



Predicting RTS,S Vaccine-Mediated Protection from Transcriptomes in a Malaria-Challenge Clinical Trial

Robert A. van den Berg^{1*}, Margherita Coccia¹, W. Ripley Ballou¹, Kent E. Kester^{2†}, Christian F. Ockenhouse^{2†}, Johan Vekemans¹, Erik Jongert¹, Arnaud M. Didierlaurent¹ and Robbert G. van der Most¹

OPEN ACCESS

¹GSK Vaccines, Rue de l'Institut, Rixensart, Belgium, ²Walter Reed Army Institute of Research, Silver Spring, MD, USA

Edited by:

Alexandre Barbosa Reis,
Universidade Federal de
Ouro Preto, Brazil

Reviewed by:

Roney Santos Coimbra,
Oswaldo Cruz Foundation, Brazil
Katie Louise Flanagan,
Monash University, Australia

*Correspondence:

Robert A. van den Berg
robert.x.van-den-berg@gsk.com

†Present address:

Kent E. Kester,
Sanofi Pasteur, Swiftwater,
PA, USA;
Christian F. Ockenhouse,
PATH Malaria Vaccine Initiative,
Washington, DC, USA

Specialty section:

This article was submitted
to Vaccines and Molecular
Therapeutics,
a section of the journal
Frontiers in Immunology

Received: 10 February 2017

Accepted: 25 April 2017

Published: 23 May 2017

Citation:

van den Berg RA, Coccia M,
Ballou WR, Kester KE,
Ockenhouse CF, Vekemans J,
Jongert E, Didierlaurent AM
and van der Most RG (2017)
Predicting RTS,S Vaccine-Mediated
Protection from Transcriptomes in a
Malaria-Challenge Clinical Trial.
Front. Immunol. 8:557.
doi: 10.3389/fimmu.2017.00557

The RTS,S candidate malaria vaccine can protect against controlled human malaria infection (CHMI), but how protection is achieved remains unclear. Here, we have analyzed longitudinal peripheral blood transcriptome and immunogenicity data from a clinical efficacy trial in which healthy adults received three RTS,S doses 4 weeks apart followed by CHMI 2 weeks later. Multiway partial least squares discriminant analysis (N-PLS-DA) of transcriptome data identified 110 genes that could be used in predictive models of protection. Among the 110 genes, 42 had known immune-related functions, including 29 that were related to the NF- κ B-signaling pathway and 14 to the IFN- γ -signaling pathway. Post-dose 3 serum IFN- γ concentrations were also correlated with protection; and N-PLS-DA of IFN- γ -signaling pathway transcriptome data selected almost all (44/45) of the representative genes for predictive models of protection. Hence, the identification of the NF- κ B and IFN- γ pathways provides further insight into how vaccine-mediated protection may be achieved.

Keywords: malaria, systems biology, transcriptome, vaccine, clinical trial, interferon gamma, RTS,S, adjuvant

INTRODUCTION

The medical burden of malaria disease remains high, most notably in Africa, where most of the deaths due to malaria occur in children under 5 years of age (1). Malaria infection is initiated by the mosquito bite, from which *Plasmodium* sporozoites pass to the liver *via* the blood to infect hepatocytes. The entry into hepatocytes is mediated by circumsporozoite protein (CSP); a protein that is highly expressed at the surface of the sporozoite (2). CSP is also the target of the RTS,S candidate malaria vaccine, in which RTS,S is a recombinant antigen derived from CSP from *Plasmodium falciparum* and the hepatitis B surface antigen. The selection of CSP was informed by the results from vaccination with inactivated sporozoites (3). Vaccination with inactivated sporozoites can result in sterile immunity, which has been associated with activation of CSP-specific cell-mediated immunity (CMI) and production of CSP-specific antibodies (4–7).

Abbreviations: 1dPI, 1dPII, and 1dPIII, 1 day after dose 1, dose 2, and dose 3, respectively; 3dPI and 14dPIII, 3 and 14 days after dose 3, respectively; AUC, area under the curve; CHMI, controlled human malaria infection; CMI, cell-mediated immunity; CSP, circumsporozoite protein; DCV, double cross-validation; DL, delayed onset of parasitemia; DOC, day of challenge; GSEA, Gene Set Enrichment Analysis; NP, non-protected against parasitemia; N-PLS-DA, multiway partial least squares discriminant analysis; PR, protected against parasitemia; prePI, prePII, and prePIII, prior to dose 1, dose 2, and dose 3, respectively; ROC, receiver operating characteristic.

RTS,S has been shown to provide partial protection against clinical and severe disease to infants and young children (8) in a phase III field trial. Protective efficacy has also been demonstrated in adults in phase II trials after controlled human malaria infection (CHMI; by bites received from *Plasmodium*-infected mosquitoes) and in the field (8). The degree of protection provided by RTS,S is dependent on the type of vaccine adjuvant in its composition (9, 10): Adjuvant System AS01 is currently selected for the RTS,S composition and has replaced AS02 (11). Both AS01 and AS02 contain the immunostimulants MPL, a toll-like receptor 4 agonist and the saponin QS-21. The two Adjuvant Systems differ in that AS01 is formulated with liposomes and AS02 is formulated as an oil-in-water emulsion (11). In AS01, the combination of MPL and QS-21 enhances antibody and T-cell responses to vaccine antigens, potentially *via* the transient stimulation of the innate immune system which induces efficient antigen-presenting dendritic cells (12, 13).

A recent analysis of the RTS,S phase III trial has identified the CSP-specific serum antibody concentration as a surrogate of RTS,S-mediated protection (14). Some, but not all, RTS,S studies suggest that CSP-specific CMI measured in peripheral blood is also associated with protection (9, 10, 15–19). In some RTS,S studies, CSP-specific interferon (IFN) gamma (IFN- γ) induction (detected by ELISPOT in cultures of PBMCs) has been associated with protection against malaria-related endpoints both in the field and in the CHMI setting (10, 20, 21). IFN- γ -ELISPOT responses to vaccination also appear to be enhanced when RTS,S is combined with AS01 or AS02 (10, 20, 22). CSP-specific CD4⁺ T cells have also been associated with protection in some CHMI and field trials (10, 18, 23, 24). However, those associations have been made with CD4⁺ T cells that predominantly express IL-2 or TNF- α , rather than IFN- γ . Therefore, questions remain as to whether there are other molecular markers in blood samples that can better predict and perhaps explain RTS,S-mediated protection.

Systems biology approaches can interrogate information from large data sets so as to identify predictive signatures of outcomes such as protection, or immunological correlates of protection (25–28). In the present study, multiway partial least squares discriminant analysis (N-PLS-DA) (29–31) was selected to identify relationships between transcriptome data and protection against parasitemia in clinical-trial recipients of RTS,S (10). In that trial, 78 malaria-naïve adult recipients of either RTS,S/AS01 or RTS,S/AS02 were challenged with CHMI, and 31/78 were protected against parasitemia (Figure 1). Transcriptome data were collected from the available archived samples, representing 39 RTS,S recipients, and covered several time points from pre-vaccination to the day of challenge (DOC) at 14 days post-dose 3 (14dPIII). Although that trial estimated efficacies at 50% for RTS,S/AS01 and 32% for RTS,S/AS02, the estimates were not statistically different ($p = 0.11$), probably reflecting the small scale of the trial. To maximize the statistical power of the present study, samples from recipients of either vaccine were considered as a single population. This approach was based on the assumption that by containing the same immunostimulants (MPL and QS-21), AS01 and AS02 have qualitatively similar mechanisms of action.

Multiway partial least squares discriminant analysis was chosen as a data analysis strategy for two main reasons. First, it is capable of analyzing biological changes over time. Second, it can build predictive models on the correlations between class information (in this case protection status) and trends in the data. Therefore, N-PLS-DA contrasts with an earlier analysis of gene expression data from the same clinical trial which considered only a single time point post sporozoite challenge for classification analysis (32, 33). In the present study, N-PLS-DA was applied to the expression patterns of 20,442 microarray probe sets and of selected probe sets of genes in the IFN-pathway. In both approaches, the results were supportive

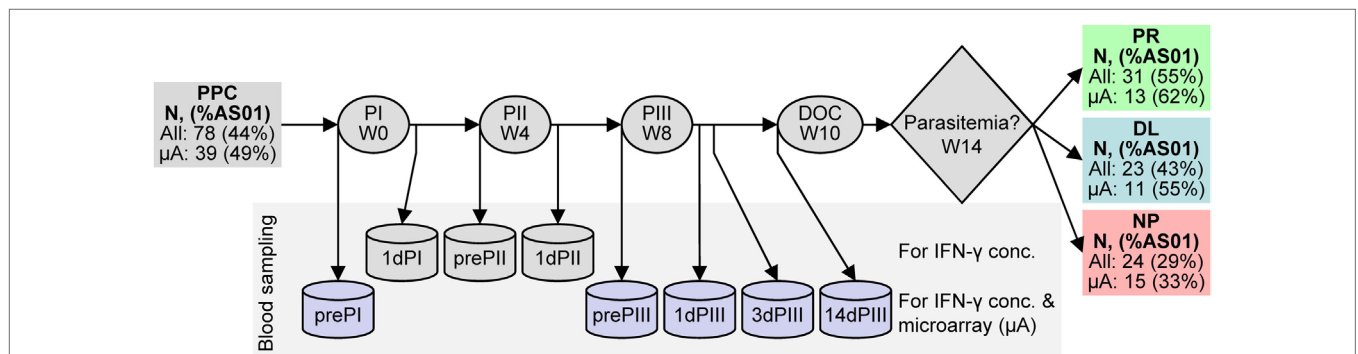


FIGURE 1 | Schematic representation of RTS,S candidate vaccine clinical trial design and efficacy results. RTS,S vaccine injections were performed at weeks (W) 0, 4, and 8. Blood was sampled (represented by can shapes) before the first, second, and third vaccine injections (prePI, prePII, and prePIII, respectively), one day after the first, second, and third vaccine injections (1dPI, 1dPII, and 1dPIII, respectively), and 3 and 14 days after the third vaccine injection (3dPIII, 10dPIII, and 14dPIII, respectively). Controlled human malaria infection (CHMI) was performed at 14dPIII (i.e., week 10; day of challenge, DOC) and the onset of parasitemia was followed up to week 14. Serum IFN- γ concentrations were measured at all of the time points, whereas RNA expression was evaluated at prePI, prePII, 1dPIII, 3dPIII, and 14dPIII (purple cans). The numbers of subjects in the per-protocol cohort (PPC) at study entry and after the outcome of CHMI (either protected (PR), non-protected (NP), or delayed onset of parasitemia (DL)) are indicated (*all* subjects) and correspond to those subjects who also provided blood samples for IFN- γ measurements. The numbers of subjects from which transcriptome data from microarrays (μ A subjects) were derived are indicated below *all* subjects. The percentages of RTS,S/AS01 recipients over recipients of either vaccine (%AS01) are indicated in parentheses for *all* subjects and μ A subjects after vaccination and after challenge. Note that all 12 non-vaccinated control subjects developed parasitemia within the follow-up period after CHMI (not shown).

of the involvement of IFN- γ - and NF- κ B-signaling pathway genes in the process of establishing RTS,S-mediated protection against CHMI.

MATERIALS AND METHODS

Clinical Trial Samples

PBMCs and serum samples were obtained from participants in the RTS,S vaccine clinical trial (ClinicalTrials.gov NCT00075049) (10).

Serum IFN- γ and CSP-Specific Antibody Concentrations and Logistic Regression Modeling

Serum IFN- γ was measured using ELISA and protocol from BD Biosciences (Belgium). The assay's limit of quantification was 1.0 pg/ml. The CSP-specific serum antibody-concentration data have been reported previously (10). Logistic regression modeling using SAS software (Version 9.2, SAS Institute Inc., Cary, NC, USA) was based on classifying CHMI outcome [protected (PR) against parasitemia, non-protected (NP) and, delayed onset of parasitemia (DL; defined as individuals in which parasites were not detected in the blood for a period longer than the longest parasitemia-free period in the control group of infected and unvaccinated subjects)] with the immunogenicity data sets (pre- and post-vaccination time points up to the DOC) and Adjuvant System selected (AS01 or AS02) as covariates. The quality of a model was determined by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.

RNA Preparation and Microarray Assay

RNA was prepared from dimethyl sulfoxide-cryopreserved PBMC samples using TRI Reagent and protocol (Molecular Research Center Inc., Cincinnati, OH, USA). The preparation, amplification, and labeling of copy DNA was performed using Ovation RNA Amplification System and protocol (NuGen Technologies, San Carlos, CA, USA). Gene expression levels were determined using Affymetrix HG-U133 Plus 2.0 arrays of 54,120 probe sets and protocol (Affymetrix, Santa Clara, CA, USA). The data were made publicly available *via* ArrayExpress under accession number E-MTAB-4629.

Microarray Data Normalization

The raw microarray data were normalized *via* GCRMA (34), and outliers were excluded. Values that were within the 99.9% of the estimated distribution of the background expression level were considered background signals. The data from 20,442 probe sets, out of 54,120 probe sets represented on the array, were entered into the analysis based on two criteria: (i) the availability of gene annotation information for the specific probe set and (ii) the presence of a signal larger than background in more than 20% of the arrays. The transcriptome data set was represented as a multiway data set in that it was defined by multiple factors [subject \times probe set (gene) \times time]. The 3-way matrix data contained 2.05% missing entries, which were handled by the

multiway N -way toolbox during modeling (35). For the data-driven model, 13 PR and 15 NP subjects were included, and for the IFN-driven model, 12 PR and 15 NP subjects were included. For classification in the data matrices, PR and NP were assigned the values of +1 and -1, respectively. Microarray normalization was performed using the R statistical software (Version 2.11.0) and Bioconductor (Version 2.6).

Model Selection and Validation

Three techniques of model validation and refinement (36) were used in this study (see the N-PLS-DA Methodological Details section and Figure S1 in Supplementary Material for more details on the methodology); (i) label permutations, (ii) building an ensemble of models (37, 38), and (iii) double cross-validation (DCV) (39, 40), where model performance was determined by the fraction of correctly classified subjects and the DQ² statistic (41). Because the modeling was based on a regression analysis, class value predictions <-1 or $>+1$ were assigned the values -1 and +1, respectively. In a cross-validation, the subjects of the data set were divided randomly in N subsets. One subset was placed aside, and the other subsets were used to build the model. In a DCV, the subjects were divided in N subsets and one subset was placed aside. $N - 1$ subsets were again divided into M subsets. The subjects in the initially excluded subset were completely independent of the model building and were used to validate the final model. DCV was used for the selection of the number of components and the probe sets to be included in the N-PLS-DA modeling. The different models were evaluated using the DQ² statistic, which is based on a least squares method for analyzing the difference between prediction and CHMI outcome (PR or NP) (41) and was a more discriminatory method for identifying differences in performance than the fraction of correctly classified outcomes. In cases where the optimum number of model refinement rounds was difficult to determine, the fraction of correctly classified outcomes and a statistic based on the mean difference between prediction score and CHMI outcome was also used. Probe sets were selected for evaluation in the DCV based on their individual prediction performance. For the IFN-driven modeling with many fewer probe sets than the data-driven modeling, a forward selection approach was used (42, 43), and model selection only involved the number of components. Label permutations were used to assess whether a model based on the true class labels (i.e., CHMI outcome) performed better than models using the same model parameters but based on randomly assorted class labels. For both the data-driven and IFN-driven modeling, respective summary models were calculated using only the first two components of the transformed data from all of the selected probe sets from the modeling. The components of the summary model were visualized by appropriate rotation on Cartesian axes (with arbitrary units) (44). The N-PLS-DA analyses were performed using Matlab R2010B (MathWorks) the Matlab Statistics Toolbox, and the N -way toolbox (Version 3.1) (35). The biological interpretation of the gene lists was aided by using Ingenuity pathway analysis and upstream transcription factor analysis. The clustering of transcriptome data for heatmap visualization was performed using Cluster 3.0 (45).

RESULTS

Serum IFN- γ Concentrations and Vaccine-Mediated Protection Status

In the CHMI clinical trial, three outcomes of sporozoite challenge (after three doses of RTS,S adjuvanted with AS01 or AS02) were defined over the 4-week observation period [Figure 1: protected (PR) against parasitemia, non-protected (NP), and DL (defined as individuals in which parasites were not detected in the blood for a period longer than the longest parasitemia-free period in the control group of infected and unvaccinated subjects)].

Cell-mediated immunity, including the production of IFN- γ in ELISPOT assays has been associated with RTS,S-mediated protection (10, 20, 21). Moreover, IFN- γ has been detected in the sera of individuals 1 day after the administration of an AS01- or AS02-adjuvanted hepatitis B vaccine candidates and tuberculosis vaccine candidates (46, 47). Given these observations, serum IFN- γ concentrations were measured prior to each dose (prePI, prePII, and prePIII, respectively), and at 1 day after each dose (1dPI, 1dPII, and 1dPIII, respectively) in the entire per-protocol cohort (PPC) (see Figure 1). Higher median concentrations were observed in PR subjects ($N = 31$) than in infected subjects (NP + DL combined, $N = 47$) at 1 day after dose 1 (1dPI; 6.3 versus 3.2 ng/ml), dose 2 (1dPII; 17.1 versus 8.3 ng/ml), and dose 3 (1dPIII; 20.9 versus 12.3 ng/ml; Figure 2A). In logistic regression modeling of sporozoite challenge outcome (PR versus NP and DL) with respect to the immunogenicity data as covariates, the IFN- γ concentrations at 1dPIII could partially explain protection status. The accuracy of the IFN- γ model was estimated at 0.71 from the calculation of the AUC of the ROC graph (Figure 2B; where an accuracy score of 0.5 indicates no accuracy). By comparison, the accuracy score for CSP-specific antibody concentrations at 14 days post-dose 3 (14dPIII; and DOC) was 0.81.

RNA Expression by IFN- γ -Signaling Pathway Genes in Relation to Protection Status

To further explore the role of IFN- γ in relation to protection against CHMI, transcriptome data were generated from PBMCs isolated from RTS,S vaccinated subjects in the microarray subset ($N = 39$; Figure 1) at prePI and prePIII, and at 1dPIII, 3 days after dose 3 (3dPIII) and 14dPIII. In the microarray subset, the percentages of RTS,S/AS01 recipients over recipients of either adjuvanted vaccine, overall and with respect to each of the post-challenge outcomes, were similar to the respective percentages in the PPC.

With an initial focus on the expression patterns of IFN-pathway genes, an analysis using the Ingenuity pathway analysis tool compared RNA expression patterns in PR and NP subjects. This analysis suggested that in both PR and NP groups, the IFN- γ -signaling pathway was active at 1dPIII (and marked by the increased expression of numerous genes) but not active at 3dPIII (Figure 2C). By contrast, at 14dPIII, the IFN- γ -signaling pathway appeared active in the PR group and not in the NP group.

Although individual PBMC samples were randomly assigned to be processed by one of two batches of microarray kits, the random allocation was not stratified according to protection status, and an imbalance in allocation was observed at 14dPIII. Therefore, by way of validation with respect to a potential batch effect, the expression of the probe sets was examined in a transcriptome data set derived from the same clinical trial, and from a single batch of microarrays (and termed the validation transcriptome) (32). Although the same time points were evaluated, the RNA was not amplified, one less subject was evaluated, and only 74 of the 116 probe sets were present in those microarrays. However, the differences in the patterns of RNA expression in the IFN pathway between PR, NP, and DL groups were similar to that observed in the present transcriptome data set, notably at 14dPIII (see Figure S2 in Supplementary Material that shows the heatmap description of the IFN-pathway gene expression using the validation-transcriptome data set). Therefore overall, the analysis suggested that IFN- γ -signaling could play a role in establishing a protective status prior to sporozoite challenge.

N-PLS-DA of Transcriptome Data Set in Relation to Protection Status

The transcriptome data set was then further analyzed using N-PLS-DA to identify RNA expression patterns that could distinguish between PR and NP subjects. The method captured the multivariate complexity [subject \times probe set (gene) \times time] of the data and was used to build mathematical models of correlations between RNA expression data and the binary outcomes of protection or non-protection (see Figure S1 in Supplementary Material that shows a flow diagram description of the N-PLS-DA). Given the onus on requiring a clear binary outcome for the analysis and the complex nature of the data set, the data from the DL subjects were excluded from the modeling. Two approaches were undertaken in the N-PLS-DA that differed by the probe sets that were analyzed. In the first approach, the entire data set was analyzed using N-PLS-DA (data-driven N-PLS-DA), with the aim of obtaining an unbiased selection of genes that could be correlated with protection status. The second approach considered whether IFN-pathway genes alone could be correlated with protection status by only including the data from a selected (i.e., IFN-related) group of probe sets (IFN-driven N-PLS-DA).

Using the data-driven approach, 100 individual models were generated, and each model typically consisted of data from 2 to 40 probe sets. Overall, these models could correctly predict sporozoite challenge outcome on average in 78% of subjects. A total of 116 probe sets were identified on the basis that each probe set was represented at least once in the 100 models. These probe sets corresponded to 110 genes (see Table S1 in Supplementary Material that lists the genes/probe sets selected by the data-driven modeling).

To explore the relationships identified in the 100 models, a summary model was calculated using only the first two components of the transformed data from all of the selected 116 probe sets (consequently, the summary model was not an optimized and

validated model because the probe set selection and the number of components were determined outside the model validation context).

In the summary model and when the data were defined by subject identity, most of the data points could be separated into distinct PR and NP clusters (Figure 3A, upper panel). The N-PLS-DA was done with data from the PR and NP subjects, leaving out the data from the DL subjects. Given that differences were detected between PR and NP subjects, it was of interest to study the data from DL subjects with respect to the two clusters. Hence, when the summary model was imposed on the data from the DL subjects, the data points were distributed among both of the PR and NP clusters. The data from six subjects (three PR and three NP; and three each from AS01 and AS02 groups, respectively) who were often misclassified in the prediction rounds during model validation were mostly positioned close the boundary between PR and NP clusters. When the data were defined by probe set identity, two distinct clusters could be identified by applying *post hoc* *k*-means clustering (48) (Figure 3A, middle panel) and further analysis revealed that these two clusters corresponded to two different types of expression kinetics in PR subjects or in NP subjects (see below). When the data were defined by sampling time point (Figure 3A, lower panel), the dispersion between consecutive time points was larger for 1dPIII to 3dPIII and 3dPIII to 14dPIII data points than for prePI to prePIII and prePIII to 1dPIII data points suggesting that the 3dPIII and 14dPIII time points contributed most to predicting the distinction between PR and NP in the modeling.

Using additional *k*-means clustering, four distinct clusters of probe sets with different expression-kinetic profiles were identified (Figure 3B and see Figure 3A, middle panel; Table S1 in Supplementary Material). This clustering was performed on the RNA expression values of the 116 probe sets from the PR subjects at different time points relative (by subtraction with log-transformed values) to the prePI time point. The same clustering was then imposed on NP and DL data sets (Figure 3B). The greatest divergence in the expression kinetics between the PR and NP groups occurred between 3dPIII and 14dPIII. The directionality of this divergence for Clusters A and C probe sets was opposite to that for Clusters B and D probe sets. For Clusters A and C probe sets, expression levels decreased in the PR group and increased in the NP group between 3dPIII and 14dPIII, whereas the reverse was observed for Clusters B and D probe sets. Interestingly, although the expression levels for the DL group at 14dPIII were similar to those of the NP group, the directionality of expression kinetics for the DL group between 3dPIII and 14dPIII was similar to those of the PR group, suggesting that the selected 116 probe sets may be able to distinguish the DL outcome from both PR and NP outcomes.

IFN- γ -Pathway Genes Feature among the Genes Identified by the Data-Driven N-PLS-DA

Of the 110 genes represented in the 100 models, 42 have been previously ascribed immune-related functions, of which 13 were

represented at least five times in the models (Figure 4; see Tables S1 and S2 in Supplementary Material that lists the references to support the characterization of the immune-related genes

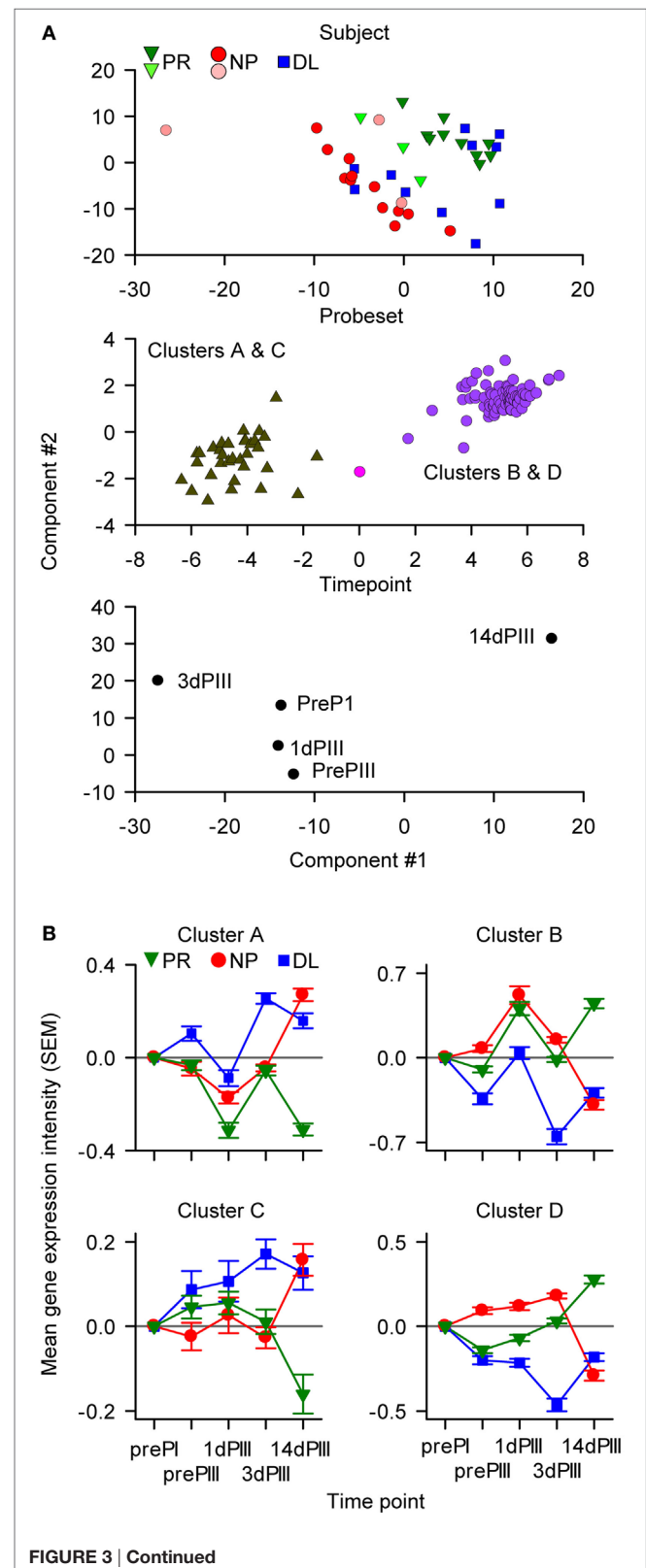


FIGURE 3 | Continued
The data-driven multiway partial least squares discriminant analysis identified 116 probe sets (110 genes) that were classified into four clusters. (A) The components of the summary model were visualized by appropriate rotation on Cartesian axes (with arbitrary units). In the upper graph, data points are distributed with respect to the protection status of subjects. The six data points representing the subjects that were often misclassified in the modeling are represented by the lighter colored symbols. In the middle graph, the data points are distributed with respect to probe set identity into two groups. Each of these two groups comprised two clusters of probe sets (Clusters A and C, and Clusters B and D) based on RNA expression patterns. In the lower graph, the data points are distributed with respect to sampling time [pre-dose 1 (prePI), at pre-dose 3 (prePIII), and 1, 3, and 14 days after dose 3 (1dPIII, 3dPIII, and 14dPIII, respectively)]. (B) Mean RNA expression levels relative to prePI, at prePIII and 1dPIII, 3dPIII, and 14dPIII, with respect to protection status of subjects [protected (PR), non-protected (NP), and non-protected with delayed parasitemia (DL)] for each of the four clusters (A–D) of probe sets among the 116 probe sets. The error bars indicate the SEM.

Gene	Cluster	Frequency	Pathway			Gene	Cluster	Frequency	Pathway		
			NF-κB	IFN-γ	Ubiq.				NF-κB	IFN-γ	Ubiq.
<i>FBXO9</i>	A	30	■	■	■	<i>IRF7</i>	B	1	■	■	■
<i>IL23A</i>	A	4	■	■	■	<i>KLHL6</i>	B	1	■	■	■
<i>CNPY4</i>	A	2	■	■	■	<i>MT2A</i>	B	1	■	■	■
<i>PDCD4</i>	A	1	■	■	■	<i>PTPN6</i>	B	1	■	■	■
<i>HLA-A</i>	B	71	■	■	■	<i>RHBDF2</i>	B	1	■	■	■
<i>RNF31</i>	B	58	■	■	■	<i>TAX1BP1</i>	B	1	■	■	■
<i>MYD88</i>	B	32	■	■	■	<i>TNIP2</i>	B	1	■	■	■
<i>BAG1</i>	B	29	■	■	■	<i>ZFP36</i>	B	1	■	■	■
<i>NCAPH2</i>	B	13	■	■	■	<i>CD302</i>	C	1	■	■	■
<i>GADD45B</i>	B	9	■	■	■	<i>UCP2</i>	D	8	■	■	■
<i>PML</i>	B	8	■	■	■	<i>STUB1</i>	D	8	■	■	■
<i>SCO2</i>	B	8	■	■	■	<i>SHARPIN</i>	D	5	■	■	■
<i>WWP2</i>	B	6	■	■	■	<i>RRAS</i>	D	3	■	■	■
<i>SP110</i>	B	3	■	■	■	<i>RPS6KA3</i>	D	2	■	■	■
<i>HLA-B</i>	B	2	■	■	■	<i>COMMD1</i>	D	1	■	■	■
<i>HSP90B1</i>	B	2	■	■	■	<i>DPP3</i>	D	1	■	■	■
<i>MT1F</i>	B	2	■	■	■	<i>ICAM2</i>	D	1	■	■	■
<i>RNASEH2C</i>	B	2	■	■	■	<i>MYBBP1A</i>	D	1	■	■	■
<i>WAS</i>	B	2	■	■	■	<i>PLIN2</i>	D	1	■	■	■
<i>HLA-C</i>	B	1	■	■	■	<i>RAD23A</i>	D	1	■	■	■
<i>HLA-DMB</i>	B	1	■	■	■	<i>STK10</i>	D	1	■	■	■

FIGURE 4 | Forty-two genes selected by the data-driven modeling were immune-related. Tabular description of the association between the 42 immune-related genes selected by the data-driven modeling and NF-κB, IFN-γ, and ubiquitination (Ubiq.) pathways (associations represented by black boxes). For each gene, the cluster allocation and the frequency it represented in the modeling process is indicated. For genes that were represented by more than one probe set, the data for the most frequently represented probe set are shown (see Table S1 in Supplementary Material). The references supporting the immune-related associations are detailed in Table S2 in Supplementary Material.

selected by the data-driven modeling). The immune-related genes were most frequently identified in Cluster-B (i.e., 25/42 genes) including two (*HLA-A* and *RNF31*) of the three most frequently represented genes. Fourteen immune-related genes have been associated with regulating IFN-γ signaling or with being

downstream targets of IFN-γ signaling. These latter genes include the major histocompatibility complex (MHC) class I and class II genes. Twenty-nine immune-related genes have been associated with the NF-κB pathway, including upstream regulators such as the frequently represented gene *MYD88*, a key mediator of TLR and IL1-R signaling. Ten immune-related genes (including frequently represented genes *RNF31*, *BAG1*, *FBXO9*, *PML*, *STUB1*, *WWP2*, and *SHARPIN*) have been associated with regulating ubiquitination, nine of which have been also associated with the NF-κB pathway. Genes such as *STUB1*, *PLIN2*, *PML*, *HSP90B1*, and *RAD23A*, in addition to the MHC class I genes, have been associated with the antigen presentation pathway. Two frequently represented genes (*NCAPH2* and *GADD45B*) have been associated with T-cell and B-cell survival, respectively. For the other 69 genes, immune-related functions have not been described. Nineteen of these genes were represented at least five times including *GTF2E2* and *GTF2F1*, which were the first and fifth most frequently represented genes overall, respectively, and code for subunits of the general transcription factor complex that is essential for transcriptional initiation.

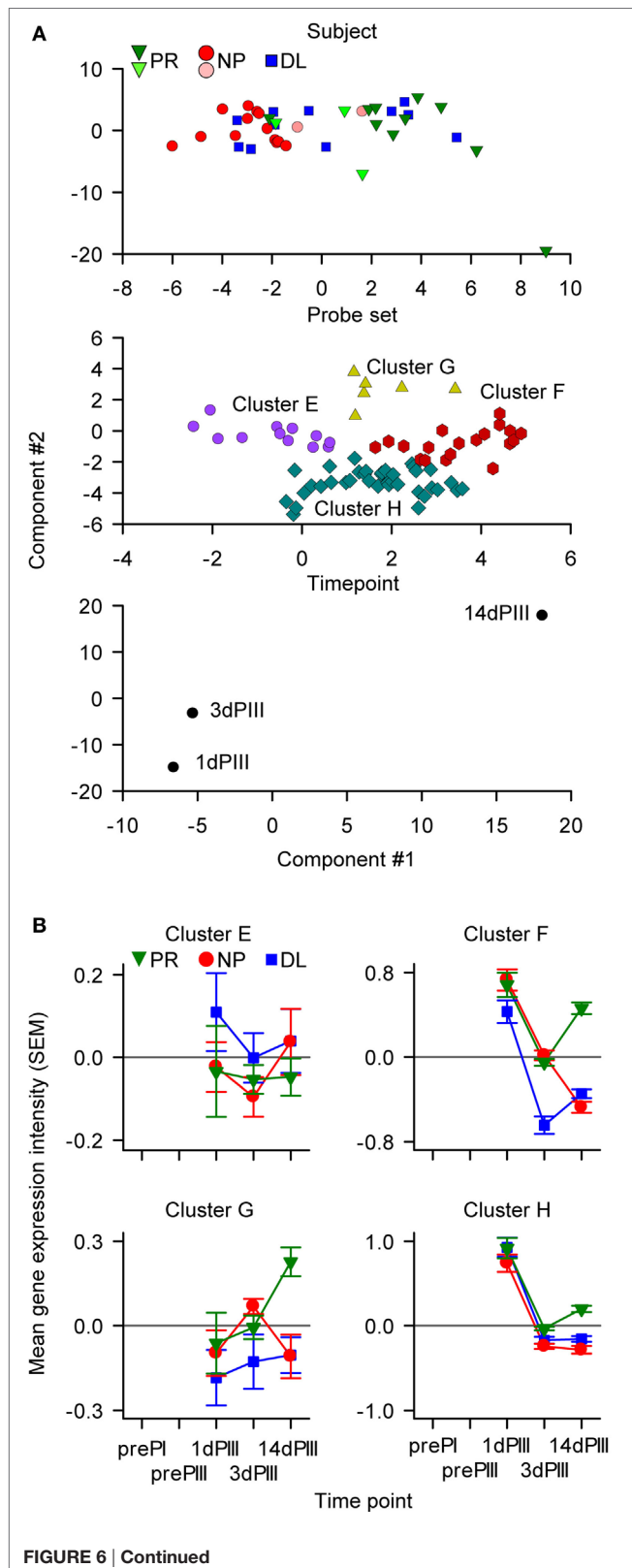
Further connections between genes were analyzed by identifying common upstream transcription factors. Thirty two of the 110 genes were allocated to one or several of 23 groups defined by a common upstream transcription factor (or complex) (Figure 5). Twenty-two of the genes identified were immune-related and each group contained at least one immune-related gene. As expected, the NF-κB complex of transcription factors was identified as regulating several genes. Other immune-related transcription factors were also identified such as CIITA, RFX5, and SATB1 [that regulate *HLA* genes (49, 50)], NR3C1 [glucocorticoid receptor (51, 52)] and REL [part of the NF-κB complex (53, 54)]. Although transcriptional regulators such as TP53, MYC, CREB1, STAT3, and CEBPA have been implicated in numerous pathways, they have also been associated with regulating innate immune pathways (55–63).

The expression patterns of the probe sets from the different clusters were also examined in the validation-transcriptome data set. A clear distinction between PR and NP subjects was observed for Cluster-B probe sets (see Figure S3 in Supplementary Material that shows the expression of Clusters A–D probe sets using the validation-transcriptome data set), with a divergence in expression kinetics from 3dPIII to 14dPIII that was similar to that in the present transcriptome data set, suggesting that the differences in expression of Cluster-B probe sets with respect to protection status were not masked by a potential batch effect at 14dPIII. However, whether batch effect abrogated differences in expression patterns for the probe sets from the other clusters could not be concluded because of the potentially inferior quality of validation-transcriptome data (e.g., absence of an RNA amplification, fewer probe sets).

Predictive Models of Protection Can Be Built from Probe Set Data Representing IFN-γ-Signaling Pathway

Based on the links between the IFN-γ secretion data (Figure 2A) and IFN-γ-signaling pathway that emerged from the pathway

The most frequently represented gene by the 74 probe sets in the modeling was *TAP1*, appearing 79 times (out of 100 models). *TAP1* and other frequently represented genes such as *JAK1*



(67 times) *STAT1* (44 times), *IRF1* (36 times) form a core part of the IFN- γ -signaling pathway, and exemplify genes for which expression levels at 14dPIII (relative to prePI) were distinctly different between the PR and NP groups, in contrast to only minor differences in expression levels at 1dPIII. Moreover, *TAP1*, *STAT1*, and *IRF1* were represented by Cluster-F probe sets and therefore may capture differences between the DL outcome from both PR and NP outcomes. The frequently represented genes *MT2A* (66 times), *IRF7* (36 times), and *PML* (8 times) were also identified in the data-driven N-PLS-DA.

DISCUSSION

Systems biology analyses are potentially advantageous in clinical trial settings where multiparametric data are often generated (28, 64, 65). Signals from both innate and adaptive immune responses can be identified in transcriptome data derived from peripheral blood from 3 days after vaccination (25, 26). Here, we have identified two gene signatures by modeling longitudinal PBMC-transcriptome data from clinical trial subjects who were vaccinated with RTS,S/AS01 or RTS,S/AS02 and then challenged with *P. falciparum* sporozoites. In having an efficacy endpoint in the vaccine trial, our study makes an advance on earlier systems biology analyses of other vaccine trials in which only immunogenicity endpoints were considered (25–27, 66–68).

Our study was hypothesis generating in its design, given that the gene signatures were not validated in an independent and similarly designed clinical trial. Also, the interpretation of the systems biology analyses was limited by three principal factors. First, the samples were derived from peripheral blood and not at the injection site or draining lymph nodes where most of the direct responses to the vaccine components would have occurred. Second, the samples contained a heterogeneous population of cells, and hence the variation in gene-expression levels may have arisen from relative differences in cell population sizes or from relative differences in a given cell population. Third, to increase the statistical power, the samples from recipients of either AS01- or AS02-adjuvanted vaccines were considered as a single population based on the assumption that by containing the same immunostimulants (MPL and QS-21), AS01 and AS02 have qualitatively similar mechanisms of action.

The two gene signatures were identified by applying N-PLS-DA both to the entire transcriptome data set and to data from selected genes within the IFN- γ -signaling pathway. In the first data-driven approach, a gene signature, consisting of 110 genes including 42 known immune-related genes was identified. Of the immune-related genes, many were associated with the IFN- γ and NF- κ B pathways; and a third of the genes associated with the NF- κ B pathway were also associated with regulating ubiquitination. Hence, the validity of the gene signature was supported by the functional connectivity of many of the immune-related genes. Other selected genes may have immune-related functions but these remain to be characterized. For example, genes such as those of the general transcription factor complex (*GTF2E2* and *GTF2F1*) may have been selected because they may have a prominent function in immune-related cell populations. The second approach confirmed that a gene signature based on IFN-pathway genes alone could have a similar predictive capacity to that of the data-driven gene signature because 44 of the 45 *a priori* selected genes were included in the modeling. Overall, the differences in serum IFN- γ concentrations and the identities of the two gene signatures suggested that the IFN- γ pathway plays a role in RTS,S vaccine-mediated protection. The gene signatures may have captured differences between subjects in terms of responses to the Adjuvant System (AS01 or AS02) because these Adjuvant Systems appear to better enhance IFN- γ -related CMI responses to vaccination in comparison with other adjuvants (20, 69).

The predictive power of the N-PLS-DA modeling was around 78% because in the two approaches, the models often misclassified four NP and four PR subjects (i.e., 16–19%). This suggests that the transcriptome data alone was inadequate in these cases to make accurate predictions, perhaps because of other unknown confounding factors. In the primary analysis of the clinical trial, a similar proportion of individuals (8/51, 16%; 5P and 3NP), but not all the same individuals, were misclassified based on assignments decided by threshold DOC antibody titers (10).

Multiway partial least squares discriminant analysis contrasts with other transcriptome data analysis in that it can incorporate both the kinetics of the immune response and the inherently multivariate nature of transcriptome data. Our conclusions extend previous transcriptome data analyses of the same vaccinated cohort, in which individual time points were considered separately (32, 33). Rinchai et al. (33) identified differences in gene expression between PR and NP subjects at 1dPIII, 3dPIII, and 14dPIII using a predefined cluster of 130 probe sets (Module 1.2) that had previously been shown to reflect type-1 interferon signaling (68). Vahey et al. (32) identified the immunoproteasome pathway using a supervised approach [Gene Set Enrichment Analysis (GSEA)] on gene-expression data at 14dPIII (DOC). Although Module 1.2 or Cluster-B (in our study) did not include probe sets representing those immunoproteasome genes, the immunoproteasome is regulated by IFN- γ (70–72), and therefore, the identification of those genes may have reflected differences in IFN- γ signaling between PR and NP subjects. Furthermore, one Cluster-B gene, *PML*, has been identified as a regulator of the immunoproteasome (73). Also, the IFN-driven N-PLS-DA confirmed that genes related to the immunoproteasome [i.e., *TAP1*

(74) and *PSMB8* (*LMP7*) (70)] can be incorporated into the predictive models.

The trends in the gene expression kinetics suggested that differences between PR and NP subjects was most notable with the divergence of the expression kinetics after 3dPII leading to clear differences in relative expression levels at 14dPIII for the four clusters of the data-driven gene signature and Cluster-F of the IFN-driven gene signature. Furthermore, the expression kinetics appeared to distinguish the DL outcome from both the PR and NP outcomes, in that the trajectory of the expression kinetics in DL subjects after 3dPIII were similar to those in PR subjects. Surprisingly, the differences between PR and NP groups in terms of expression levels of the clusters of IFN-driven signature genes was not apparent at 1dPIII and 3dPIII, even though the differences in serum IFN- γ concentrations between PR and NP groups at 1dPIII appeared to contribute toward protection. Rather, the expression data suggested that the IFN- γ pathway was transiently activated in the peripheral circulation in all groups, but signs of its reactivation occurred primarily in PR subjects, at 14dPIII.

Therefore, our hypothesis is that the gene signatures reflect the presence of activated effector/effector memory (E/EM) CD4⁺ T cells that are induced after immunization and are responsible for IFN- γ signaling in a different manner in PR subjects to that in NP subjects. The potential role of CD4⁺ T cells in the release of IFN- γ into the serum is suggested by the observed increase in the serum IFN- γ concentrations after repeated vaccination in this study and in a previous study of AS01- and AS02-adjuvanted hepatitis B vaccines (46). Although it is possible that antigen-specific CD4⁺ or CD8⁺ T cells were directly responsible for IFN- γ production, this cytokine was not prominently produced by antigen-specific CD4⁺ T cells, and antigen-specific CD8⁺ T cells were not detected in PBMCs after vaccination in the same clinical trial (10, 23). Alternatively, other cells, such as NK cells, may have produced IFN- γ . NK cells have been shown to produce IFN- γ in response to signals from antigen-specific CD4⁺ T cells (75–77) and from other innate immune cells activated by AS01 in the lymph nodes draining the vaccine injection site (Margherita Coccia, unpublished data). Given that antigen-specific CD4⁺ T cells can stimulate IFN- γ production in NK cells in an IL-2-dependent manner (75–77), it is notable that in a previous analysis of the same clinical trial, the frequencies of IL-2 producing T_{E/EM} and central memory T cells in PR subjects were significantly higher than in NP subjects (23). Also CD4⁺ T cell-mediated activation of NK cells may depend on CD40–CD40L interactions (78) as well as NF- κ B signaling (79). Another intriguing possibility is that NK cells may be differently primed in PR versus NP individuals by the time of the third RTS,S dose, thus contributing to the differences in IFN- γ production after the third dose. Kazmin et al. performed a systems biology analysis (80) of a more recent CHMI efficacy trial of RTS,S (19), in which individual time points pre- and post-vaccination were considered, and which used the same validation-transcriptome data set as our study. In their study, several models predictive of protection were identified at prePIII (the day of the third RTS,S dose). The frequently represented genes in those models and other GSEA analyses identified an inverse correlation between NK-cell-related gene expression

in the blood and protection, suggesting that in those individuals who were subsequently protected, there was a greater efflux of NK cells from the blood expressing homing receptors to the draining lymph node or injection site between the second and third RTS,S doses.

Hence, an enticing corollary to our hypothesis is that the interactions between NK cells, CSP-specific antibodies, and CD4⁺ T cells may be directly relevant in the subsequent clearance of sporozoite-infected hepatocytes through mechanisms such as antibody-mediated cell cytotoxicity. A role of IFN- γ produced by NKT cells to suppress sporozoite-infected hepatocytes was recently described in a mouse model (81). It is possible, therefore, that RTS,S-induced memory responses were capable of mobilizing NKT cells or other sources of IFN- γ production in the protective response to sporozoite challenge. Furthermore, and despite the hypothesis-generating nature of our study, the genes identified by the systems biology analyses shed useful light on understanding how protection against malaria parasitemia is achieved by RTS,S.

ETHICS STATEMENT

The clinical trial was conducted in accordance with all applicable regulatory requirements, including the Declaration of Helsinki (1996). The clinical trial was approved by institutional review boards from the Walter Reed Army Institute of Research (WRAIR) Human Use Review Committee (HURC) and the United States Army Medical Research and Materiel Command (USAMRMC) Human Subjects Research Review Board (HSRRB). Written-informed consent was obtained from each participant prior to the performance of any study-specific procedures in accordance with relevant ICH Guidelines, US Army Regulations, and principles of the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

WRB, KK, and CO participated in the design and conduct of the clinical trial from which the samples were acquired. RB, JV, EJ, and RM conceived, and/or designed and implemented the

transcriptome data analyses. RB, RM, AD, and MC conceived, designed, and implemented the immunogenicity logistical-regression analysis. All authors contributed to the interpretation of the results, as well as to the development of this manuscript. All authors had full access to the data and approved the manuscript before it was submitted by the corresponding author. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official of as reflecting the views of the United States Department of the Army or the Department of Defense.

ACKNOWLEDGMENTS

We thank all the clinical trial participants, the investigators, the nurses, clinicians, laboratory personnel, and other staff members for conducting the trial. We thank those who contributed to the present study, and in particular, Walthère Dewé (GSK Vaccines, Belgium) for his help with the statistical analysis. Matthew Morgan (MG Science Communications, Belgium) provided scientific writing services and Ulrike Krause (GSK Vaccines, Belgium) provided editorial assistance and manuscript coordination.

FUNDING

This study was sponsored by GlaxoSmithKline Biologicals SA using samples acquired from the clinical trial (ClinicalTrials.gov NCT00075049). The costs associated with the development and publishing of the manuscript, including scientific writing assistance, were provided by GlaxoSmithKline Biologicals SA. MC received funding from the Marie Curie Intra European grant PIEF-GA-2012-330788 ADJSYN and from the Bill & Melinda Gates Foundation grant OPP1220977.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fimmu.2017.00557/full#supplementary-material>.

REFERENCES

1. Organization WH. *World Malaria Report 2015*. Geneva: WHO Press (2015).
2. Cerami C, Frevert U, Sinnis P, Takacs B, Clavijo P, Santos MJ, et al. The basolateral domain of the hepatocyte plasma membrane bears receptors for the circumsporozoite protein of *Plasmodium falciparum* sporozoites. *Cell* (1992) 70:1021–33. doi:10.1016/0092-8674(92)90251-7
3. Cohen J, Nussenzweig V, Nussenzweig R, Vekemans J, Leach A. From the circumsporozoite protein to the RTS, S/AS candidate vaccine. *Hum Vaccin* (2010) 6:90–6. doi:10.4161/hv.6.1.9677
4. Oliveira GA, Kumar KA, Calvo-Calle JM, Othoro C, Altszuler D, Nussenzweig V, et al. Class II-restricted protective immunity induced by malaria sporozoites. *Infect Immun* (2008) 76:1200–6. doi:10.1128/IAI.00566-07
5. Seder RA, Chang LJ, Enama ME, Zephir KL, Sarwar UN, Gordon IJ, et al. Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. *Science* (2013) 341:1359–65. doi:10.1126/science.1241800
6. Behet MC, Foquet L, Van Gemert GJ, Bijker EM, Meuleman P, Leroux-Roels G, et al. Sporozoite immunization of human volunteers under chemoprophylaxis induces functional antibodies against pre-erythrocytic stages of *Plasmodium falciparum*. *Malar J* (2014) 13:136. doi:10.1186/1475-2875-13-136
7. Ishizuka AS, Lyke KE, Dezure A, Berry AA, Richie TL, Mendoza FH, et al. Protection against malaria at 1 year and immune correlates following PfSPZ vaccination. *Nat Med* (2016) 22(6):614–23. doi:10.1038/nm.4110
8. Agnandji ST, Fernandes JF, Bache EB, Ramharther M. Clinical development of RTS,S/AS malaria vaccine: a systematic review of clinical phase I-III trials. *Future Microbiol* (2015) 10:1553–78. doi:10.2217/fmb.15.90
9. Stoute JA, Slaoui M, Heppner DG, Momin P, Kester KE, Desmons P, et al. A preliminary evaluation of a recombinant circumsporozoite protein vaccine against *Plasmodium falciparum* malaria. *N Engl J Med* (1997) 336:86–91. doi:10.1056/NEJM199701093360202
10. Kester KE, Cummings JF, Ofori-Anyinam O, Ockenhouse CF, Krzych U, Moris P, et al. Randomized, double-blind, phase 2a trial of falciparum malaria vaccines RTS,S/AS01B and RTS,S/AS02A in malaria-naïve adults: safety, efficacy, and immunologic associates of protection. *J Infect Dis* (2009) 200:337–46. doi:10.1086/600120
11. Garçon N, Van Mechelen M. Recent clinical experience with vaccines using MPL- and QS-21-containing adjuvant systems. *Expert Rev Vaccines* (2011) 10:471–86. doi:10.1586/erv.11.29
12. Dendouga N, Fochesato M, Lockman L, Mossman S, Giannini SL. Cell-mediated immune responses to a varicella-zoster virus glycoprotein E vaccine

- using both a TLR agonist and QS21 in mice. *Vaccine* (2012) 30:3126–35. doi:10.1016/j.vaccine.2012.01.088
13. Didierlaurent AM, Collignon C, Bourguignon P, Wouters S, Fierens K, Fochesato M, et al. Enhancement of adaptive immunity by the human vaccine adjuvant AS01 depends on activated dendritic cells. *J Immunol* (2014) 193:1920–30. doi:10.4049/jimmunol.1400948
 14. White MT, Verity R, Griffin JT, Asante KP, Owusu-Agyei S, Greenwood B, et al. Immunogenicity of the RTS,S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial. *Lancet Infect Dis* (2015) 15:1450–8. doi:10.1016/S1473-3099(15)00239-X
 15. Schwenk R, Asher LV, Chalom I, Lanar D, Sun P, White K, et al. Opsonization by antigen-specific antibodies as a mechanism of protective immunity induced by *Plasmodium falciparum* circumsporozoite protein-based vaccine. *Parasite Immunol* (2003) 25:17–25. doi:10.1046/j.1365-3024.2003.00495.x
 16. Pinder M, Reece WH, Plebanski M, Akinwunmi P, Flanagan KL, Lee EA, et al. Cellular immunity induced by the recombinant *Plasmodium falciparum* malaria vaccine, RTS,S/AS02, in semi-immune adults in the Gambia. *Clin Exp Immunol* (2004) 135:286–93. doi:10.1111/j.1365-2249.2004.02371.x
 17. Olotu A, Lusingu J, Leach A, Lievens M, Vekemans J, Msham S, et al. Efficacy of RTS,S/AS01E malaria vaccine and exploratory analysis on anti-circumsporozoite antibody titres and protection in children aged 5–17 months in Kenya and Tanzania: a randomised controlled trial. *Lancet Infect Dis* (2011) 11:102–9. doi:10.1016/S1473-3099(10)70262-0
 18. Olotu A, Moris P, Mwacharo J, Vekemans J, Kimani D, Janssens M, et al. Circumsporozoite-specific T cell responses in children vaccinated with RTS,S/AS01E and protection against *P. falciparum* clinical malaria. *PLoS One* (2011) 6:e25786. doi:10.1371/journal.pone.0025786
 19. Ockenhouse CF, Regules J, Tosh D, Cowden J, Kathcart A, Cummings J, et al. Ad35.CS.01 – RTS,S/AS01 heterologous prime boost vaccine efficacy against sporozoite challenge in healthy malaria-naïve adults. *PLoS One* (2015) 10:e0131571. doi:10.1371/journal.pone.0131571
 20. Sun P, Schwenk R, White K, Stoute JA, Cohen J, Ballou WR, et al. Protective immunity induced with malaria vaccine, RTS,S, is linked to *Plasmodium falciparum* circumsporozoite protein-specific CD4⁺ and CD8⁺ T cells producing IFN- γ . *J Immunol* (2003) 171:6961–7. doi:10.4049/jimmunol.171.12.6961
 21. Reece WH, Pinder M, Gothard PK, Milligan P, Bojang K, Doherty T, et al. A CD4⁺ T-cell immune response to a conserved epitope in the circumsporozoite protein correlates with protection from natural *Plasmodium falciparum* infection and disease. *Nat Med* (2004) 10:406–10. doi:10.1038/nm1009
 22. Ansong D, Asante KP, Vekemans J, Owusu SK, Owusu R, Brobby NA, et al. T cell responses to the RTS,S/AS01(E) and RTS,S/AS02(D) malaria candidate vaccines administered according to different schedules to Ghanaian children. *PLoS One* (2011) 6:e18891. doi:10.1371/journal.pone.0018891
 23. Lumsden JM, Schwenk RJ, Rein LE, Moris P, Janssens M, Ofori-Anyinam O, et al. Protective immunity induced with the RTS,S/AS vaccine is associated with IL-2 and TNF- α producing effector and central memory CD4⁺ T cells. *PLoS One* (2011) 6:e20775. doi:10.1371/journal.pone.0020775
 24. Ndungu FM, Mwacharo J, Kimani D, Kai O, Moris P, Jongert E, et al. A statistical interaction between circumsporozoite protein-specific T cell and antibody responses and risk of clinical malaria episodes following vaccination with RTS,S/AS01E. *PLoS One* (2012) 7:e52870. doi:10.1371/journal.pone.0052870
 25. 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) 10:116–25. doi:10.1038/ni.1688
 26. 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) 12:786–95. doi:10.1038/ni.2067
 27. Li S, Roupheal N, Duraisingham S, Romero-Steiner S, Presnell S, Davis C, et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat Immunol* (2014) 15:195–204. doi:10.1038/ni.2789
 28. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* (2014) 157:499–513. doi:10.1016/j.cell.2014.03.031
 29. Hall GJ, Clow KE, Kenny JE. Estuarial fingerprinting through multidimensional fluorescence and multivariate analysis. *Environ Sci Technol* (2005) 39:7560–7. doi:10.1021/es0503074
 30. Castro C, Manetti C. A multiway approach to analyze metabonomic data: a study of maize seeds development. *Anal Biochem* (2007) 371:194–200. doi:10.1016/j.ab.2007.08.028
 31. Wopereis S, Rubingh CM, Van Erk MJ, Verheij ER, Van Vliet T, Cnubben NH, et al. Metabolic profiling of the response to an oral glucose tolerance test detects subtle metabolic changes. *PLoS One* (2009) 4:e4525. doi:10.1371/journal.pone.0004525
 32. Vahey MT, Wang Z, Kester KE, Cummings J, Heppner DG Jr, Nau ME, et al. Expression of genes associated with immunoproteasome processing of major histocompatibility complex peptides is indicative of protection with adjuvanted RTS,S malaria vaccine. *J Infect Dis* (2010) 201:580–9. doi:10.1086/650310
 33. Rinchai D, Presnell S, Chaussabel D. Blood interferon signatures putatively link lack of protection conferred by the RTS,S recombinant malaria vaccine to an antigen-specific IgE response. *F1000Res* (2015) 4:919. doi:10.12688/f1000research.7093.1
 34. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F. A model-based background adjustment for oligonucleotide expression arrays. *J Am Stat Assoc* (2004) 99:909–17. doi:10.1198/016214504000000683
 35. Andersson CA, Bro R. The N-way toolbox for MATLAB. *Chemom Intell Lab Syst* (2000) 52:1–4. doi:10.1016/S0169-7439(00)00071-X
 36. Westerhuis JA, Hoefsloot HC, Smit S, Vis A, Smilde AK, Van Velzen EJ, et al. Assessment of PLS-DA cross validation. *Metabolomics* (2008) 4:81–9. doi:10.1007/s11306-007-0099-6
 37. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer (2009).
 38. Yang P, Yang YH, Zhou BB, Zomaya AY. A multi-filter enhanced genetic ensemble system for gene selection and sample classification of microarray data. *Curr Bioinform* (2010) 5:296–308. doi:10.1186/1471-2105-11-S1-S5
 39. Stone M. Cross-validated choice and assessment of statistical predictions. *J R Stat Soc Series B Stat Methodol* (1974) 36:111–47.
 40. Smit S, Van Breemen MJ, Hoefsloot HC, Smilde AK, Aerts JM, De Koster CG. Assessing the statistical validity of proteomics based biomarkers. *Anal Chim Acta* (2007) 592:210–7. doi:10.1016/j.aca.2007.04.043
 41. Westerhuis JA, Van Velzen EJ, Hoefsloot HC, Smilde AK. Discriminant Q² (DQ²) for improved discrimination in PLS-DA models. *Metabolomics* (2008) 4:293–6. doi:10.1007/s11306-008-0126-2
 42. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* (1997) 97:273–324. doi:10.1016/S0004-3702(97)00043-X
 43. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* (2007) 23:2507–17. doi:10.1093/bioinformatics/btm344
 44. Kiers HAL. Some procedures for displaying results from three-way methods. *J Chemometr* (2000) 14:151–70. doi:10.1002/1099-128X(200005/06)14:3<151::AID-CEM585>3.0.CO;2-G
 45. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* (2004) 20:1453–4. doi:10.1093/bioinformatics/bth078
 46. Vandepapelière P, Horsmans Y, Moris P, Van Mechelen M, Janssens M, Koutsoukos M, et al. Vaccine adjuvant systems containing monophosphoryl lipid A and QS21 induce strong and persistent humoral and T cell responses against hepatitis B surface antigen in healthy adult volunteers. *Vaccine* (2008) 26:1375–86. doi:10.1016/j.vaccine.2007.12.038
 47. Leroux-Roels I, Forgue S, De Boever F, Clement F, Demoitie MA, Mettens P, et al. Improved CD4⁺ T cell responses to Mycobacterium tuberculosis in PPD-negative adults by M72/AS01 as compared to the M72/AS02 and Mtb72F/AS02 tuberculosis candidate vaccine formulations: a randomized trial. *Vaccine* (2013) 31:2196–206. doi:10.1016/j.vaccine.2012.05.035
 48. Kaufman L, Rousseeuw PJ. *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: John Wiley & Sons Inc (1990).
 49. Mach B, Steimle V, Martinez-Soria E, Reith W. Regulation of MHC class II genes: lessons from a disease. *Annu Rev Immunol* (1996) 14:301–31. doi:10.1146/annurev.immunol.14.1.301
 50. Kumar PP, Bischof O, Purbey PK, Notani D, Urlaub H, Dejean A, et al. Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nat Cell Biol* (2007) 9:45–56. doi:10.1038/ncb1516
 51. Dejager L, Pinheiro I, Puimege L, Fan YD, Gremeaux L, Vankelecom H, et al. Increased glucocorticoid receptor expression and activity mediate the LPS resistance of SPRET/EI mice. *J Biol Chem* (2010) 285:31073–86. doi:10.1074/jbc.M110.154484

52. Chinenov Y, Gupte R, Rogatsky I. Nuclear receptors in inflammation control: repression by GR and beyond. *Mol Cell Endocrinol* (2013) 380:55–64. doi:10.1016/j.mce.2013.04.006
53. Kieran M, Blank V, Logeat F, Vandekerckhove J, Lottspeich F, Le Bail O, et al. The DNA binding subunit of NF- κ B is identical to factor KBF1 and homologous to the *rel* oncogene product. *Cell* (1990) 62:1007–18. doi:10.1016/0092-8674(90)90275-J
54. Baltimore D. NF- κ B is 25. *Nat Immunol* (2011) 12:683–5. doi:10.1038/ni.2072
55. Greenhill CJ, Rose-John S, Lissilaa R, Ferlin W, Ernst M, Hertzog PJ, et al. IL-6 trans-signaling modulates TLR4-dependent inflammatory responses via STAT3. *J Immunol* (2011) 186:1199–208. doi:10.4049/jimmunol.1002971
56. Menendez D, Shatz M, Azzam K, Garantzios S, Fessler MB, Resnick MA. The toll-like receptor gene family is integrated into human DNA damage and p53 networks. *PLoS Genet* (2011) 7:e1001360. doi:10.1371/journal.pgen.1001360
57. Yim HC, Li JC, Pong JC, Lau AS. A role for c-Myc in regulating antimicrobial responses. *Proc Natl Acad Sci U S A* (2011) 108:17749–54. doi:10.1073/pnas.1104892108
58. Fu H, Yang G, Wei M, Liu L, Jin L, Lu X, et al. The RNA-binding protein QKI5 is a direct target of C/EBP α and delays macrophage differentiation. *Mol Biol Cell* (2012) 23:1628–35. doi:10.1091/mbc.E11-05-0412
59. Mayer TZ, Simard FA, Cloutier A, Vardhan H, Dubois CM, McDonald PP. The p38-MSK1 signaling cascade influences cytokine production through CREB and C/EBP factors in human neutrophils. *J Immunol* (2013) 191:4299–307. doi:10.4049/jimmunol.1301117
60. Menendez D, Shatz M, Resnick MA. Interactions between the tumor suppressor p53 and immune responses. *Curr Opin Oncol* (2013) 25:85–92. doi:10.1097/CCO.0b013e32835b6386
61. Welner RS, Bararia D, Amabile G, Czibere A, Benoukraf T, Bach C, et al. C/EBP α is required for development of dendritic cell progenitors. *Blood* (2013) 121:4073–81. doi:10.1182/blood-2012-10-463448
62. Wen AY, Landaw EM, Ochoa R, Cho M, Chao A, Lawson G, et al. Increased abscess formation and defective chemokine regulation in CREB transgenic mice. *PLoS One* (2013) 8:e55866. doi:10.1371/journal.pone.0055866
63. Zhang H, Hu H, Greeley N, Jin J, Matthews AJ, Ohashi E, et al. STAT3 restrains. *Nat Commun* (2014) 5:5798. doi:10.1038/ncomms6798
64. Six A, Bellier B, Thomas-Vaslin V, Klatzmann D. Systems biology in vaccine design. *Microb Biotechnol* (2012) 5:295–304. doi:10.1111/j.1751-7915.2011.00321.x
65. Li S, Nakaya HI, Kazmin DA, Oh JZ, Pulendran B. Systems biological approaches to measure and understand vaccine immunity in humans. *Semin Immunol* (2013) 25:209–18. doi:10.1016/j.smim.2013.05.003
66. Gaucher D, Therrien R, Kettaf N, Angermann BR, Boucher G, Filali-Mouhim A, et al. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *J Exp Med* (2008) 205:3119–31. doi:10.1084/jem.20082292
67. Zak DE, Andersen-Nissen E, Peterson ER, Sato A, Hamilton MK, Borgerding J, et al. Merck Ad5/HIV induces broad innate immune activation that predicts CD8⁺ T-cell responses but is attenuated by preexisting Ad5 immunity. *Proc Natl Acad Sci U S A* (2012) 109:E3503–12. doi:10.1073/pnas.1208972109
68. Obermoser G, Presnell S, Domico K, Xu H, Wang Y, Anguiano E, et al. Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. *Immunity* (2013) 38:831–44. doi:10.1016/j.immuni.2012.12.008
69. Leroux-Roels G, Marchant A, Levy J, Van Damme P, Schwarz TF, Horsmans Y, et al. Impact of adjuvants on CD4⁺ T cell and B cell responses to a protein antigen vaccine: results from a phase II, randomized, multicenter trial. *Clin Immunol* (2016) 169:16–27. doi:10.1016/j.clim.2016.05.007
70. Griffin TA, Nandi D, Cruz M, Fehling HJ, Kaer LV, Monaco JJ, et al. Immunoproteasome assembly: cooperative incorporation of interferon γ (IFN- γ)-inducible subunits. *J Exp Med* (1998) 187:97–104. doi:10.1084/jem.187.1.97
71. Kloetzel PM, Ossendorp F. Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Curr Opin Immunol* (2004) 16:76–81. doi:10.1016/j.coi.2003.11.004
72. Heink S, Ludwig D, Kloetzel P-M, Krüger E. IFN- γ -induced immune adaptation of the proteasome system is an accelerated and transient response. *Proc Natl Acad Sci U S A* (2005) 102:9241–6. doi:10.1073/pnas.0501711102
73. Yang XW, Wang P, Liu JQ, Zhang H, Xi WD, Jia XH, et al. Coordinated regulation of the immunoproteasome subunits by PML/RAR α and PU.1 in acute promyelocytic leukemia. *Oncogene* (2014) 33:2700–8. doi:10.1038/nc.2013.224
74. Hwang LY, Lieu PT, Peterson PA, Yang Y. Functional regulation of immunoproteasomes and transporter associated with antigen processing. *Immunol Res* (2001) 24:245–72. doi:10.1385/IR:24:3:245
75. Horowitz A, Behrens RH, Okell L, Fooks AR, Riley EM. NK cells as effectors of acquired immune responses: effector CD4⁺ T cell-dependent activation of NK cells following vaccination. *J Immunol* (2010) 185:2808–18. doi:10.4049/jimmunol.1000844
76. Horowitz A, Newman KC, Evans JH, Korbel DS, Davis DM, Riley EM. Cross-talk between T cells and NK cells generates rapid effector responses to *Plasmodium falciparum*-infected erythrocytes. *J Immunol* (2010) 184:6043–52. doi:10.4049/jimmunol.1000106
77. Horowitz A, Hafalla JC, King E, Lusingu J, Dekker D, Leach A, et al. Antigen-specific IL-2 secretion correlates with NK cell responses after immunization of Tanzanian children with the RTS,S/AS01 malaria vaccine. *J Immunol* (2012) 188:5054–62. doi:10.4049/jimmunol.1102710
78. Bihl F, Pecheur J, Bréart B, Poupon G, Cazareth J, Julia V, et al. Primed antigen-specific CD4⁺ T cells are required for NK cell activation in vivo upon *Leishmania major* infection. *J Immunol* (2010) 185:2174–81. doi:10.4049/jimmunol.1001486
79. Tato CM, Villarino A, Caamano JH, Boothby M, Hunter CA. Inhibition of NF- κ B activity in T and NK cells results in defective effector cell expansion and production of IFN- γ required for resistance to *Toxoplasma gondii*. *J Immunol* (2003) 170:3139–46. doi:10.4049/jimmunol.170.6.3139
80. Kazmin D, Nakaya HI, Johnson M, Van Der Most R, Van Den Berg RA, Ballou WR, et al. Systems analysis of protective immune responses to RTS,S malaria vaccination in humans. *Proc Natl Acad Sci U S A* (2017) 114:2425–30. doi:10.1073/pnas.1621489114
81. Miller JL, Sack BK, Baldwin M, Vaughan AM, Kappe SH. Interferon-mediated innate immune responses against malaria parasite liver stages. *Cell Rep* (2014) 7:436–47. doi:10.1016/j.celrep.2014.03.018

Conflict of Interest Statement: All authors declared the following interests: RB, MC, WRB, AD, EJ, and RM are employees of the GSK group of companies. JV was an employee of the GSK group of companies at the time of the study but is now affiliated with WHO, which was not involved in the study. The opinions or assertions contained herein are not to be construed as reflecting the views of the WHO. RB, WRB, AD, EJ, and RM report ownership of GSK shares and/or restricted shares. KK is an employee of Sanofi Pasteur. CO declares no competing financial interests.

Copyright © 2017 van den Berg, Coccia, Ballou, Kester, Ockenhouse, Vekemans, Jongert, Didierlaurent and van der Most. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.