to be vital in modular function than peripheral genes [30–32]. We identified the hubs within the pertussis Th2 module and compared them to the hubs in the atopy Th2 module.

The module position of each gene was determined by calculating the module eigengene-based connectivity (k_{ME}), which is the correlation between the expression profile of a gene and the module eigengene. These values are a measure of intramodular connectivity or module membership, and genes that highly correlate with the module eigengene are likely to be hub genes [23]. The correlation between the respective intramodular connectivity measures across all genes in the Th2 modules from both networks was highly significant ($r = 0.87$, $p < 2.2\text{e}^{-16}$; [Supplementary Fig. 6A](#)). Genes which were highly connected in the pertussis Th2 module were also highly connected in the atopic Th2 module, indicating the preservation of hubs in both modules. This is illustrated in [Fig. 4](#), which visualizes the Th2 modules from both responses by depicting the strongest gene–gene connections. The hubs are shown as enlarged nodes, surrounded by the peripheral genes of low connectivity. There were a higher proportion of genes with strong connectivity in the atopy Th2 module compared to the pertussis Th2 module which is demonstrated by the large number of hub genes in the atopy Th2 module ([Fig. 4A](#)). However the connectivity of these genes in the pertussis Th2 module is still relatively strong ([Supplementary Fig. 6A](#)). Interestingly, several genes appeared to be uniquely connected within the pertussis Th2 module only. These genes, which were also uniquely expressed in the pertussis response ([Supplementary Fig. 6B](#)) include MYB, CSF2 (GM-CSF) and IL-3, the latter being a hub in the pertussis Th2 module ([Fig. 4B](#)).

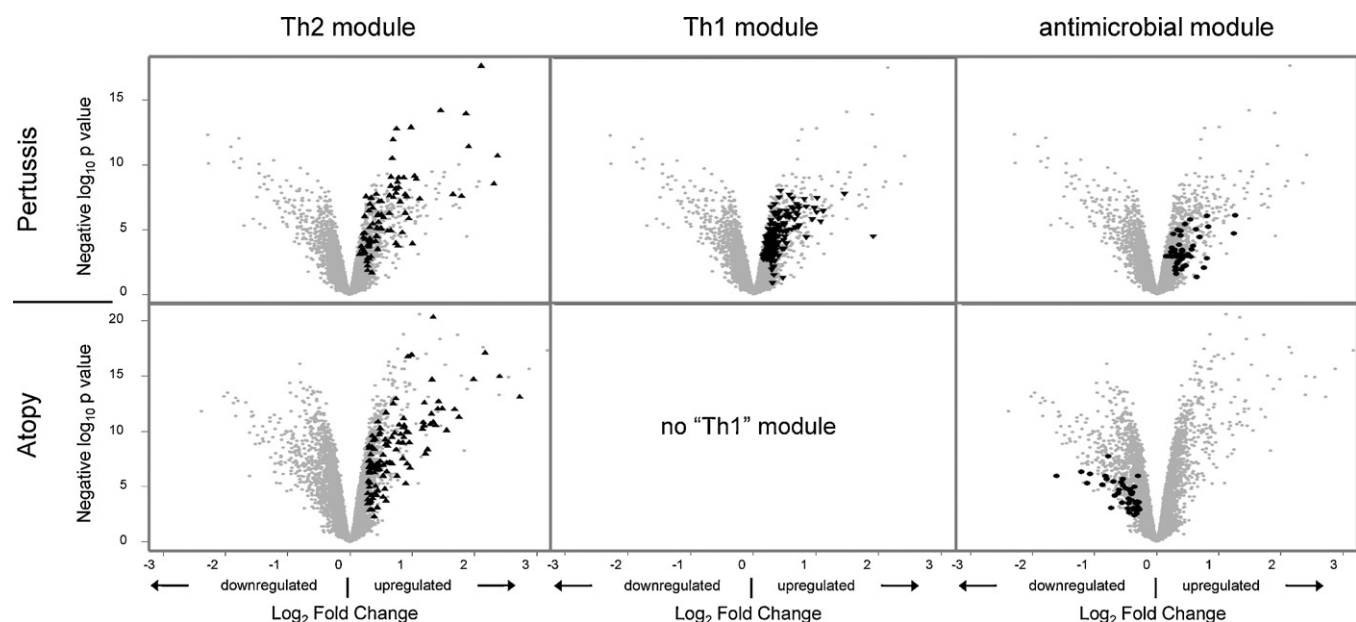


Fig. 2. Gene expression profiles of the modules in the pertussis and atopy networks. Volcano plots show the pertussis response (top panel), and the atopy response (bottom panel). Differential expression magnitude is plotted on the x-axis (stimulated/unstimulated on the log₂ scale (log₂ fold change)) and the significance of differential expression (negative log₁₀-transformed p values) plotted on the y-axis. Genes from the pertussis and atopy Th2 modules (triangles), Th1 modules (inverted triangles) and antimicrobial modules (circles) are highlighted in black. There is no Th1 module in the atopy response. Positive and negative values on the x-axis indicate upregulated and downregulated expression, respectively.

4. Discussion

The use of genome wide expression profiling to analyse vaccine responses in humans is in its infancy and few published studies are yet available. Lahdenpera et al. [33] measured the kinetics of gene expression from vaccinated infants following *in vitro* re-stimulation of PBMCs with *Bordetella pertussis* acellular vaccine antigens. However, there were several limitations to this study. cDNA arrays representing only 1176 genes were utilized, the sample size was small (*n*=5 subjects) and pooled, and the authors focused their analysis on immune response genes with association to asthma and allergy. More recently, a systems biology approach has been used to study the immune response following influenza vaccination. Bucacas and colleagues [34] measured ongoing global gene expression pattern in freshly prepared PBMC from 92 individuals

before, and at three timepoints after, influenza vaccination using Illumina arrays representing over 25,000 genes. These authors found maximum expression changes within the first 24 h following vaccination, which strongly correlated with the magnitude of the antibody response. The use of whole PBMC, however, may mask responses observed in individual cell subsets. Indeed, in a separate study [35] Nakaya demonstrated gene expression changes to influenza vaccination for various cell types, including B cells, monocytes and dendritic cells. Furthermore, this latter group identified a molecular gene signature to predict influenza vaccine immunogenicity. Other studies employing a systems biology approach have identified molecular gene signatures that predict the magnitude of adaptive responses in humans vaccinated against yellow fever [36] and protection against influenza-like illness in the first two weeks after vaccination with live attenuated influenza vaccine [37],

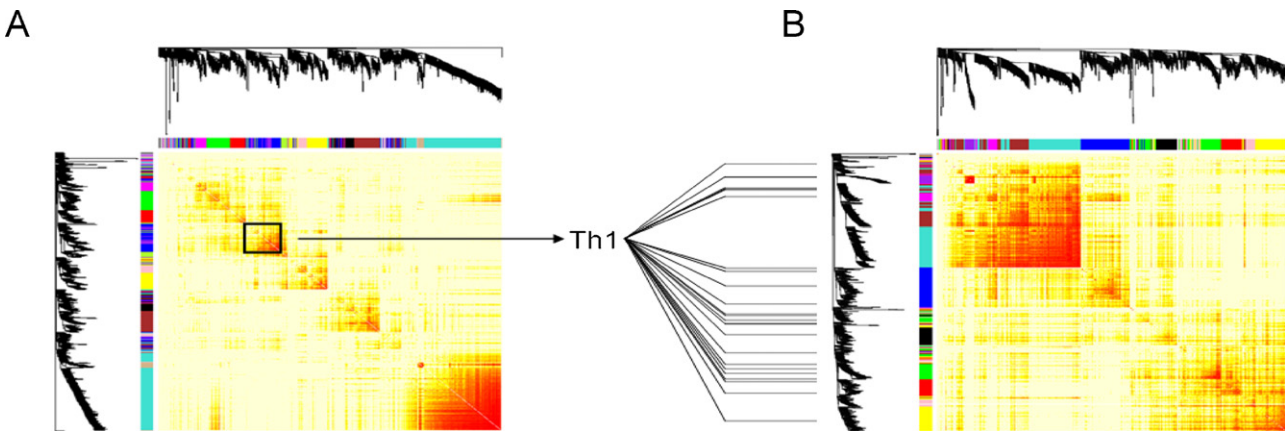


Fig. 3. Topological overlap matrix plots for the pertussis and atopy networks. (A) TOMplot for pertussis network. (B) TOMplot for atopy network. The TOMplot is a symmetric heatmap plot of gene connectivity within the network. Each row and column corresponds to all genes in the network, which have been sorted by hierarchical clustering into branches representing modules. Each module is assigned a colour displayed in the colour bars below and to the right of the dendrograms. In the heatmap, high connectivity is indicated by the progressively saturated colour. Modules correspond to squares on the diagonal. The Th1 module within the pertussis network is highlighted. Only ~35 genes from this module are expressed in the atopy network (see [Supplementary Table II](#)) and these genes do not form a tightly connected module, but are scattered across the network (black horizontal lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

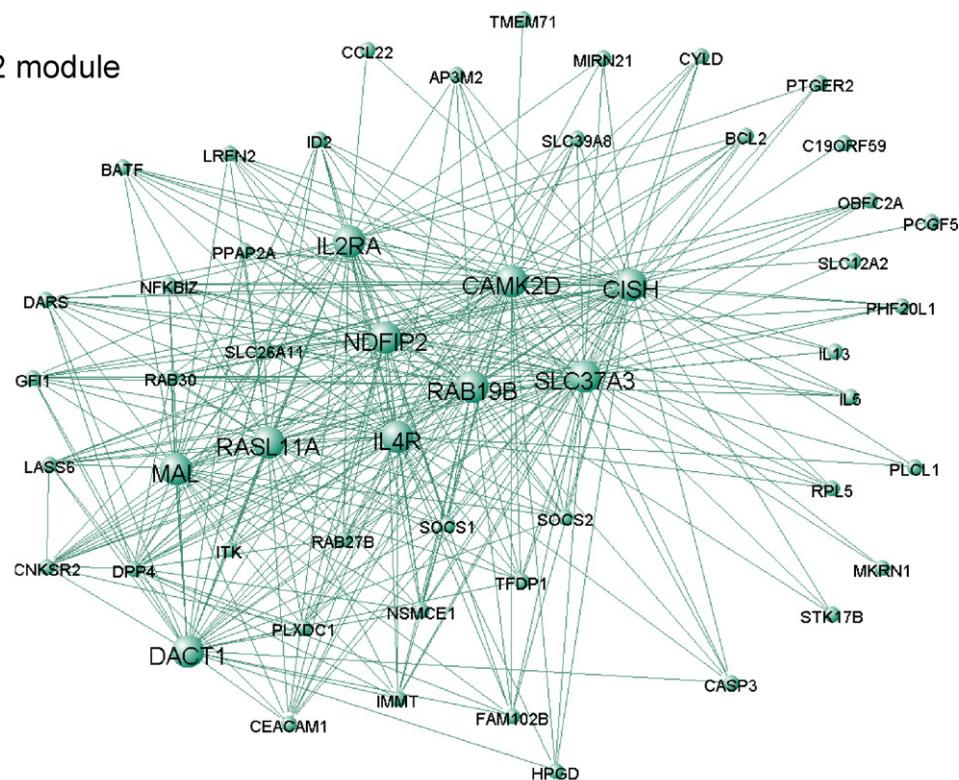
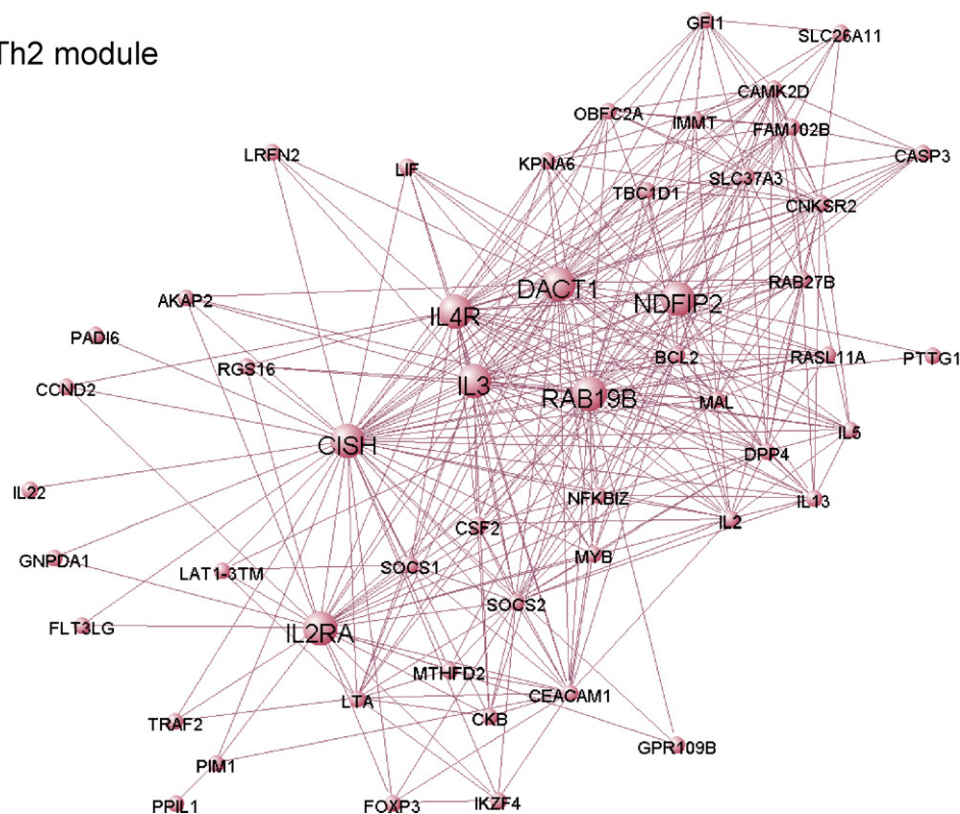
A: Atopy Th2 module**B: Pertussis Th2 module**

Fig. 4. Th2 module depictions from the atopy and pertussis networks allow visualization of intramodular connections and hub genes. Visualization was performed using VisANT to plot the top 350 positive gene–gene connections with the strongest topological overlap within the (A) atopy Th2 module, and (B) pertussis Th2 module. Enlarged nodes represent hub genes with more than 28 connections. The length of each line and node positioning was determined by selecting the “elegant relaxing” layout algorithm within VisANT to highlight module structure. Any duplicate or unannotated genes were removed.

highlighting the potential power of this type of analysis in determining underlying mechanisms of vaccine immunity. However, to our knowledge this is the first study to incorporate genome wide co-expression network analysis of a vaccine response, providing an additional level of insight (based on gene–gene correlations) into the modular architecture of the response.

Our aim in the present study was to evaluate the potential of this approach for assessment of vaccine safety/immunogenicity. The vaccine of initial interest in this test-of-concept was DTaP, which typically generates immunological memory with a prominent Th2 component [4,15,16,38] which in some situations has been associated with adverse reactions in recipients. This history underscores the desirability for immunological tools to predict such responses at an earlier stage of vaccine development, and this represents one of our long term aims.

In this study we have mapped gene networks expressed during reactivation of Th-memory to DTaP in CD4⁺ T-cells from preschool children vaccinated during infancy, and compared these to Th2-associated memory responses to HDM aeroallergen in sensitized children which is associated with allergic inflammatory disease [39]. Our salient findings are as follows:

- (i) the pertussis-specific memory response to DTaP vaccine involves over 1000 genes and includes major Th1- and Th2-associated components localized principally within three network co-expression modules; this finding is consistent with the mixed nature of the response to this vaccine in children as reported earlier [4,15,16,38];
- (ii) the Th2-associated genes in the pertussis response are essentially restricted to one of the network modules, and include a broad range of classical Th2 effector and regulatory genes as well as the bulk of the novel Th2-associated genes we have defined in recent studies on immunity to HDM allergen in atopics [12,13];
- (iii) the architecture of the pertussis-specific Th2 network module closely resembles that observed in the HDM response including concordance with respect to the majority of hyperconnected “hub” genes; the pertussis-specific module displayed three additional hyperconnected genes which have been previously associated with expression of Th2 immunity in other systems [40,41], notably IL-3, GM-CSF and the GATA3-associated transcription factor MYB;
- (iv) the pertussis-specific response exhibits two additional upregulated coexpression network modules which contain respectively the bulk of the classical Th1-associated genes, and another containing a broad range of genes downstream of the Type 1/111 IFNs (“antimicrobial” module) indicative of bystander activation of IFN-secreting cells in the initial PBMC cultures; direct immunodulatory effects of these IFNs on T-cell populations have been reported in several experimental systems [42,43];
- (v) the pattern of expression of the genes comprising these three distinct network modules within the memory response to DTaP vaccine in Caucasian Australian children was closely reproduced within the response to PCV vaccine in Melanesian children from Papua New Guinea who had been primed in infancy;
- (vi) it was noteworthy that the Th1 module was not present within the overall atopic response network, and moreover that the genes comprising the antimicrobial module are actively down-regulated within the atopic response.

Thus while the Th2 component of the memory responses primed via these paediatric infectious disease vaccines closely resembles the corresponding component of allergic responses, the lack of allergy-related side effects in association with their current usage

demonstrates that the presence of an active Th2 network module in vaccine-specific memory response profiles does not *per se* confer risk. It has been established in a range of studies in the clinical allergy literature (e.g. [39]) that one of the major determinants of risk for development of inflammatory reactions amongst sensitized children is the intensity of their Th2 memory responses to allergens, and a similar rule may apply in relation to Th2-associated anti-vaccine responses. An illustration of this can be seen in our earlier report on injection site side effects following the preschool DTaP booster in children vaccinated during the late 1990s under the original infant priming schedule of 2, 4, 6 and 18 months. Delayed local reactions (swelling/redness >50 mm diameter) were observed at that time in the order of 40% of children primed exclusively with DTaP [4] and these have since declined at least 4-fold following removal of the 18 month booster dose from the schedule. Analysis of vaccine-specific Th-memory responses in children primed under the original schedule demonstrated increased levels of production of IL-5, IL-6 and IL-13 in vaccine antigen-stimulated PBMC samples from subjects experiencing severe reactions [4], which supports the notion that quantitative differences in Th2-associated effector gene expression were relevant in this context. While such differences were readily detectable retrospectively in a case:control study design comparing reactors:non-reactors, it is unlikely that using this approach in isolation would have successfully predicted the potential for such response patterns in unstratified populations at an earlier stage in vaccine development.

The lack of evidence of such reactions amongst challenged vaccinees under the less intense priming schedule in current use infers that the Th2 component of their memory responses are endogenously regulated such that production of pro-inflammatory cytokines is maintained below the threshold level associated with clinically significant allergic inflammation. The present study demonstrates that broadening assessment of host anti-vaccine responses to the genome wide/gene network level can provide additional layers of information which help to understand the basis for such endogenous regulation. In particular, it is evident that in contrast to the Th2-dominant response of atopics to HDM allergen, the Th2 component of pertussis-specific memory is counterbalanced by the presence of two additional gene network modules associated with Th1-mediated and Types 1/3 IFN-mediated (antimicrobial) immunity. The tight clustering of these three network modules observed *via* PCA of the vaccine-specific memory response (Supplementary Fig. 2) indicates they are strongly associated *i.e.* there is a high likelihood that the component genes of the modules are coexpressed within the same microenvironments. In this regard, in addition to its known antagonistic effects on expression of Th1 immunity, the principle growth factor driving Th2 differentiation (IL-4) also negatively regulates multiple IFN-inducible genes within the antimicrobial module [44], and this may explain the key difference between the overall response profiles. Thus, in the case of the pathogenic Th2 response to allergen in which expression of Th2 genes exemplified by IL-4 reaches levels of intensity sufficient to induce allergic inflammation, the latter is also sufficient to suppress/dysregulate “bystander” Th1 and antimicrobial network modules. In contrast, within the pertussis response these mutually antagonistic gene networks appear to coexist in a state of dynamic equilibrium in which the intensity of the Th2 component remains below this critical threshold. The net result may be a “balanced” immune response in which Th2 cytokines necessary for optimal antibody responses coexist with Th1 response elements central to sterilising adaptive immunity and the broad range of IFN-inducible genes constituting the innate antimicrobial pathway.

The capacity to demonstrate these essential differences between two superficially very similar manifestations of Th2-associated immunity is a direct function of the systems level

analyses undertaken here. In contrast, quantitative expression profiling in isolation in the absence of information on constituent gene network modules operating within the respective immune responses, would be hard pressed to detect the relatively subtle differences involved on a gene-by-gene basis. It is not until the relevant genes within these responses are resolved into (function-associated) coexpression modules and then considered as a group, as exemplified by the antimicrobial module in Fig. 2, that these data become readily interpretable. The second illustrative example in this regard involves the constituent genes of the Th1 module. As noted in Fig. 2 no contiguous Th1 coexpression module could be detected within the atopy response, however a range of the constituent genes are upregulated but are distributed randomly throughout the overall network. This disrupted architecture may be an initial step towards active downregulation of the Th1 network as a whole, a pattern in keeping with the extreme Th2 dominance that is frequently observed in atopics.

In summary, this study illustrates the potential of a systems level approach to analysis of host responses to new vaccine entities. The use of genome wide profiling in conjunction with gene coexpression network analysis and allied bioinformatics provides novel opportunities for elucidation of vaccine-induced host defense mechanisms, and also for the early detection of covert interactions between vaccine-activated pathways that have potential for induction of pathogenic reactions in the host. Further pursuit of this approach may yield potent new screening tools for assessment of vaccine safety and efficacy.

Acknowledgement

This work was supported by the National Health and Medical Research Council, Australia.

Conflict of interest statement: The authors have no conflicts of interest to declare.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.vaccine.2011.12.118.

References

- [1] Siegrist CA. Neonatal and early life vaccinology. *Vaccine* 2001;19(May (25–26)):3331–46.
- [2] Guiso N, Njamkepo E, Vie le Sage F, Zepp F, Meyer CU, Abitbol V, et al. Long-term humoral and cell-mediated immunity after acellular pertussis vaccine compares favourably with whole-cell vaccines 6 years after booster vaccination in the second year of life. *Vaccine* 2007;25:1390–7.
- [3] Knuf M, Schmitt HJ, Jacquet JM, Collard A, Kieninger D, Meyer CU, et al. Booster vaccination after neonatal priming with acellular pertussis vaccine. *J Pediatr* 2010;165:675–8.
- [4] Rowe J, Yerkovich ST, Richmond P, Suriyaarachchi D, Fisher E, Feddema L, et al. Th2-associated local reactions to the acellular diphtheria-tetanus-pertussis vaccine in 4- to 6-year-old children. *Infect Immun* 2005;73(December (12)):8130–5.
- [5] Scheifele DW, Ochnio JJ, Halperin SA. Cellular immunity as a potential cause of local reactions to booster vaccination with diphtheria and tetanus toxoids and acellular pertussis antigens. *Pediatr Infect Dis J* 2009;28(November (11)):985–9.
- [6] White OJ, Rowe J, Richmond P, Marshall H, McIntyre P, Wood N, et al. Th2-polarisation of cellular immune memory to neonatal pertussis vaccination. *Vaccine* 2010;28(March (14)):2648–52.
- [7] Rosenkrands I, Vingsbo-Lundberg C, Bundgaard TJ, Lindenstrom T, Enouf V, van der Werf S, et al. Enhanced humoral and cell-mediated immune responses after immunization with trivalent influenza vaccine adjuvanted with cationic liposomes. *Vaccine* 2011;(June).
- [8] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005;4(1). Article 17.
- [9] Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4(5):P3.
- [10] Hu Z, Mellor J, Wu J, Yamada T, Holloway D, Delisi C. VisANT: data-integrating visual framework for biological networks and modules. *Nucleic Acids Res* 2005;33(July (Web Server issue)):W352–7.
- [11] Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, et al. Extracting biological meaning from large gene lists with DAVID. *Curr Protoc Bioinform* 2009;(September). Chapter 13: Unit 13.1.
- [12] Bosco A, McKenna KL, Devitt CJ, Firth MJ, Sly PD, Holt PG. Identification of novel Th2-associated genes in T memory responses to allergens. *J Immunol* 2006;176(April (8)):4766–77.
- [13] Bosco A, McKenna KL, Firth MJ, Sly PD, Holt PG. A network modeling approach to analysis of the Th2 memory responses underlying human atopic disease. *J Immunol* 2009;182:6011–21.
- [14] Ausiello CM, Urbani F, la Sala A, Lande R, Cassone A. Vaccine- and antigen-dependent type 1 and type 2 cytokine induction after primary vaccination of infants with whole-cell or acellular pertussis vaccines. *Infect Immun* 1997;65(June (6)):2168–74.
- [15] Rowe J, Macaubas C, Monger T, Holt BJ, Harvey J, Poolman JT, et al. Heterogeneity in diphtheria-tetanus-acellular pertussis vaccine-specific cellular immunity during infancy: relationship to variations in the kinetics of postnatal maturation of systemic th1 function. *J Infect Dis* 2001;184(July (1)):80–8.
- [16] Ryan EJ, Nilsson L, Kjellman N, Gotheffors L, Mills KH. Booster immunization of children with an acellular pertussis vaccine enhances Th2 cytokine production and serum IgE responses against pertussis toxin but not against common allergens. *Clin Exp Immunol* 2000;121(August (2)):193–200.
- [17] van den Biggelaar A, Pomat W, Bosco A, Phuanukoonnon S, Devitt C, Nadal-Sims M, et al. Pneumococcal conjugate vaccination at birth in a high-risk setting: no evidence for neonatal T-cell tolerance. *Vaccine* 2011;29:5414–20.
- [18] National Health and Medical Research Council. The Australian Immunisation Handbook. 9th ed Australian Government Publishing Service; 2008.
- [19] Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3(1). Article 3.
- [20] Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behaviour genetics research. *Behav Brain Res* 2001;125(1–2):279–84.
- [21] Cui X, Churchill GA. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol* 2003;4(4):210.
- [22] Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 2008;24(March (5)):719–20.
- [23] Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* 2008;4(8):e1000117.
- [24] Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 2000;97(August (18)):10101–6.
- [25] Fuller TF, Ghazalpour A, Aten JE, Drake TA, Lusis AJ, Horvath S. Weighted gene coexpression network analysis strategies applied to mouse weight. *Mammal Genome* 2007;18(July (6–7)):463–72.
- [26] Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? *PLoS Comput Biol* 2011;7(1):e1001057.
- [27] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 2004;5(February (2)):101–13.
- [28] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95(December (25)):14863–8.
- [29] Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, et al. A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* 2011;472(April (7344)):481–5.
- [30] Carlson MR, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF. Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks. *BMC Genomics* 2006;7:40.
- [31] Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci U S A* 2006;103(November (46)):17402–7.
- [32] Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, et al. Functional organization of the transcriptome in human brain. *Nat Neurosci* 2008;11(November (11)):1271–82.
- [33] Lahdenpera AI, Nilsson LJ, Regnstrom K. Kinetics of asthma- and allergy-association immune response gene expression in peripheral blood mononuclear cells from vaccinated infants after *in vitro* re-stimulation with vaccine antigen. *Vaccine* 2008;26:1725–30.
- [34] Bucacas KL, Franco LM, Shaw CA, Bray MS, Wells JM, Nino D, et al. Early patterns of gene expression correlate with the humoral immune response to influenza vaccination in humans. *J Infect Dis* 2011;203(April (7)):921–9.
- [35] Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, et al. Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol* 2011;12(8):786–95.
- [36] Querec TD, Akondy RS, Lee EK, Cao W, Nakaya HI, Teuwen D, et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nat Immunol* 2009;10(January (1)):116–25.
- [37] Zhu W, Higgs BW, Morehouse C, Streicher K, Ambrose CS, Woo J, et al. A whole genome transcriptional analysis of the early immune response induced by live attenuated and inactivated influenza vaccines in young children. *Vaccine* 2010;28(April (16)):2865–76.
- [38] Ryan M, Murphy G, Ryan E, Nilsson L, Shackley F, Gotheffors L, et al. Distinct T-cell subtypes induced with whole cell and acellular pertussis vaccines in children. *Immunology* 1998;93(January (1)):1–10.

- [39] Hollams EM, Devereill M, Serralha M, Suriyaarachchi D, Parsons F, Zhang G, et al. Elucidation of asthma phenotypes in atopic teenagers through parallel immunophenotypic and clinical profiling. *J Allergy Clin Immunol* 2009;124(September (3)):463–70, 70 e1–16.
- [40] Dan Y, Katakura Y, Ametani A, Kaminogawa S, Asano Y. IL-3 augments TCR-mediated responses of type 2 CD4 T cells. *J Immunol* 1996;156(January (1)):27–34.
- [41] Nakata Y, Brignier AC, Jin S, Shen Y, Rudnick SI, Sugita M, et al. c-Myb, Menin, GATA-3, and MLL form a dynamic transcription complex that plays a pivotal role in human T helper type 2 cell development. *Blood* 2010;116(August (8)):1280–90.
- [42] Hervas-Stubbs S, Perez-Gracia JL, Rouzaut A, Sanmamed MF, Le Bon A, Melero I. Direct effects of type I interferons on cells of the immune system. *Clin Cancer Res* 2011;17(May (9)):2619–27.
- [43] Pappas DJ, Coppola G, Gabatto PA, Gao F, Geschwind DH, Oksenberg JR, et al. Longitudinal system-based analysis of transcriptional responses to type I interferons. *Physiol Genomics* 2009;38(August (3)):362–71.
- [44] Chen Z, Lund R, Aittokallio T, Kosonen M, Nevalanien O, Lahesmaa R. Identification of novel IL-4/Stat6-regulated genes in T lymphocytes. *J Immunol* 2003;171:3627–35.