Title

Transcriptome-wide changes in gene expression, splicing, and lncRNAs in response to a live attenuated dengue virus vaccine

Journal

Cell Reports, 38(6)

ISSN

2639-1856

Authors

Kim, Eun-Young
Che, Yan
Dean, Hansi J
et al.

Publication Date

2022-02-01

DOI

10.1016/j.celrep.2022.110341

Peer reviewed

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2022 April 09.

Published in final edited form as:

Cell Rep. 2022 February 08; 38(6): 110341. doi:10.1016/j.celrep.2022.110341.

Transcriptome-wide changes in gene expression, splicing, and lncRNAs in response to a live attenuated dengue virus vaccine

Eun-Young Kim[1], **Yan Che**[1], **Hansi J. Dean**[2], **Ramon Lorenzo-Redondo**[1], **Michael Stewart**[1], **Caroline K. Keller**[1], **Daniel Whorf**[1], **Dawson Mills**[1], **Nikita N. Dulin**[1], **Tiffany Kim**[1], **Megan Votoupal**[1], **Miriam Walter**[1], **Ana Fernandez-Sesma**[3], **Heejin Kim**[1], **Steven M. Wolinsky**[1,4,*]

[1]Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60011, USA

[2]Takeda Vaccines, Inc, Cambridge, MA 02139, USA

[3]Department of Microbiology and Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[4]Lead contact

SUMMARY

The tetravalent dengue vaccine candidate, TAK-003, induces a functional antibody response, but the titers of antibodies against the four serotypes of the dengue virus (DENV) can vary. Here, through a transcriptomic analysis on whole blood collected from recipients of a two-dose schedule of TAK-003, we examine gene expression, splicing, and transcript isoform-level changes for both protein-coding and noncoding genes to broaden our understanding of the immune response. Our analysis reveals a dynamic pattern of vaccine-associated regulation of long noncoding RNAs (lncRNAs), differential splicing of interferon-stimulated gene exons, and gene expression changes related to multiple signaling pathways that detect viral infection. Co-expression networks isolate immune cell-type-related and interferon-response modules that represent specific biological processes that correlate with more robust antibody responses. These data provide insights into the early determinants of the variable immune response to the vaccine, highlighting the significance of splicing and isoform-level gene regulatory mechanisms in defining vaccine immunogenicity.

In brief

*Correspondence: s-wolinsky@northwestern.edu.

AUTHOR CONTRIBUTIONS

S.M.W., E.-Y.K., and H.J.D. conceived, planned, and supervised the experiments. H.J.D. contributed samples and case reviews. E.-Y.K., R.L.R., M.S., C.K.K., D.W., M.V., M.W., and D.M. performed the experiments and validated the results. Y.C., N.N.D., T.K., H.K., and E.-Y.K. performed the data analyses. S.M.W., E.-Y.K., and Y.C. wrote the manuscript with support from H.J.D. and A.F.-S.

DECLARATION OF INTERESTS

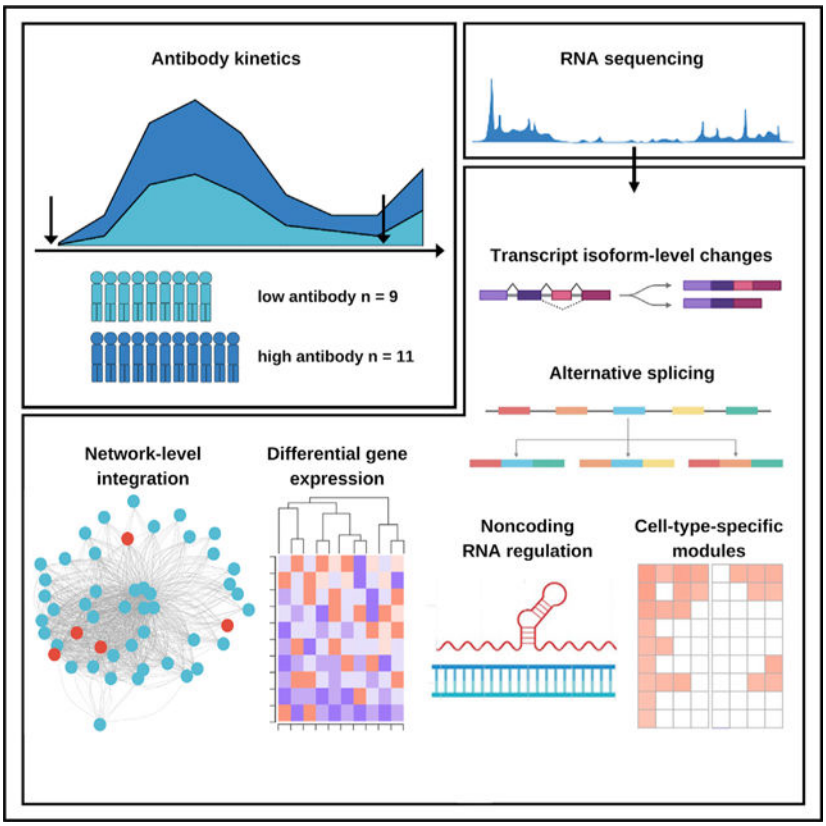
H.J.D. was an employee (retired) of Takeda Pharmaceutical Company Ltd. The content is solely the responsibility of the authors. It does not necessarily represent the views of Takeda Pharmaceutical Company Ltd or NIH. The other authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110341.

Kim et al. use genome-wide transcriptome analysis to reveal the molecular pathways, cell-type-related and interferon-response modules, defining immune mechanisms associated with the titers of neutralizing antibody. These data contribute to a coherent emerging picture of vaccine-elicited immunity, highlighting the significance of splicing and isoform-level gene regulation in vaccine response.

Graphical Abstract



INTRODUCTION

Dengue is a mosquito-borne viral infection that causes global outbreaks and epidemics in tropical and sub-tropical regions with expanding ranges, creating an imminent threat to public health (Bhatt et al., 2013). Human illness caused by dengue virus (DENV) infection is clinically classified as uncomplicated dengue, dengue with warning signs, and severe dengue (Bhatt et al., 2013). Although most people who become infected with the virus recover on their own without requiring medical treatment, an estimated 500,000 people will develop severe dengue each year, which can be fatal (Bhatt et al., 2013). DENV belongs to the Flaviviridae family of single-stranded RNA viruses that includes other arthropod-borne viruses, including yellow fever virus, Zika virus, and Japanese encephalitis virus. DENV has four genetically related but distinct serotypes (DENV-1 through DENV-4), each of which is further classified into genotypes (Holmes and Twiddy, 2003; Katzelnick et al., 2015). A second infection with a different serotype heightens the risk of severe

dengue (Gibbons et al., 2007; Olkowski et al., 2013). Antibody-dependent enhancement of disease occurs when non-neutralizing cross-reactive antibodies bind to virus particles of the second infecting serotype to facilitate virus entry into cells through Fc-receptor-mediated endocytosis (Bournazos et al., 2020). The increased virus production and suppressed antiviral defenses can cause severe illness and pose a risk to vaccine recipients who were previously exposed to DENV (Deng et al., 2020).

The vaccine TAK-003, a formulation of four live attenuated chimeric viruses in which the prM and E proteins of each DENV-1, DENV-3, and DENV-4 replace DENV-2, is under active investigation in humans. It has demonstrated safety, immunogenicity, and effectiveness in preventing dengue illness that varied according to serotype (Biswal et al., 2019). Live attenuated vaccines induce immunity by stimulating the innate immune system and providing antigen to specify the adaptive immune responses (Iwasaki and Medzhitov, 2015). Pattern-recognition receptors (PRRs), including the endosome-associated Toll-like receptor 3 (TLR3) and TLR7 and the cytoplasmic retinoic acid-inducible gene I (RIG-I), melanoma differentiation-associated protein 5 (MDA5), and NOD-like receptor protein 3 (NLRP3)-specific inflammasome, activate downstream signaling pathways through recognition of their ligands (Medzhitov, 2007). The activation of downstream signaling pathways can induce and effect type I and type III interferons (IFNs) and proinflammatory cytokine productions and hundreds of interferon-stimulated genes (ISGs) that contribute to the early control of DENV infection (Hur, 2019; Lazear et al., 2019). DENV, however, can interfere with interferon regulator factor (IRF) and nuclear factor-κB (NF-κB) activation to replicate its genome within its human host (Garcia-Sastre, 2017).

In addition to direct antiviral activity and innate immune cell activation, type I IFN signaling can regulate adaptive immune responses by promoting or inhibiting the activation, proliferation, and differentiation of various immune cell types (Iwasaki and Medzhitov, 2015). The transcriptional and post-transcriptional regulation of gene expression in signaling pathways that trigger IFN production is intrinsic to the activation of DENV-specific adaptive immunity (Carpenter et al., 2013). The ISGs induced by vaccination may correlate with and sometimes predict the variations in the adaptive immune responses (Chaussabel and Baldwin, 2014; Gaucher et al., 2008; Li et al., 2014; Querec et al., 2009). However, information about how regulation of innate immunity affects the kinetics and magnitude of the protective adaptive immune response for a vaccine and how they correlate with transcriptional activation of the primary immune response genes, post-transcriptional regulators, and coregulatory interactions remain poorly defined.

In this study, we performed extensive analyses of the protein-coding, noncoding, and alternatively spliced transcriptome and gene and isoform co-expression networks in blood samples collected from healthy adults without previous dengue exposure who received two doses, 90 days apart, of the TAK-003 vaccine candidate. Our analysis reveals distinct gene expression and transcript isoform expression patterns and differential splicing for protein-coding and noncoding genes in defining the early transcriptomic signatures elicited by TAK-003. Our data suggest that signatures of innate immunity correlate with the neutralizing antibody titers. Co-expression networks isolate immune celltype-related and IFN-response modules, implicating early B cell activation and differentiation, antigen

presentation by dendritic cells, and cell cycle phase transition in the induction of a more robust antibody response. Together, these data demonstrate the specific transcriptional pathways and the immune cell types involved, emphasizing the significance of splicing and isoform-level gene regulatory mechanisms for inducting IFNs and ISGs in promoting an effective immune response.

RESULTS

Study design

Twenty recipients of a two-dose regimen of the TAK-003 vaccine candidate were selected for study from a phase 1 clinical trial involving healthy adults from an admixed Columbian population without pre-existing immunity to DENV or other flaviviruses (Osorio et al., 2014). An overview of the study design is provided in Figure S1. Blood samples were collected at baseline and days 2, 4, and 7 after the first dose and day 2 after the second dose (92 days after the first dose). The first dose induced functional antibody responses that were significantly higher in the high response group when compared with the low response group (445.9 ± 162.5 versus 25.5 ± 17.0, respectively; Mann-Whitney *U* Test, p = 0.0003) and persistent in both groups (Figure S1B). The vaccine elicited higher serotype-specific antibody titers to DENV-2 than the other serotypes (Osorio et al., 2014). A booster dose of TAK-003 given 90 days after the primary dose elicited a modest increase of neutralizing antibodies against DENV-1, DENV-2, and DENV-3. The high and low response groups (n = 11 and n = 9, respectively) were matched for sex, age, and genetic ancestry (data not shown). We detected transient low-level viremia by quantitative serotype-specific RT-PCR more frequently after the first dose, particularly among those in the high response group (Fisher's exact test, p = 0.0379) (Osorio et al., 2014), suggesting that the attenuated vaccine strains replicated more efficiently in those vaccine recipients and led to a more vigorous immune response. No correlation was found between viremia and the titer of neutralizing antibody in the high response group. The minimal viremia after the second dose may involve more limited virus replication and spread by the innate and adaptive immune responses induced by the first dose (Reinhardt et al., 1998).

Identification of the differentially expressed genes and transcripts

RNA sequencing (RNA-seq) results from blood specimens collected over time enabled comprehensive assessments of gene expression and transcript isoform expression for protein-coding and noncoding genes (Figure S1F). The proportion of total reads mapping to the DENV reference was low but comparable between groups (Figure S1C). The lowest proportions observed were at day 4. Differentially expressed genes (DEGs; DESeq2 negative binomial model) were identified for all samples across a four-point time course with the baseline for each vaccine recipient as their control. Only genes with an absolute log-transformed fold-change ($|\log_2 FC|$) >1 at false discovery rate (FDR) < 0.05 and identified in at least 60% of the vaccine recipients were considered DEGs (Table S1A). To validate the differential gene expression results, we performed a multiplex analysis of immune response-related gene expression (both protein-coding and noncoding) on 347 genes on a custom nCounter panel (nanoString Technologies). We found significant concordance in

log_2FC compared with those for RNA-seq (average Pearson correlation coefficient across replicates, $r^2 = 0.89$, $p \leq 2.2 \times 10^{-16}$)(Figures S1D and S1E; Table S2).

We compared the high and low response groups to identify genes whose baseline expression ($|log_2FC| \geq 1$, FDR <0.05) might predict the outcome for TAK-003 vaccination. We identified five genes – *GIMAP2*, *ZBTB34*, *ZNF354A*, *SOCS5*, and *hnRNPF* – that were significantly associated with the magnitude of the functional antibody response (Figure S2A). This specific gene subset was a baseline predictor for our cohort (Tsang et al., 2020). We observed pervasive differential gene expression changes, compared with baseline, across the four-point time course (n = 480, n = 13, n = 101, and n = 2,308 genes at days 2, 4, 7, and 92, respectively). Those DEGs were more likely to be upregulated than downregulated (n = 400 versus 80, n = 9 versus 4, n = 84 versus 17, and n = 1981 versus 327 genes at days 2, 4, 7, and 92, respectively) (Figure S2B). To capture the transcription kinetics, we generated *z*-scores from the transformed DEGs. We segregated the expression data into clusters by the *k*-means algorithm. We identified eight distinct clusters with similar temporal patterns across the four-point time course (Figure S2D and Table S1B). Those clusters related to innate immunity exhibited substantial enrichment for known molecular pathways involved in type I IFN signaling, immunometabolism, and the inflammasome complex (Figure S2C).

Differential expression analysis revealed both distinctive and shared genes that were significant after adjustment for multiple comparisons (FDR <0.05) when comparing the two groups, with higher expression levels after the first dose associated with higher antibody titers (Figures 1A and S3A–S3C; Table S1C). We identified 3,108 DEGs with activity after the first and second doses. The kinetics of these DEGs varied between time points in the two groups without a temporal delay in their expression in the low response group. These DEGs were mostly observed at days 2 and 92 (n = 406 and n = 2,819, respectively) compared with days 4 and 7 (n = 62 and n = 162, respectively). Those gene signatures consisted of a similar set of DEGs involved in the type I IFN response, including *DDX58* (also known as RIG-I) and *IFIH1* (also known as MDA-5), the primary cytosolic receptors responsible for the recognition of RNA, and the IFN-inducible genes *RSAD2*, *OAS2*, *OASL*, *PARP9*, *MX1*, *IFI6*, *ZBP1*, *XAF1*, and *ZBP1*, that have essential roles in viral restriction and apoptosis-promoting activities (Figures 1B and S3D; Table S1C). Negative regulators of the type I IFN response were identified, including *USP18*, *HERC5*, and *ISG15*, which code for ubiquitin-like proteins that maintain cellular homeostasis mechanisms triggered by viruses (Ivashkiv and Donlin, 2014). Compared with previously published yellow fever vaccine YF-17D clinical trials datasets, there were few similarities in their transcriptional signatures of vaccination that correlate with and predict the subsequent adaptive immune response (Figures S4A and S4B) (Kotliarov et al., 2020; Querec et al., 2009).

To better understand how the transcriptomic response might relate to biological processes specific to the first or second dose, we performed Gene Ontology (GO) enrichment analysis of DEGs ($|log_2FC| \geq 1$, FDR <0.05) between the two groups or genes selectively upregulated or downregulated when compared across the four-point time course. A more robust functional antibody response showed a significant correlation ($-log_{10}p > 10$) with the upregulated DEGs in the GO enrichment analysis (Figure 1C). We found that genes involved in TLR-mediated My88-dependent signaling, TIR domain-containing adaptors

signaling, type I IFN and antiviral innate immune response, and inflammatory response were upregulated in both groups (Frishberg et al., 2019; Uhlen et al., 2019). When comparing the enriched gene sets shared between the two groups, the GO analysis revealed a significant upregulation of genes involved in biological processes related to the activation of an IFN-regulated gene expression program (Figures S4C–S4F). We found both groups displayed elevated expression of the genes involved in the type I IFN signaling responses, various responses to stress, ribosomal RNA processing, and mitochondrial translation datasets most significantly after the second dose (Figure 1C). The Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathway analyses on these genes confirmed that an IFN-regulated gene expression program is highly enriched in the high response group, most significantly at day 2 compared with day 92.

We expanded these analyses to the transcript isoform-level to discover additional gene expression alterations. We observed substantial sharing of the differential transcript expression (DTE) and DEG signals and a transcriptome gradient reflecting the $\log_2$FC from baseline across the four-point time course (Tables S1C and S1D). The isoform-level changes (DTE) showed a larger differential expression effect size ($|\log_2 FC|$) compared with gene expression changes (two-sided Kolmogorov-Smirnov test statistic, $p < 2.2 \times 10^{-16}$), especially for the protein-coding genes (Figure 2A). The most significant changes were observed 2 days after the second dose. DTE and DEG showed similar GO term enrichments ($-\log_{10} p > 30$) for upregulation in viral replication, immunometabolism, and catabolic process pathways and downregulation in cell-type-related activation and cell process pathways (Figure 2B). The combined GO and KEGG functional networks and Reactome pathways demonstrated significant enrichment for RIG-I-like receptor signaling, type I IFN signaling, response to type I IFN, antiviral mechanisms, cellular response to type I IFN, and negative regulation of viral genome replication.

Expression of the noncoding transcriptome

We mapped 445 long noncoding RNAs (lncRNAs) that matched the annotated loci in Ensembl, passed filter for $|\log_2 FC| \geq 1$ (FDR <0.05), and were differentially expressed (Figure 3A and Table S1E). Noncoding RNAs have been shown to participate in various stages of immune cell development and PRR pathways, with direct functional effects through the regulation of chromatin structure, gene expression, and mRNA translation (Zhang and Cao, 2016). Notable examples are *NRAV*, which suppresses ISG transcription; *NEAT1*, which promotes inflammasome activation; and *BISPR*, which stimulates ISG production (Carpenter et al., 2014).

We found high correlations between the expression of 254 candidate lncRNAs and their related 702 protein-coding genes (Pearson correlation test, $p < 0.01$), adding evidence to the idea of the regulatory potential of individual lncRNAs in the cooperative dynamics of gene activation (Tables S3A and S3B). We found distinct candidate lncRNAs, most of which do not have functional annotation, related to one group or the other ($n = 155$ and $n = 67$ for the high and low response groups, respectively) (Peng et al., 2010). A correlation matrix revealed significant (Pearson correlation test, $p < 0.01$) associations between lncRNAs and their proximal protein-coding gene (Figure 3B). Multiplex analysis of

candidate lncRNA-related gene expression examining 14 noncoding RNAs (*AC012645.3, SPDYE11, CIRBP-AS1, AL022311.1, SNHG9, AC012615.4, EP300-AS1, A1BG-AS1, AC136475.3, NALT1, AC245052.4, AC069281.2, AC138028.1,* and *AC005387.1*) revealed a substantial correlation with 30 protein-coding genes across the two groups (Table S3B). Protein-protein interaction (PPI) enrichment analysis revealed candidate lncRNAs that are potential local regulators of DEGs in peptide chain elongation, the metabolism of RNA, and G protein coupled receptor (GPCR) ligand binding (Figure 3C). Notably, *DDX49*, *RPS5*, *RSP15*,and *TBL3* ($\log_{10}p < -1.90$) are involved in the major pathway of rRNA processing in the nucleolus and cytosol. The significant overlap of lncRNAs and DTE suggests that the lncRNAs may have a role in RNA processing and gene transcription, involving diverse mechanisms that include splicing regulation (Engreitz et al., 2016).

We transformed the data to a topographical overlap matrix using Weighted Gene Co-expression Network Analysis (WGCNA) for hierarchical clustering of the transcripts to identify clusters of highly correlated candidate lncRNA and protein-coding genes (Figure 3D; Tables S4A and S4B) (Langfelder and Horvath, 2008). The most substantial evidence for highly connected DEGs and lncRNAs with similar patterns of expression (cutoff value of 0.2, $\log_{10}p < -1.34$) was found in the annotation networks groups for the metabolism of RNA, type I IFN signaling, regulation of cytokine production, and myeloid cell activation functions. Notable examples of significant lncRNA hub genes included *NRIR*, which produces a functional lncRNA transcript for a negative regulator of the IFN response, *AC004551.1*, which is antisense to *OAS1*, *OAS2*, and *OAS3*; *AL445490.1*, which is antisense to *IFI6*; and *AP001610.2*, which is antisense to the reverse strand of *MX1* (Figure 3D). The noncoding genes had a more comprehensive representation. The incomplete overlap between the two networks is in keeping with the observation that splicing provides another layer of gene regulation for the innate immune-regulated cellular machinery (Ip et al., 2007).

Alternative splicing

Post-transcriptional regulation of gene expression, including differential splicing (DS) of precursor mRNAs (pre-mRNAs), is essential for controlling cell-type- and tissue-specific expression of the immune response (Braunschweig et al., 2013; Carpenter et al., 2014). Alterations in intron usage through exon skipping, alternative 5' exon inclusion, alternative 3' splice site usage, and intron retention can increase the complexity of gene expression and lead to distinct functional roles in regulating a coordinated, IFN-regulated gene expression program (Braunschweig et al., 2013; Carpenter et al., 2014).

We used LeafCutter to identify variable splicing events and detect shifts in splicing patterns (Li et al., 2018; Reyes and Huber, 2018). We distinguished eight sets of significant (FDR <0.05) DS intron clusters across the four-point time course for the two groups (Figures S5A–S5C; Tables S5A and S5B; supplemental information). The annotated changes in intron usage identified exon skipping, alternative 5' exon inclusion, and alternative 3' splice site usage (Table S5A; supplemental information). The overlap among DEGs (both upregulated or downregulated) and DTE across the two groups was significant (hypergeometric test of enrichment, $p < 2.2 \times 10^{-16}$, Figure 4A). DS overlapped

significantly with DTE 2 days after the first and second dose, suggesting that splicing changes are coordinated (Baralle and Giudice, 2017; Bhatt et al., 2012). There was a substantial correlation in DS measured by differences in intron excision events and, as a result, overlap among DS genes (Figure 4B). DS genes exhibited significant enrichment (FDR <0.05) for regulation of mRNA processing, RNA splicing, RNA catabolic processes, and transmembrane receptor protein kinase signaling (Figures 4C and S5C). We validated the differences in alternative splice isoform expression for *IFIT1* and *IFIT3* using real-time RT-PCR (Pearson correlation test, $p < 0.05$; data not shown).

DS is regulated by RNA-binding factors that interact with *cis*-acting RNA elements to affect cellular spliceosome assembly at nearby splice sites and can vary across immune cell types (Amit et al., 2009). The relative abundance of the multiple transcript isoforms is influenced further by expression levels of *trans*-acting splicing factors (Baralle and Giudice, 2017). Removal of introns from nascent RNA and exon ligation catalyzed by the spliceosome can determine splice site selection, influence transcript isoform profiles and differential gene expression, and regulate transcription (Shi, 2017). We measured the expression levels of 230 candidate regulatory splicing factors identified through PPI networks and their GO term enrichments and found high correlations between splicing factor expression and DS events (Figures S5D–S5F and Table S5C) (Hegele et al., 2012). We observed substantial enrichment for many transcription factors and RNA-binding proteins, which are primary regulators of gene expression and RNA processing essential for gene function (Figure S5E and Table S5C) (Amit et al., 2009). Hierarchical clustering exhibited changes in distinct clusters of immune genes that may be subject to transcriptional and post-transcriptional regulation, suggesting that coordinated DS events can regulate their cognate biological processes (Figure S5D).

Gene correlation networks

After regressing out all covariates, we performed a correlation-based expression analysis to construct correlation networks using WGCNA to capture biologically significant gene expression patterns (Figure 5A and Table S4A). We then partitioned the complex networks into sets of highly correlated temporally co-expressed genes or modules using hierarchical clustering (Table S4B). After the first dose, the relatedness dendrograms in the hierarchical clustering identified nine modules describing temporal and diverse GO term pathway activities. We clarified the module interconnection patterns by combining the module correlations that overlap within and between groups to identify community structures (Figure 5B) (Yip and Horvath, 2007). Though we built the protein-coding and noncoding gene correlation networks independently, the separate networks largely mirrored similar biological processes in the hierarchical clustering.

The H8-L8 community module captured the biology of the overlap for known elements of type I IFN signaling mechanisms. These include the cytoplasmic PRRs, the transcriptional regulators, and the molecular pathways associated with the type I IFN response, the cellular response to viral infection, and cytokine signaling. We observed enrichment for *DDX58*, *IFIH1*, and *STAT1*, and several ISG effectors, including *HERC5*, *MX1*, and *USP18*, that have direct antiviral activities. The top hub genes represented in the

transcriptional regulatory networks with the highest degree of connectivity within a module included the IRF7 and STAT1 transcription factors (Figure 5C). Those genes exhibited the strongest correlation across all modules (Pearson correlation test, $p < 0.01$). Notably, the innate immunity interactome for type I IFN is associated with a robust functional antibody response shown in the network analysis (Figures 5D and S6). The protein clusters with significant correlations (confidence score >0.7) include CXCR3 chemokine receptor binding, postchaperonin tubulin folding pathway, neddylation, and transcription factor AP-1 complexes. These findings exemplify that independent gene- and isoform-level networks refine the cellular processes and interactions.

The modular transcriptional repertoire

To resolve the functionally related and co-regulated genes expressed in response to TAK-003 vaccination, we performed gene set enrichment analysis (GSEA) on genes annotated in a modular transcriptional repertoire according to their biological functions (Tables S6A and S6B) (Chaussabel and Baldwin, 2014; Gaucher et al., 2008; Li et al., 2014; Querec et al., 2009). The modular transcriptional signatures of vaccine recipients after the first dose identified a conserved pattern of gene expression related to antiviral innate immunity and type I IFN response, including M127, M150, and M75; myeloid cell/inflammation, including M11, M32.0, and M32.1; and cell cycle activity, including M4.0, M4.1, and M32.2, blood transcription modules (BTMs) (Figures 6, 7, and S7; Table S6).

The first dose led to a substantially broader innate immune response in the high response group than the low response group. In addition to the TAK-003-induced IFN-response modules, the high response group had significant increases (FDR <0.05; normalized enrichment score [NES]) in gene expression in cell-type-related activation and cell cycle activity-related modules at day 2. It is notable that the first dose led to significant differences in the directionality of enrichment (upregulated versus downregulated NES) for 17 modules with unassigned transcriptional program activity, including intracellular transport (M147), Golgi membrane (M237), and inositol phosphate metabolism (M129), and other modules that do not have a functional assignment (Figures 6 and S7). This pattern contrasts with the module activity at day 7, when most BTMs exhibited shared enrichment patterns across the two groups. Notable exceptions are modules that had significant increases for antigen presentation (M95) and dendritic cell activation (M168) in the high response group, and chemokine cluster (M27.1), and cell division in stimulated T cells (M4.6) and signaling in T cells (M35.0) in the low response group.

On day 2, the first dose induced a transcriptional signature of B cell transmembrane signaling receptor and co-stimulatory molecule activity in the high response group (Figure 7A). The fold changes of the genes within those modules, including M47.2, M47.3, M69, M182, and M9, showed an increase in *CD79B*, which is related to the B cell antigen receptor complex, and *IGKV4* and *IGKV1–5*, which are related to the V region of the variable domain of immunoglobulin light chains that take part in antigen recognition. One module, M152, which contains genes related to B cell function and inflammatory response, was negatively enriched in the low response group. In addition, the cell cycle activity-related modules that were significantly enriched, including M144, M250, M37.3, M23, and M145,

mirror the genes implicated in cell cycle phase transitions and the regulation of DNA replication dynamics of B cells (Braun et al., 2018; Franco et al., 2013; Li et al., 2009, 2014; Nakaya et al., 2011; Obermoser et al., 2013). This pattern of module activity for B cell signatures was not captured in the low response group on day 4 or 7, suggesting that the trajectory of rapid plasma cell development and expansion was not delayed (Figure 6).

The high response group tended to have more significant changes in the myeloid cell (M37.0, M11.0, S4, M118, M37.1, and M16) and dendritic/antigen-presenting cell (M47.0, M47.1, M47.2, M69, M182, and M9) modules. The average fold changes of those modules were consistent at days 4 and 7 regardless of the enrichment of modules at day 2 that differed between the two groups (Figure 7A). Collectively, these data suggest that significant early gene expression changes related to B cell activation and plasma cell development, antigen presentation by dendritic cells, and other unassigned transcriptional program modules are required for a robust functional antibody response (Figure 7B).

Compared with the modules characterizing the first dose, the second dose increased cell-type-related module expression reflective of the same, or distinct, immune processes in the high and low response groups. Significant associations with both groups were observed for monocyte signatures (M37, M11, S4, M118), platelet activation (M32.1, M32.0), TLR and inflammatory signaling (M16), neutrophil activity (M37.1), inflammatory response (M33), dendritic cell activation (M165, M95, M43.1), antigen presentation (M71), lysosomal and endosomal proteins (M139), cell division (M46), T cell activation (M7.0, M7.1, M7.3, M52), and NK cell activation (M7.2). We identified BTMs associated with the high response group, including proinflammatory dendritic cell response (M86.1), chemokines and receptors (M38), ubiquitination (M138), transmembrane transport (M87), dendritic cell activation (M64), T cell activation (M5.1), and cell cycle (M167) modules. BTMs that showed association with the low response group, including lipopolysaccharide-activated dendritic cell surface signature (S11) and enriched in NK cell (M61.2) modules, emphasize the different recognition pathways of the innate immune system that can instruct the adaptive immune response.

DISCUSSION

In this study, we have performed a longitudinal RNA-seq analysis of cells, including immune cells, in blood obtained from 20 healthy adults to broaden our understanding of the early determinants of the variable immune response to the TAK-003 vaccine. Extensive analyses of gene expression, splicing, transcript isoform expression, and co-expression networks allowed the identification of protein-coding and noncoding genes with differential expression in the various immune cell types. Cell-type-related enrichment and molecular pathway analysis demonstrate that most changes in gene expression, splicing, and lncRNAs in response to vaccination involve type I IFN signaling, immunometabolism, and inflammatory response pathways. These data reveal an early transcriptional signature of antiviral innate immunity and cell-type-related activation, highlighting the significance of alternative splicing and isoform-level gene regulatory mechanisms in defining the immune response to the TAK-003 vaccine.

Post-transcriptional regulation is a critical mechanism for tuning and modulating the innate immune response (Ashraf et al., 2019; Carpenter et al., 2014). By creating multiple transcript isoforms from a single gene, differential transcript utilization expands the complexity of the proteome, leading to potential functional consequences for the immune response (Carpenter et al., 2013). The tight correlation for the coordinated changes in DEGs and DTE emphasizes the significance of differential isoform usage and pre-mRNA splicing mechanisms in post-transcriptional gene regulation and immune function (Gandal et al., 2018; Martinez and Lynch, 2013; Rotival et al., 2019; Schaub and Glasmacher, 2017). The pervasive DS isoforms and the array of functional gene groups that it affects emphasize the importance of splicing as an essential mechanism of gene regulation, allowing exons to be expressed at different levels across different immune cell types (Baralle and Giudice, 2017). Notably, the TLR and RIG-I signaling pathways are regulated by diverse transcripts for fine-tuning and modulating signal transduction (Ashraf et al., 2019; Carpenter et al., 2014). Moreover, the expression levels of candidate regulatory splicing elements showed substantial enrichment for transcription factors and RNA-binding proteins that can affect RNA processing and regulation of gene expression. It is notable that DENV, like other RNA viruses, disrupts splicing to suppress host defenses (Banerjee et al., 2020; De Maio et al., 2016).

Co-expression of lncRNAs with DEGs suggests that these noncoding RNAs are involved in similar biological functions and may dynamically regulate gene expression. Functional elements in the noncoding genome, including upstream molecular regulators, promoters, and distant-acting enhancers, can control innate immune defense, inflammation, and immune cell development (Chen et al., 2017; Consortium et al., 2007; Wang and Chang, 2011). Notable examples of lncRNAs with proven immunity roles include *NRAV*, which negatively regulates ISG expression, *NEAT1*, which affects both RIG-I signaling and the cGAS-STING-IRF3 pathway, and *BISPR*, which positively regulates the expression of *BST2* (Qiu et al., 2018). We identified candidate lncRNAs as potential *cis*-acting regulatory elements for ISGs with roles in immunity and inflammation, providing good evidence that they may contribute toa coordinated, IFN-regulated gene expression program. We cannot exclude other mechanisms, such as alternative gene activation states in innate immune cells or many environmental and genetic factors that may affect functional diversity in gene expression patterns. The substantial overlap between lncRNAs and DS signals highlights pre-mRNA splicing as an essential mechanism for post-transcriptional regulation of gene expression, allowing immune cell functions to proceed more rapidly (Ashraf et al., 2019; Carpenter et al., 2014).

Application of a modular repertoire framework to the analysis of transcriptional data reduces the bias of using predefined functional pathways and increases the correlation analysis information, assigning biological significance of the co-expressed genes to the observed trends (Gaucher et al., 2008; Querec et al., 2009). We identified a substantially broader innate immune response with increased enrichment of BTMs at day 2 significantly associated with a higher magnitude functional antibody response. Notably, a signature associated with the upregulation of B cell and cell cycle activity-related modules reflects the B cells' clonal expansion and differentiation into plasma cells. Whether the altered early innate immune response affects the kinetics or regulation of effector B cell priming will

require further investigation. Most modules exhibited shared enrichment patterns across the two groups at days 4 and 7, suggesting that the generation and function of plasma cells are not delayed in the low response group. Other modules implicate antigen presentation by dendritic cells, NK cell or T cell feedback, and cell cycle phase transition. In addition to identifying an early transcriptional signature of antiviral innate immune responses, we recognized 17 modules with unassigned transcriptional program activity that may help identify previously unexamined grouping of genes of immunological relevance.

In conclusion, we provide insights into the molecular pathways and immune cell types involved, emphasizing the importance of splicing and isoform-level gene-regulators of immunity elicited by a live attenuated DENV vaccine. These results illustrate how genes involved in regulating the innate immune response have a substantial role in conditioning adaptive immunity that could be exploited to elicit the production of functional antibodies and inform vaccine formulation optimization. These results also have important implications for clinical trials. They suggest an early correlate of vaccine immunogenicity to identify those who respond poorly to primary vaccination and inform mechanisms by which the vaccine generates the adaptive immune responses.

Limitations of the study

Although neutralizing antibodies are the principal mechanistic correlate of protection against yellow fever and Japanese encephalitis, the precise immune parameters that confer protection against DENV are unknown (Plotkin, 2010). Because of the limited blood volumes, we did not assess effector T cell responses that have essential roles in mediating immunity to DENV or plasma levels of cytokines, chemokines, and acutephase proteins. Another constraint is that we did not characterize cell identity and state or dissect cell-type-related molecular circuitry of single cells, as those multimodal omics data were unavailable. Nevertheless, we merged gene expression, transcript structure, and regulatory variation data using statistical and computational methods for an in-depth understanding of gene and isoform expression changes, differential expression of the noncoding transcriptome, alternative splicing, and immune celltype-related signals. We envision those future studies can leverage larger numbers of vaccine recipients with viable cell samples at even earlier time points to understand better post-transcriptional regulation of antiviral innate immunity in controlling the antibody response.

STAR*METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to the lead contact, Steven Wolinsky (s-wolinsky@northwestern.edu)

Materials availability—This study did not generate new unique reagents.

Data and code availability

- The raw files for the datasets generated in this study are available at GEO: GSE146658. The software referenced in the paper is provided via a link to GitHub (https://github.com/eykim909/TDV_paper_Scripts).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study design—We collected and analyzed blood specimens and data from twenty healthy adults with no serologic evidence of previous flavivirus infection (mean age overall, 20 years range, 18 to 26 years; 50% female) who enrolled in phase 1, double-blind, placebo-controlled trial of a two-dose schedule of the TAK-003 vaccine candidate manufactured by Takeda Pharmaceutical Company (Osorio et al., 2014). TAK-003 is an admixture composed of DENV-2 and three chimeric viruses with the DENV-2 prM and E proteins exchanged for DENV-1, DENV-3, and DENV-4 in the DENV-2 genome. One dose of TAK-003 contained about 3.70, 3.90, 4.00, and 5.30 $\log_{10}$ plaque-forming units of DENV-1, DENV-2, DENV-3, and DENV-4, respectively. Serotype-specific serum neutralization titers were measured by microneutralization testing and vaccine plasma viremia by quantitative serotype-specific RT-PCR. The demographic characteristics of the vaccine recipients were similar in the two groups, which were well-balanced for age and sex and not confounded by ancestry (chi-squared test for a trend in proportions, $p = 0.2446$; data not shown).

METHOD DETAILS

RNA-seq—We extracted total RNA from blood collected in PAXgene Blood RNA Tubes with the use of PAXgene Blood RNA Kits (BD Biosciences) according to the manufacturers' instructions. The extracted RNA underwent DNase digestion and cleanup using RNeasy MinElute kits (Qiagen). The purified RNA samples were quantified by RNA High Sensitivity Assay kits (ThermoFisher) with a Qubit 3.0 fluorometer. We assessed RNA integrity number (RIN) and yield using RNA High Sensitivity ScreenTape on the TapeStation 2200 (Agilent Technologies). Only RNA samples with RIN greater than eight were used for downstream cDNA synthesis and library preparation. The ribosomal RNA and globin mRNA were removed. Then the enriched RNA was fragmented and primed for cDNA synthesis using TruSeq Total Stranded RNA HT kit with Ribo-Zero Globin on a Microlab STAR automated liquid handling system (Hamilton). TruSeq HT indices allowed for multiplexing. The barcoded libraries were enriched by ligation-mediated PCR for 15 cycles and purified with the Agencourt AMPure XP beads system (Beckman Coulter). We assessed the libraries for quality with a high-sensitivity DNA ScreenTape assay on the 2200 TapeStation System (Agilent) and quantity with KAPA Library Quantification Kits for Illumina platforms (Kapa Biosystems). The libraries were diluted to 2 nM and combined equimolarly in pools of 12. They were then clustered using an Illumina cBot with a HiSeq 3000/4000 paired-end cluster kit on a patterned flow cell, one pool per lane, and a HiSeq 3000/4000 SBS kit (300 cycles, Illumina) on the HiSeq 4000 sequencing platform.

We obtained more than 61 million high-quality paired-end reads (range, 61 to 80 million) for each blood sample after quality filtering. This depth of coverage is ample for robust and reproducible differential expression and splicing analyses (Mehmood et al., 2019; Shen et al., 2012). Measurements performed at the same time minimized batch effects. Technical replicates are highly correlated with one another (average Spearman rank correlation coefficient across replicates, $\rho = 0.93$). The biological replicates clustered with small dispersion distributions across all genes for equal sequencing depth and coherent genome sequence and assembly quality. We corrected for individual effect by adding the sample information as a covariate.

RNA-seq data processing—We converted the sequencer-generated base call (BCL) files to multiplexed FASTQ files, separating the individual files using bcl2fastq conversion software. We used FastQC (version 0.11.4) to calculate quality control metrics (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The FASTQ files were quality filtered using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) with the invocation fastq_quality_filter -q 30 -p 50 -v -Q 33. Reads were mapped to the GRCh38.p10 (Ensembl release 90) reference genome using HISAT2 (Kim et al., 2015). We used SAMtools (version 0.1.19) to sort and convert the SAM files to BAM files (Li et al., 2009). We then assembled the aligned sequences into potential transcripts, including splice variants, and the abundance of the transcripts and genes in each sample quantified using StringTie (version 1.2.2) (Pertea et al., 2015).

Population stratification analysis—Variants were inferred from the processed RNA-seq data using GATK HaplotypeCaller (version 3.5.0) and filtered by setting to missing any genotype with a genotype quality threshold for calling variants to ≤20 (https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-variant-discovery-SNPs-Indels-). Ancestry estimates were obtained from running the model-based clustering program ADMIXTURE (version 1.3) with the *k* value for this dataset populations (based on low cross-validation error) fixed for the number of postulated ancestral on these vaccine recipients with 1000 Genomes Project Native American, African, Asian, and European superpopulations as reference data (Alexander et al., 2009; Galanter et al., 2012; Purcell et al., 2007).

Digital multiplexed gene expression analysis—The expression of immune response-related genes was measured with a nanoString custom nCounter panel for 347 target protein-coding and noncoding genes. Copies of individual RNA molecules were labeled with gene-specific barcodes, and each one was counted with the nCounter system (nanoString Technologies). The gene expression data were normalized to negative and positive control lanes and reference gene lanes on the panel using nSolver Analysis software (version 4.0). The significantly differentially expressed genes were adjusted for multiple statistical comparisons in the analysis using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

RT-qPCR validation of differential isoform usage—Validation of DS was performed using isoform-specific primers that bridge a region that can be used to differentiate between

isoforms for known immune response genes. We used the lactate dehydrogenase A gene for internal normalization with the transcript primer set. RT-qPCR primers are detailed in the key resources table. All RT-qPCR data were analyzed in technical duplicate.

Differential gene expression analysis—DEGs and differentially expressed transcripts normalized on the fragments per kilobase of transcript per million (FPKM) mapped reads were identified across all samples with DESeq2 multivariate negative binomial generalized linear model (version 1.10.1) (Love et al., 2014) and a likelihood ratio test (Anders and Huber, 2010; Frazee et al., 2015; Robinson et al., 2010). Principal component analysis (PCA) identified the top 500 upregulated and downregulated genes found in at least 60% of the vaccine recipients with the highest coefficient of variation (standard deviation to the mean) after variance stabilizing transformation. We accounted for noise or changes that might have occurred after some lag time despite being time matched. The p values were adjusted for multiple comparisons (Benjamini and Hochberg, 1995).

Gene expression and antibody correlation analyses—To investigate the discriminate predictive capability of each DEG, we performed a Pearson correlation analysis of the $\log_2$FC enrichments. The correlation analysis provided how genes behave differently in each group and whether the difference is far from the null. The direction of the coefficient depends on relative changes of $\log_2$FC between the two groups, and it shows whether a gene exhibited greater or lower $\log_2$FC in the high than the low response group. The strength of correlation indicates the extent of changes of $\log_2$FC from each group and if the extent of changes is statistically different. Note that the direction and the strength of correlation do not necessarily represent characteristics of a particular group. A positive correlation coefficient does not always represent a correlation with the high response group. Instead, those two quantities depict relative expression level changes from baseline in each group.

lncRNA quantification and differential expression analysis—We employed the same pipeline for quantifying and differential expression testing of known genes, using the lncRNA Ensembl gene annotation from the Ensembl regulatory build (Guttman et al., 2009).

Co-expression analysis of lncRNAs and neighboring genes—We annotated the chromosomal coordinates of differentially expressed lncRNA and mRNA transcripts aligned to the GRCh38.p10 (Ensembl release 90) reference genome with the Ensembl BioMart tool (Durinck et al., 2009; Guberman et al., 2011). For each differentially expressed lncRNA, neighboring RNAs within 500 kb upstream or downstream of transcription start and stop sites were extracted using a custom R script. We performed a Pearson correlation test on each candidate lncRNA's expression level. We defined significant associations as having a p value <0.05.

Alternative splicing analysis—We used LeafCutter to count aggregate changes in intron usage that affect exons shared by multiple transcripts, regardless of genome annotation, and group introns into clusters of potential splicing disruption rather than exon skipping (Li et al., 2018; Reyes and Huber, 2018). We called intron clusters supported by at least 50 split reads across all samples and have a maximum length of 500 kb, resulting in 78,993 intron clusters, and then filtered the intron clusters to have at least one read in 5 or more samples.

We identified exon-inclusion ratio in terms of PSI values of greater than or equal to 20%. Pairwise differential splicing analysis was then performed by comparing its corresponding value at baseline (day 0). For each pairwise comparison, intron clusters were required to have at least 20 reads in at least three samples in each comparison. The p values were corrected for multiple comparisons. We defined a significant cluster as having an FDR <0.05. We performed a hypergeometric test of enrichment for significant overlap among DEGs and DTE ($|\log_2 FC| > 1$, FDR <0.05) and DS (likelihood ratio test, FDR <0.05). The background pool was the number of genes filtered from the gene count matrix.

Hierarchical clustering of splicing factors—For the correlation clustering of splicing factors, gene expression of 230 splicing factors was extracted from the gene count data (Hegele et al., 2012). We calculated the Pearson correlation coefficients for all possible combinations of splicing factors and clustered those correlations using unsupervised clustering based on Euclidean distance and complete linkage.

Pathway, process, PPI enrichment, and network analyses—For the pathway and process enrichment analysis, we carried out the analyses on each significant list of genes with the open-access ontology sources for Gene Set (GO) and Biochemical Pathway (KEGG, Reactome, Canonical Pathways, Panther, Pathview, Metascape, and CORUM) databases (Luo and Brouwer, 2013; Ogata et al., 1999; Reimand et al., 2019; Thomas et al., 2003; Zhou et al., 2019). We chose the most statistically significant term within a cluster as the one representing the cluster. For the GSEA method (Subramanian et al., 2005), we ranked all genes by $\log_2 FC$ by comparing its corresponding value at baseline. We then compared the distribution of gene ranks from the gene set to the other genes using the NES (one-sided Kolmogorov-Smirnov statistic). The biological interpretation of the lists of genes was performed with ClueGO (Bindea et al., 2009), which integrates GO and KEGG functional networks and BioCarta pathways to create a functionally organized GO/pathway term network. Network visualization of the functionally grouped terms was performed in Cytoscape (version 3.2.1.) (Shannon et al., 2003), where each node represented an enriched term and colored by its cluster identifier annotated by their p value.

We used a pipeline written with a Python script for network analysis and visualized it with Cytoscape applications. First, we extracted DEGs ($|\log_2 FC|>1$, adjusted p value <0.05), then constructed a network based on interactions between the proteins coded for by these genes across days 2, 4, and 7. We selected the STRING - Human Protein Links - High Confidence (Score R0.7) interactome (Reimand et al., 2007) to extract PPIs between DEGs (Szklarczyk et al., 2017). This network was then imported to Cytoscape, where we identified densely connected areas of the network (that is, community detection) using OSLOM (Lancichinetti et al., 2011), via the CDAPS (Singhal et al., 2020) app in Cytoscape. This new network featured communities of proteins represented by network nodes that were annotated using gProfiler (Reimand et al., 2007) through the functional enrichment feature of CDAPS. The edges signify the containment of one community by another.

Weighted gene co-expression network analysis (WGCNA)—We constructed gene co-expression networks with the WGCNA R package (Langfelder and Horvath, 2008). We used the counts per million normalizations with measured values in at least 60% of all

the samples analyzed (n = 17,588 and n = 17,646 for the high and low response groups, respectively). We chose the cutoff to eliminate noise from very rarely detected genes or whose expression was difficult to correlate with other genes due to the few samples measured. An initial hierarchical clustering with the expression data confirmed that there were no outliers. We used a scale-free topology criterion to choose the soft threshold power β. The β, six, resulted in approximate scale-free topology using a free topology fitting index. A signed hybrid network was constructed using the adjacency matrix of the biweight mid-correlation of all pairwise comparisons of gene expression values, that is, the matrix of connection strengths by using the soft threshold power. Modules were identified using the cutreeDynamic function, setting the branch height cutoff to 0.99, the minimum module size to 50, and the merging cut height, that is, the dissimilarity threshold below which separate modules would be merged, to 0.35.

We calculated eigengene-based connectivity, also known as module membership kME measures for a particular gene within a given non-preserved module to identify highly connected or hub genes representative of the module's overall function with a high likelihood to be critical components within the module. We determined module membership by correlating the gene and the module eigengene value (the first principal component of each module output value). This value quantified how close a gene is within a given module, and we applied this measure to detect the hub genes. Genes with absolute kME membership ≥0.9 were considered hub genes to the respective module (Langfelder and Horvath, 2008).

Identification of module communities—We calculated pairwise biweight mid-correlations between module eigengenes, determining correlation significance with the Student's t-test. We also determined overlap in gene content between modules in the two groups, calculating the significance of this overlap using a one-tailed Fisher's exact test. Both correlation and overlap p values were adjusted for multiple comparisons by control of FDR. We created correlation and overlap module networks, defining module correlations or overlaps with FDR <0.01 as connections between individual modules. The module communities were detected based on greedy optimization of modularity implemented in igraph (https://igraph.org/r/).

Modular transcriptional signature and antibody correlation analyses—GSEA (Subramanian et al., 2005) was performed using modules that contain at least ten genes (Li et al., 2014). The dimension reduction circumvents the mixed expression signals that reduce the sensitivity of the analysis. We ranked transcripts according to their $\log_2$FC by comparison with their corresponding value at baseline. The modules were grouped further into seven families (Braun et al., 2018). Housekeeping genes enabled the deconvolution of cell-type-related gene expression signatures to estimate the celltype-related frequency and expression signals. We determined the module's correlation by calculating each module's NES. We then compared the NES obtained by GSEA for the genes in each module with its corresponding value at each time point, controlling the proportion of false positives. Significant modules were defined using p <0.05 after adjustment made for multiple comparisons (FDR <0.05). We calculated Pearson's correlation of NES with antibody response. Significant correlation was defined as having a p value <0.05. We compared the

NES obtained by GSEA for each transcription module to its corresponding value at baseline to identify predictive transcriptional signatures.

QUANTIFICATION AND STATISTICAL ANALYSIS

Information on specific quantification methods is described in associated Method details, or main text. Statistical tests were performed using R software (version 4.1.1), parameters, statistical test, and significance, are reported in Figures and Figure legends or Method details section. We did not use statistical tests to predetermine sample size. The experiments were not randomized, and we were not blinded to distribution during experiments and outcome measurement.

ADDITIONAL RESOURCES

The trial was conducted in a region of Columbia, where dengue is endemic, in accordance with the Declaration of Helsinki principles, Good Clinical Practice guidelines, and applicable local regulations. Written informed consent was obtained from all participants before enrollment. Full details of the trial design, conduct, oversight, and analyses can be found in (Osorio et al., 2014). This study is registered with ClinicalTrials.gov identifier NCT01224639.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This study was supported by NIH grants U19 AI118610-01, U19 AI135972, and U01 AI035039, Federal Contract HHSN272201200003I Task Order IA-013, and the James B. Pendleton Charitable Trust. We thank Aleksandar Zivkovic for providing technical support.

REFERENCES

Alexander DH, Novembre J, and Lange K (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664. 10.1101/gr.094052.109. [PubMed: 19648217]

Amit I, Garber M, Chevrier N, Leite AP, Donner Y, Eisenhaure T, Guttman M, Grenier JK, Li W, Zuk O, et al. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science 326, 257–263. 10.1126/science.1179050. [PubMed: 19729616]

Anders S, and Huber W (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106. 10.1186/gb-2010-11-10-r106. [PubMed: 20979621]

Ashraf U, Benoit-Pilven C, Lacroix V, Navratil V, and Naffakh N (2019). Advances in analyzing virus-induced alterations of host cell splicing. Trends Microbiol. 27, 268–281. 10.1016/j.tim.2018.11.004. [PubMed: 30577974]

Banerjee AK, Blanco MR, Bruce EA, Honson DD, Chen LM, Chow A, Bhat P, Ollikainen N, Quinodoz SA, Loney C, et al. (2020). SARSCoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. Cell 183, 1325–1339.e1. 10.1016/j.cell.2020.10.004. [PubMed: 33080218]

Baralle FE, and Giudice J (2017). Alternative splicing as a regulator of development and tissue identity. Nat. Rev. Mol. Cell Biol. 18, 437–451. 10.1038/nrm.2017.27. [PubMed: 28488700]

Benjamini Y, and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodological) 57, 289–300. 10.1111/j.2517-6161.1995.tb02031.x.

Bhatt DM, Pandya-Jones A, Tong AJ, Barozzi I, Lissner MM, Natoli G, Black DL, and Smale ST (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell 150, 279–290. 10.1016/j.cell.2012.05.043. [PubMed: 22817891]

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, et al. (2013). The global distribution and burden of dengue. Nature 496, 504–507. 10.1038/nature12060. [PubMed: 23563266]

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, and Galon J (2009). ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25, 1091–1093. 10.1093/bioinformatics/btp101. [PubMed: 19237447]

Biswal S, Reynales H, Saez-Llorens X, Lopez P, Borja-Tabora C, Kosalaraksa P, Sirivichayakul C, Watanaveeradej V, Rivera L, Espinoza F, et al. (2019). Efficacy of a tetravalent dengue vaccine in healthy children and adolescents. N. Engl. J. Med. 381, 2009–2019. 10.1056/NEJ-Moa1903869. [PubMed: 31693803]

Bournazos S, Gupta A, and Ravetch JV (2020). The role of IgG Fc receptors in antibody-dependent enhancement. Nat. Rev. Immunol. 20, 633–643. 10.1038/s41577-020-00410-0. [PubMed: 32782358]

Braun RO, Brunner L, Wyler K, Auray G, Garcia-Nicolas O, Python S, Zumkehr B, Gaschen V, Stoffel MH, Collin N, et al. (2018). System immunology-based identification of blood transcriptional modules correlating to antibody responses in sheep. NPJ Vaccines 3, 41. 10.1038/s41541-018-0078-0. [PubMed: 30302283]

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, and Blencowe BJ (2013). Dynamic integration of splicing within gene regulatory pathways. Cell 152, 1252–1269. 10.1016/j.cell.2013.02.034. [PubMed: 23498935]

Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, Byron M, Monks B, Henry-Bezy M, Lawrence JB, et al. (2013). A long noncoding RNA mediates both activation and repression of immune response genes. Science 341, 789–792. 10.1126/science.1240925. [PubMed: 23907535]

Carpenter S, Ricci EP, Mercier BC, Moore MJ, and Fitzgerald KA (2014). Post-transcriptional regulation of gene expression in innate immunity. Nat. Rev. Immunol. 14, 361–376. 10.1038/nri3682. [PubMed: 24854588]

Chaussabel D, and Baldwin N (2014). Democratizing systems immunology with modular transcriptional repertoire analyses. Nat. Rev. Immunol. 14, 271–280. 10.1038/nri3642. [PubMed: 24662387]

Chen YG, Satpathy AT, and Chang HY (2017). Gene regulation in the immune system by long noncoding RNAs. Nat. Immunol. 18, 962–972. 10.1038/ni.3771. [PubMed: 28829444]

Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447, 799–816. 10.1038/nature05874. [PubMed: 17571346]

De Maio FA, Risso G, Iglesias NG, Shah P, Pozzi B, Gebhard LG, Mammi P, Mancini E, Yanovsky MJ, Andino R, et al. (2016). The dengue virus NS5 protein intrudes in the cellular spliceosome and modulates splicing. PLoS Pathog. 12, e1005841. 10.1371/journal.ppat.1005841. [PubMed: 27575636]

Deng SQ, Yang X, Wei Y, Chen JT, Wang XJ, and Peng HJ (2020). A review on dengue vaccine development. Vaccines (Basel) 8, 63. 10.3390/vaccines8010063.

Durinck S, Spellman PT, Birney E, and Huber W (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. 4, 1184–1191. 10.1038/nprot.2009.97. [PubMed: 19617889]

Engreitz JM, Haines JE, Perez EM, Munson G, Chen J, Kane M, McDonel PE, Guttman M, and Lander ES (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. Nature 539, 452–455. 10.1038/nature20149. [PubMed: 27783602]

Franco LM, Bucasas KL, Wells JM, Nino D, Wang X, Zapata GE, Arden N, Renwick A, Yu P, Quarles JM, et al. (2013). Integrative genomic analysis of the human immune response to influenza vaccination. Elife 2, e00299. 10.7554/eLife.00299. [PubMed: 23878721]

Frazee AC, Pertea G, Jaffe AE, Langmead B, Salzberg SL, and Leek JT (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat. Biotechnol. 33, 243–246. 10.1038/nbt.3172. [PubMed: 25748911]

Frishberg A, Peshes-Yaloz N, Cohn O, Rosentul D, Steuerman Y, Valadarsky L, Yankovitz G, Mandelboim M, Iraqi FA, Amit I, et al. (2019). Cell composition analysis of bulk genomics using single-cell data. Nat. Methods 16, 327–332. 10.1038/s41592-019-0355-5. [PubMed: 30886410]

Galanter JM, Fernandez-Lopez JC, Gignoux CR, Barnholtz-Sloan J, Fernandez-Rozadilla C, Via M, Hidalgo-Miranda A, Contreras AV, Figueroa LU, Raska P, et al. (2012). Development of a panel of genome-wide ancestry informative markers to study admixture throughout the Americas. PLoS Genet. 8, e1002554. 10.1371/journal.pgen.1002554. [PubMed: 22412386]

Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, Won H, van Bakel H, Varghese M, Wang Y, et al. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science 362, eaat8127. 10.1126/science.aat8127. [PubMed: 30545856]

Garcia-Sastre A (2017). Ten strategies of interferon evasion by viruses. Cell Host Microbe 22, 176–184. 10.1016/j.chom.2017.07.012. [PubMed: 28799903]

Gaucher D, Therrien R, Kettaf N, Angermann BR, Boucher G, Filali-Mouhim A, Moser JM, Mehta RS, Drake DR 3rd, Castro E, et al. (2008). Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. J. Exp. Med. 205, 3119–3131. 10.1084/jem.20082292. [PubMed: 19047440]

Gibbons RV, Kalanarooj S, Jarman RG, Nisalak A, Vaughn DW, Endy TP, Mammen MP Jr., and Srikiatkhachorn A (2007). Analysis of repeat hospital admissions for dengue to estimate the frequency of third or fourth dengue infections resulting in admissions and dengue hemorrhagic fever, and serotype sequences. Am. J. Trop. Med. Hyg. 77, 910–913. [PubMed: 17984352]

Guberman JM, Ai J, Arnaiz O, Baran J, Blake A, Baldock R, Chelala C, Croft D, Cros A, Cutts RJ, et al. (2011). BioMart Central Portal: an open database network for the biological community. Database (Oxford) 2011, bar041. 10.1093/database/bar041. [PubMed: 21930507]

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 223–227. 10.1038/nature07672. [PubMed: 19182780]

Hegele A, Kamburov A, Grossmann A, Sourlis C, Wowro S, Weimann M, Will CL, Pena V, Luhrmann R, and Stelzl U (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell 45, 567–580. 10.1016/j.molcel.2011.12.034. [PubMed: 22365833]

Holmes EC, and Twiddy SS (2003). The origin, emergence and evolutionary genetics of dengue virus. Infect Genet. Evol. 3, 19–28. 10.1016/s1567-1348(03)00004-2. [PubMed: 12797969]

Hur S (2019). Double-stranded RNA sensors and modulators in innate immunity. Annu. Rev. Immunol. 37, 349–375. 10.1146/annurev-im-munol-042718-041356. [PubMed: 30673536]

Ip JY, Tong A, Pan Q, Topp JD, Blencowe BJ, and Lynch KW (2007). Global analysis of alternative splicing during T-cell activation. RNA 13, 563–572. 10.1261/rna.457207. [PubMed: 17307815]

Ivashkiv LB, and Donlin LT (2014). Regulation of type I interferon responses. Nat. Rev. Immunol. 14, 36–49. 10.1038/nri3581. [PubMed: 24362405]

Iwasaki A, and Medzhitov R (2015). Control of adaptive immunity by the innate immune system. Nat. Immunol. 16, 343–353. 10.1038/ni.3123. [PubMed: 25789684]

Katzelnick LC, Fonville JM, Gromowski GD, Bustos Arriaga J, Green A, James SL, Lau L, Montoya M, Wang C, VanBlargan LA, et al. (2015). Dengue viruses cluster antigenically but not as discrete serotypes. Science 349, 1338–1343. 10.1126/science.aac5017. [PubMed: 26383952]

Kim D, Langmead B, and Salzberg SL (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360. 10.1038/nmeth.3317. [PubMed: 25751142]

Kotliarov Y, Sparks R, Martins AJ, Mule MP, Lu Y, Goswami M, Kardava L, Banchereau R, Pascual V, Biancotto A, et al. (2020). Broad immune activation underlies shared set point signatures for vaccine responsiveness in healthy individuals and disease activity in patients with lupus. Nat. Med. 26, 618–629. 10.1038/s41591-020-0769-8. [PubMed: 32094927]

Lancichinetti A, Radicchi F, Ramasco JJ, and Fortunato S (2011). Finding statistically significant communities in networks. PLoS One 6, e18961. 10.1371/journal.pone.0018961. [PubMed: 21559480]

Langfelder P, and Horvath S (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559. 10.1186/1471-2105-9-559. [PubMed: 19114008]

Lazear HM, Schoggins JW, and Diamond MS (2019). Shared and distinct functions of type I and type III interferons. Immunity 50, 907–923. 10.1016/j.immuni.2019.03.025. [PubMed: 30995506]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. 10.1093/bioinformatics/btp352. [PubMed: 19505943]

Li S, Rouphael N, Duraisingham S, Romero-Steiner S, Presnell S, Davis C, Schmidt DS, Johnson SE, Milton A, Rajam G, et al. (2014). Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat. Immunol. 15, 195–204. 10.1038/ni.2789. [PubMed: 24336226]

Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, and Pritchard JK (2018). Annotation-free quantification of RNA splicing using LeafCutter. Nat. Genet. 50, 151–158. 10.1038/s41588-017-0004-9. [PubMed: 29229983]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]

Luo W, and Brouwer C (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29, 1830–1831. 10.1093/bioinformatics/btt285. [PubMed: 23740750]

Martinez NM, and Lynch KW (2013). Control of alternative splicing in immune responses: many regulators, many predictions, much still to learn. Immunol. Rev. 253, 216–236. 10.1111/imr.12047. [PubMed: 23550649]

Medzhitov R (2007). Recognition of microorganisms and activation of the immune response. Nature 449, 819–826. 10.1038/nature06246. [PubMed: 17943118]

Mehmood A, Laiho A, Venalainen MS, McGlinchey AJ, Wang N, and Elo LL (2019). Systematic evaluation of differential splicing tools for RNA-seq studies. Brief Bioinform. 21, 2052–2065. 10.1093/bib/bbz126.

Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, Means AR, Kasturi SP, Khan N, Li GM, et al. (2011). Systems biology of vaccination for seasonal influenza in humans. Nat. Immunol. 12, 786–795. 10.1038/ni.2067. [PubMed: 21743478]

Obermoser G, Presnell S, Domico K, Xu H, Wang Y, Anguiano E, Thompson-Snipes L, Ranganathan R, Zeitner B, Bjork A, et al. (2013). Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. Immunity 38, 831–844. 10.1016/j.immuni.2012.12.008. [PubMed: 23601689]

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, and Kanehisa M (1999). KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 27, 29–34. 10.1093/nar/27.1.29. [PubMed: 9847135]

Olkowski S, Forshey BM, Morrison AC, Rocha C, Vilcarromero S, Halsey ES, Kochel TJ, Scott TW, and Stoddard ST (2013). Reduced risk of disease during postsecondary dengue virus infections. J. Infect Dis. 208, 1026–1033. 10.1093/infdis/jit273. [PubMed: 23776195]

Osorio JE, Velez ID, Thomson C, Lopez L, Jimenez A, Haller AA, Silengo S, Scott J, Boroughs KL, Stovall JL, et al. (2014). Safety and immunogenicity of a recombinant live attenuated tetravalent dengue vaccine (DENVax) in flavivirus-naive healthy adults in Colombia: a randomised, placebo-controlled, phase 1 study. Lancet Infect Dis 14, 830–838. 10.1016/S1473-3099(14)70811-4. [PubMed: 25087476]

Peng X, Gralinski L, Armour CD, Ferris MT, Thomas MJ, Proll S, Bradel-Tretheway BG, Korth MJ, Castle JC, Biery MC, et al. (2010). Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. MBio 1, e00206. 10.1128/mBio.00206-10. [PubMed: 20978541]

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, and Salzberg SL (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33, 290–295. 10.1038/nbt.3122. [PubMed: 25690850]

Plotkin SA (2010). Correlates of protection induced by vaccination. Clin. Vaccin. Immunol. 17, 1055–1065. 10.1128/CVI.00131-10.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, and Sham PC (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. 10.1086/519795. [PubMed: 17701901]

Qiu L, Wang T, Tang Q, Li G, Wu P, and Chen K (2018). Long non-coding RNAs: regulators of viral infection and the interferon antiviral response. Front. Microbiol. 9, 1621. 10.3389/fmicb.2018.01621. [PubMed: 30072977]

Querec TD, Akondy RS, Lee EK, Cao W, Nakaya HI, Teuwen D, Pirani A, Gernert K, Deng J, Marzolf B, et al. (2009). Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat. Immunol. 10, 116–125. 10.1038/ni.1688. [PubMed: 19029902]

Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C, et al. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat. Protoc 14, 482–517. 10.1038/s41596-018-0103-9. [PubMed: 30664679]

Reimand J, Kull M, Peterson H, Hansen J, and Vilo J (2007). g:Profiler–a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res. 35, W193–W200. 10.1093/nar/gkm226. [PubMed: 17478515]

Reinhardt B, Jaspert R, Niedrig M, Kostner C, and L'Age-Stehr J (1998). Development of viremia and humoral and cellular parameters of immune activation after vaccination with yellow fever virus strain 17D: a model of human flavivirus infection. J. Med. Virol. 56, 159–167. 10.1002/(sici)1096-9071. [PubMed: 9746073]

Reyes A, and Huber W (2018). Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. Nucleic Acids Res. 46, 582–592. 10.1093/nar/gkx1165. [PubMed: 29202200]

Robinson MD, McCarthy DJ, and Smyth GK (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. 10.1093/bioinformatics/btp616. [PubMed: 19910308]

Rotival M, Quach H, and Quintana-Murci L (2019). Defining the genetic and evolutionary architecture of alternative splicing in response to infection. Nat. Commun. 10, 1671. 10.1038/s41467-019-09689-7. [PubMed: 30975994]

Schaub A, and Glasmacher E (2017). Splicing in immune cells-mechanistic insights and emerging topics. Int. Immunol. 29, 173–181. 10.1093/intimm/dxx026. [PubMed: 28498981]

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. 10.1101/gr.1239303. [PubMed: 14597658]

Shen S, Park JW, Huang J, Dittmar KA, Lu ZX, Zhou Q, Carstens RP, and Xing Y (2012). MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-seq data. Nucleic Acids Res. 40, e61. 10.1093/nar/gkr1291. [PubMed: 22266656]

Shi Y (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nat. Rev. Mol. Cell Biol. 18, 655–670. 10.1038/nrm.2017.86. [PubMed: 28951565]

Singhal A, Cao S, Churas C, Pratt D, Fortunato S, Zheng F, and Ideker T (2020). Multiscale community detection in cytoscape. PLoS Comput. Biol. 16, e1008239. 10.1371/journal.pcbi.1008239. [PubMed: 33095781]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 45, D362–D368. 10.1093/nar/gkw937. [PubMed: 27924014]

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, and Narechania A (2003). PANTHER: a library of protein families and subfamilies indexed by function. Genome Res. 13, 2129–2141. 10.1101/gr.772403. [PubMed: 12952881]

Tsang JS, Dobano C, VanDamme P, Moncunill G, Marchant A, Othman RB, Sadarangani M, Koff WC, and Kollmann TR (2020). Improving vaccine-induced immunity: can baseline predict outcome? Trends Immunol. 41, 457–465. 10.1016/j.it.2020.04.001. [PubMed: 32340868]

Uhlen M, Karlsson MJ, Zhong W, Tebani A, Pou C, Mikes J, Lakshmikanth T, Forsstrom B, Edfors F, Odeberg J, et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. Science 366, eaax9198. 10.1126/science.aax9198. [PubMed: 31857451]

Wang KC, and Chang HY (2011). Molecular mechanisms of long noncoding RNAs. Mol. Cell 43, 904–914. 10.1016/j.molcel.2011.08.018. [PubMed: 21925379]

Yip AM, and Horvath S (2007). Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinformatics 8, 22. 10.1186/1471-2105-8-22. [PubMed: 17250769]

Zhang Y, and Cao X (2016). Long noncoding RNAs in innate immunity. Cell Mol. Immunol. 13, 138–147. 10.1038/cmi.2015.68. [PubMed: 26277893]

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, and Chanda SK (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. 10, 1523. 10.1038/s41467-019-09234-6. [PubMed: 30944313]

Highlights

- Genome-wide transcriptome analysis of dengue vaccine-elicited immune responses
- A temporality of changes in lncRNA, splicing, and gene expression patterns
- Genes strongly enriched for pathways involved in antiviral innate immunity
- Cell-type-related modules exhibit a significant correlation with antibody titers

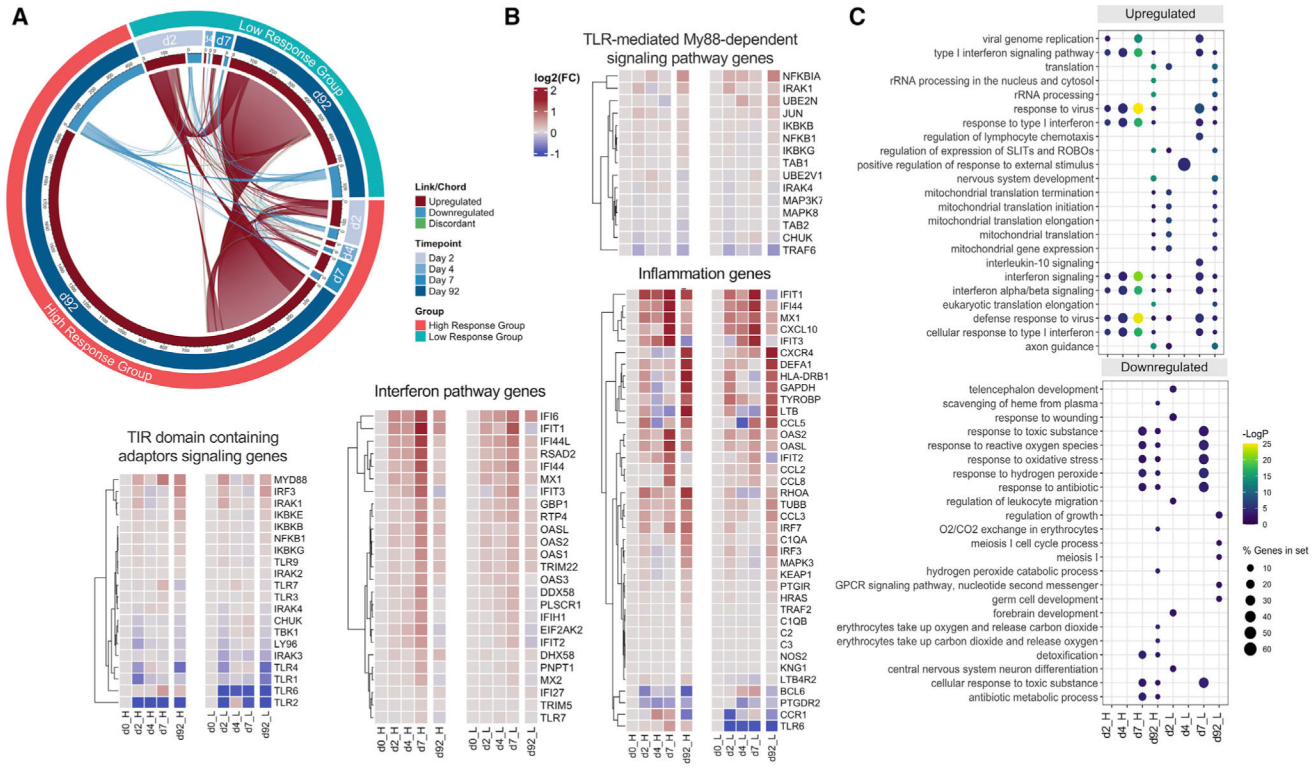


Figure 1. Immune gene expression changes in blood sampled from vaccine recipients

(A) Circos plot of the overlap across the two groups in genes expressed after each dose. The two segments of the outer circle represent the two groups. Each segment of the middle circle represents the four-point time course. Inner circle line colors represent the upregulated or downregulated genes. Lines connect upregulated (red), downregulated (blue), or conflictingly regulated (green) genes.

(B) Heatmap depicting the scaled $\log_2$FC expression for the DEGs represented in inflammation, TLR-mediated My88-dependent signaling pathway, TIR domain-containing adaptors signaling, and innate and type I IFN signaling pathways. The dendrograms alongside the heatmap cluster the DEGs based on the hierarchal clustering.

(C) Bubble maps depicting the GO term enrichments corresponding to the top DEGs. The top upregulated and downregulated DEGs enriched pathways (top and bottom panels, respectively) are shown for the two groups. Scale bar exhibits the negative $\log_{10}$p value, and the circle's size represents the percentage of genes in each set.

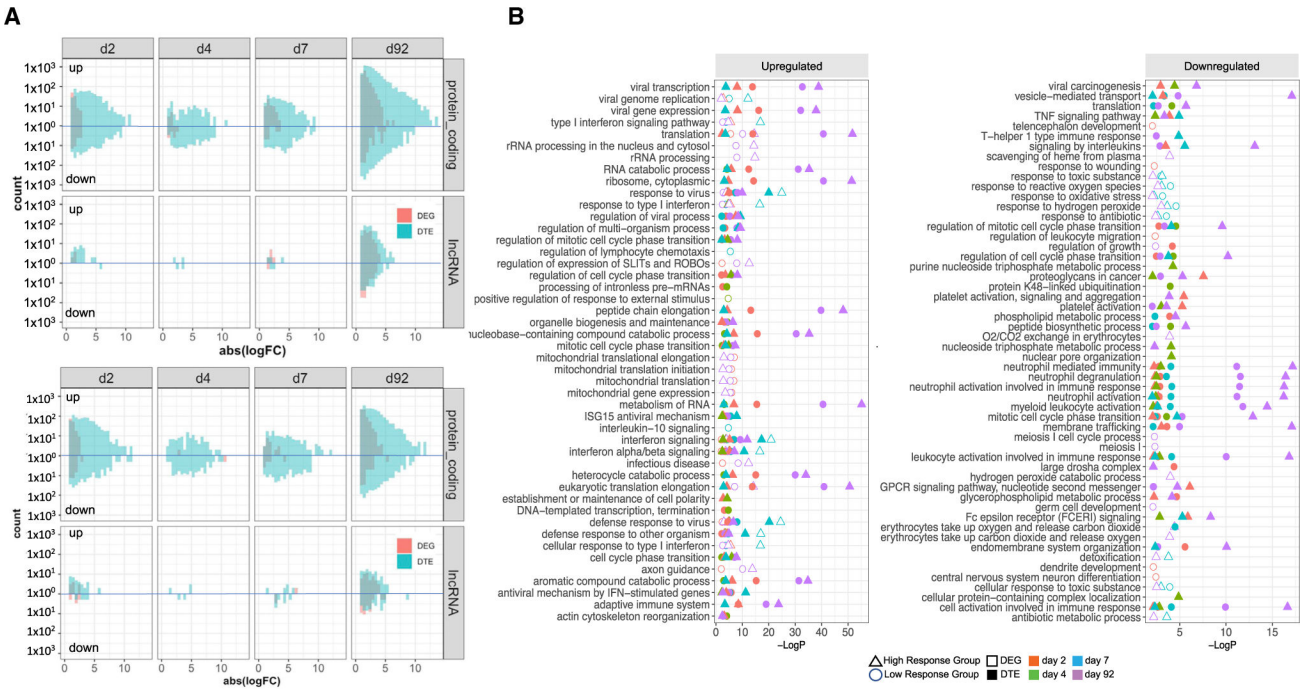


Figure 2. Gene and isoform expression changes in blood samples

(A) Differential expression effect size histograms showing the significant number of upregulated or downregulated protein-coding genes and lncRNAs. The isoform-level (DTE) changes exhibit a larger effect size than the gene-level (DEG) changes in the high and low response groups (top and bottom panels, respectively).

(B) GO term enrichments showing the top significant upregulated and downregulated DEGs and enriched pathways (left and right panels, respectively) across the four-point time course in each group.

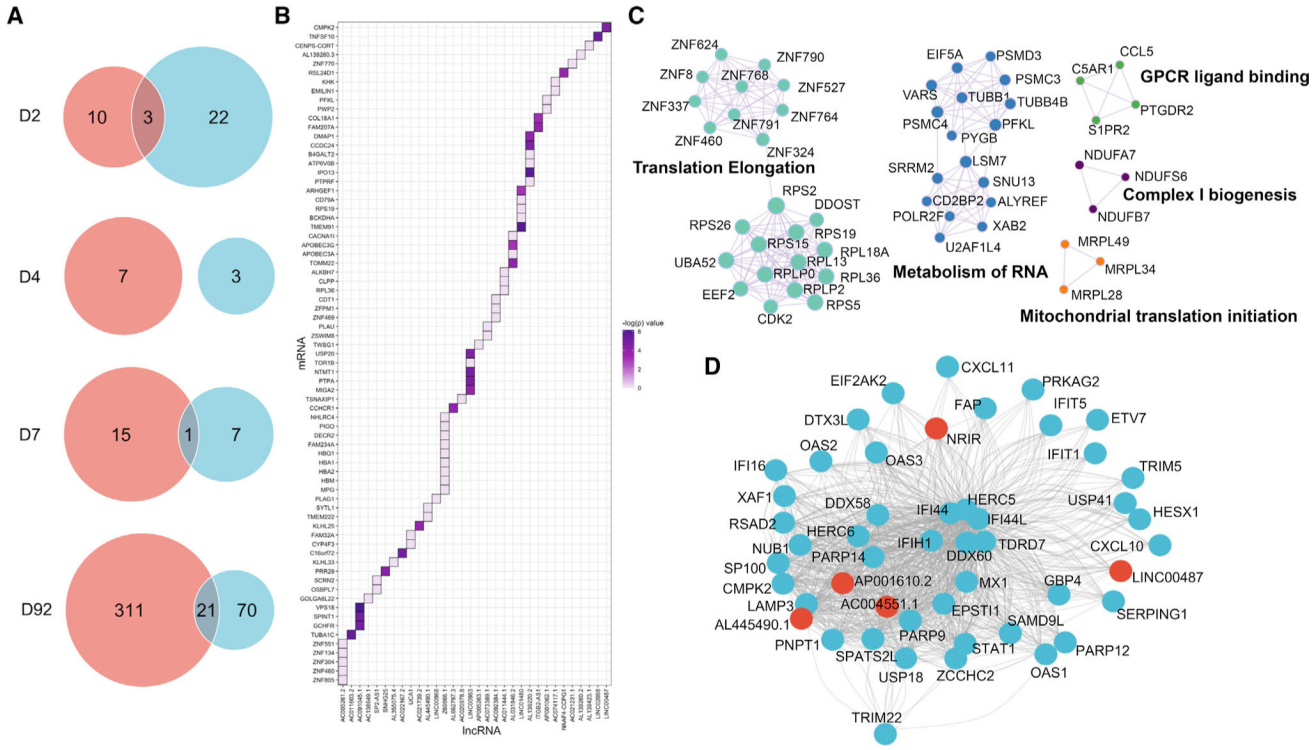


Figure 3. Overlap and enrichment of lncRNAs

(A) Venn diagrams showing the overlaps between the significant differentially expressed lncRNAs.

(B) Correlation matrix showing both significant (Pearson correlation test, $p < 0.05$; purple) and nonsignificant (Pearson correlation test, $p > 0.05$; gray) associations between lncRNAs and their proximal protein-coding genes.

(C) Pathway networks for lncRNA-associated PPI modules showing the subset of proteins that interact in translation elongation, metabolism of RNA, GPCR ligand binding, mitochondrial translation, and complex 1 biogenesis.

(D) WGCNA identifying the network of protein-coding and noncoding genes in the magenta module (Table S4) involved in inflammatory and immune-related pathways based on GO and KEGG enrichment analyses.

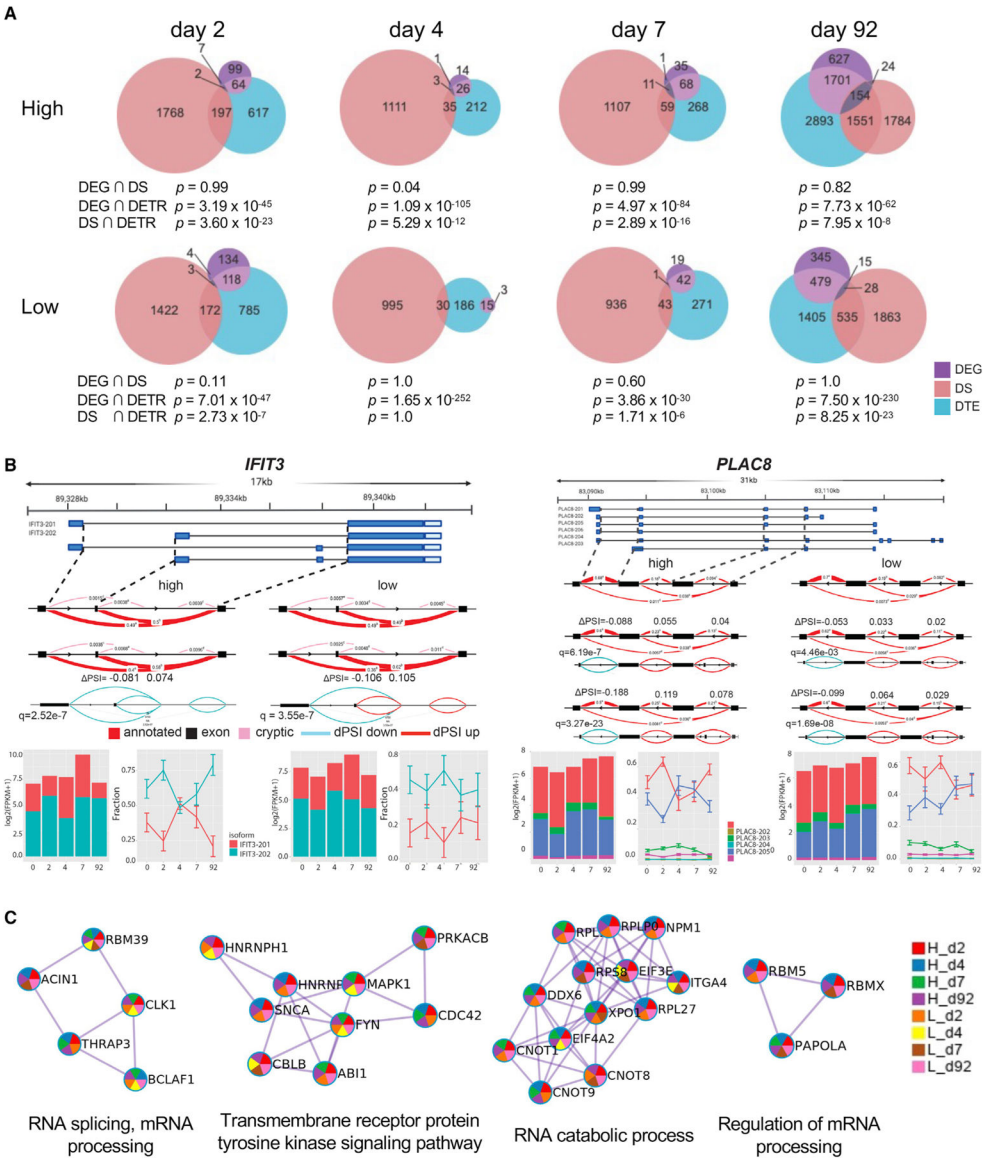


Figure 4. DEGs, splicing, and transcript isoform usage in high and low response groups

(A) Venn diagrams showing the overlaps between the significant genes across the four-point time course for the two groups (p values were from hypergeometric test).

(B) Significant intron clusters in *IFIT3* (chr10: 89,328,078–89,338,661) and *PLAC8* (chr2:86,503,430–86,563,497) illustrating differentially excised introns and exon skipping in the two groups. Splicing events for *IFIT3* and *PLAC8* measured in terms of the change in the percent spliced in (ΔPSI) values indicate an increase (red) or decrease (turquoise) for intron exclusion. FDR-corrected p values are shown for each comparison. Histograms depicting the distribution of $\log_2$FPKM values for isoform-level changes at baseline and across the four-point time course for *IFIT3* and *PLAC8*. Line plots show significant (FDR <0.05) fractional changes in transcript usage for *IFIT3* and *PLAC8*.

(C) Examples of specific PPIs of overlapping DEGs and DTEs exhibiting physical interactions in the network ($-\log_{10}p > 10$). Pie chart represents the proportion of nodes with significant enrichment in the pathway network analysis.

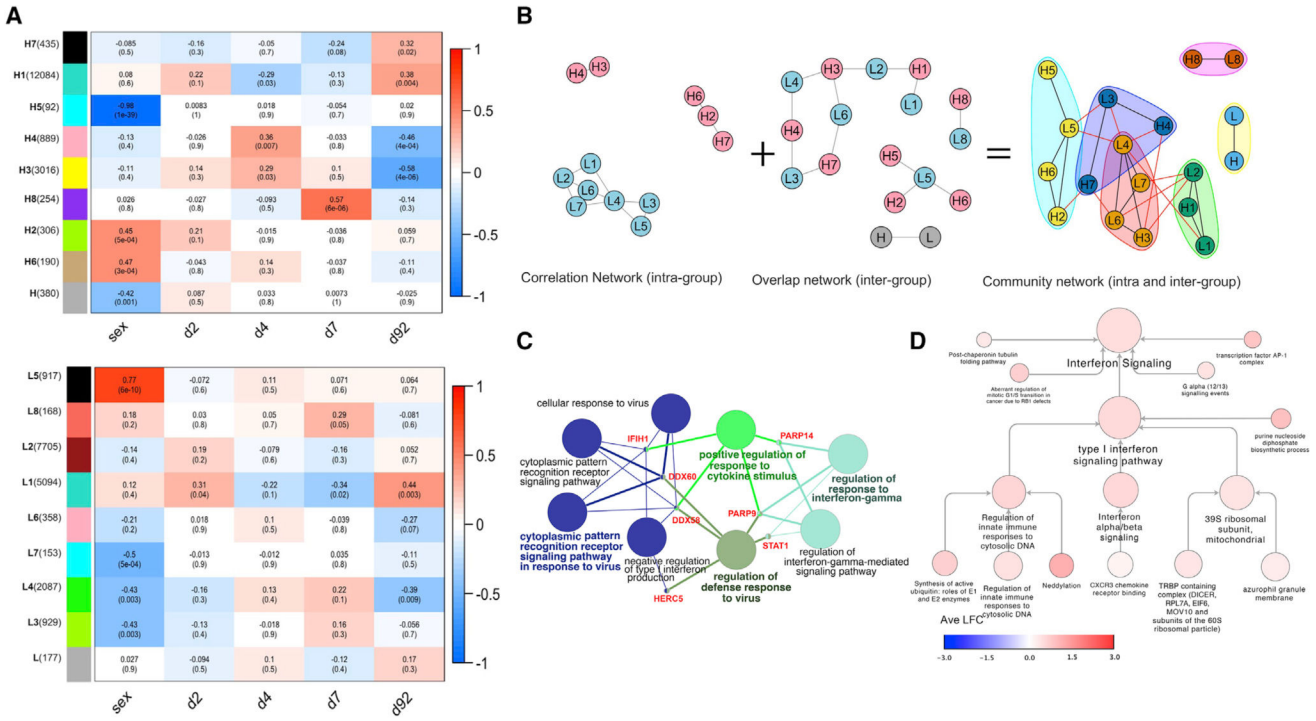


Figure 5. Co-expression networks capture shared and unique biological processes and interactions

(A) Heatmap depicting the adjacency matrix (pairwise correlations) of eigengenes (including the trait weight) from the WGCNA. The gene-level network relationships reveal both positive and negative correlations in a two-group comparison. The color module clustering for nine distinct modules in the high and low response groups (H1-H9 and L1-L9, respectively) is based on the weighted correlations (minimum module size set to 50). Each color-coded module contains a group of highly connected genes for a height cutoff of 0.30, corresponding to a similarity of 0.70 to merge. The colors represent the scaled expression values of the topological overlap matrix for the module-trait relationships based on module-trait positive correlation, negative correlation, or no correlation (red, blue, and white, respectively). The p values are adjusted for multiple comparisons in the analysis.

(B) Co-expression modules capture correlation, overlap, and community networks for the high and low response groups (pink and blue, respectively). Communities of related modules are identified in a merged network of gene correlations and gene overlap with p values corrected for multiple comparisons (FDR <0.01). The network construct demonstrates four core-centric communities with a set of isolated communities (H8 and L8).

(C) Functional enrichment analysis displaying the hub genes for the isolated H8 and L8 communities.

(D) A network map of the 15 identified subnets displaying 655 high confidence protein interactors (score >0.7 STRING). The network comprises interconnected nodes that represent the proteins, and the edges represent the PPIs for the biological process. The average $\log_2$FC of each community network is color coded.

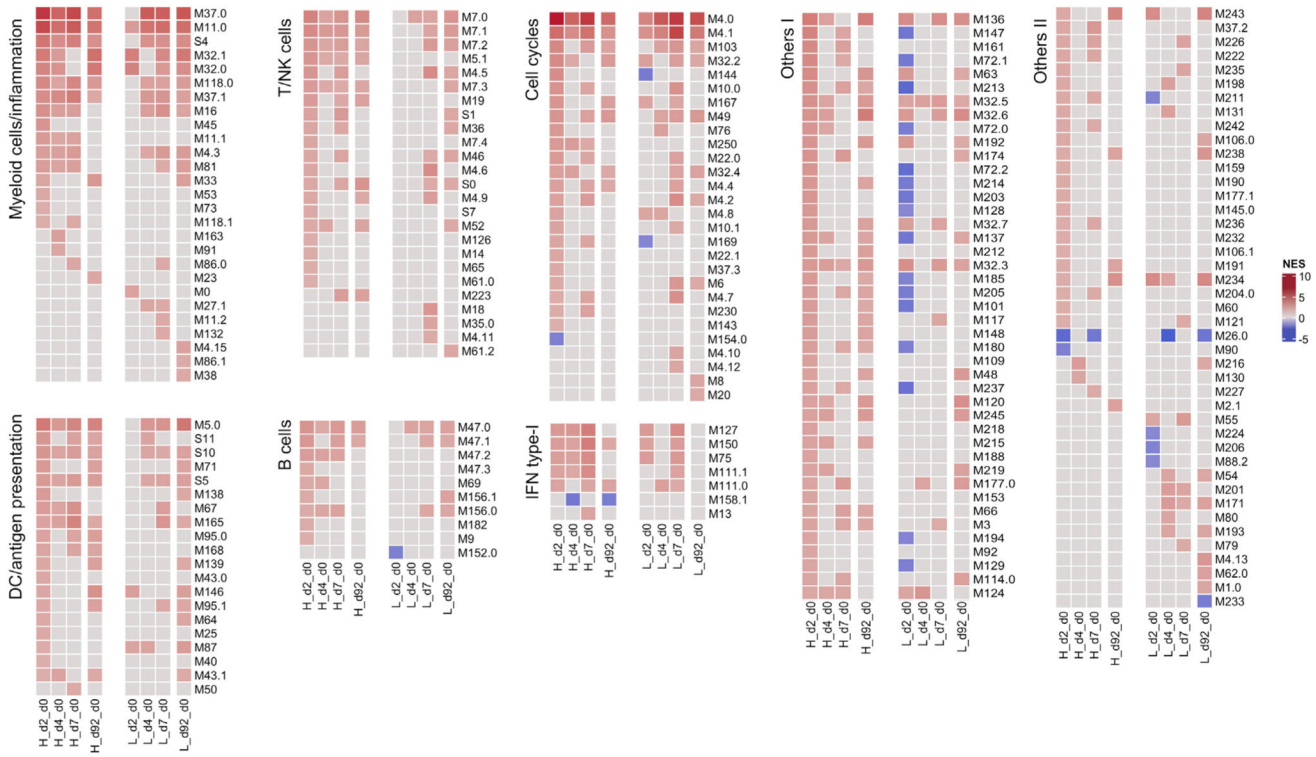


Figure 6. Transcriptomic signatures induced by TAK-003 vaccination

Heatmaps show the enrichment in the gene module-trait relationships. The NES reflects the degree to which a gene set is overrepresented at the top or bottom of the ranked gene list. The modules demonstrate enrichment within gene lists ranked by correlation with $\log_2$FC for distinct biological processes and identify meaningful gene expression patterns and dynamic progression. Days 2, 4, 7, and 92 were compared against day 0. Only modules with gene lists correlated with the functional antibody response were plotted.

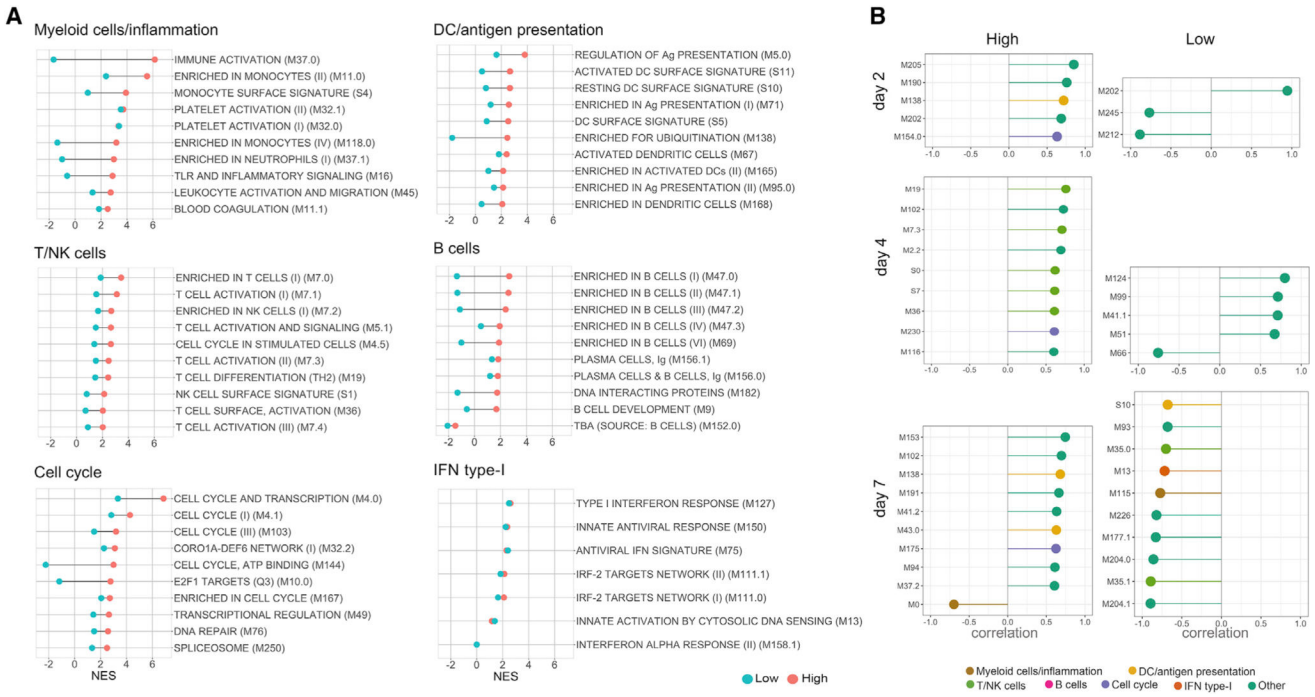


Figure 7. BTM correlations to TAK-003 functional antibody responses

(A) Cleveland dot plots depict the modules that were significantly enriched (FDR <0.05) at day 2 after the first dose. GSEA was used to identify enrichmentof modules within gene lists (Tables S6A and S6B) ranked by correlation with the functional antibody response at day 30. Day 2 was compared against day 0. The response at days 4 and 7 after the first dose and at day 2 after the second dose were broadly similar between the two groups. Days 2, 4, 7, and 92 were compared against day 0.

(B) Lollipop plots comparing Pearson's correlation between neutralizing antibody titers on day 30 and the modules significantly enriched by GSEA at days 2, 4, and 7 after the first dose.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
PAXgene samples from humans: Safety and Immunogenicity Study to Assess TDV, a Live Attenuated Tetravalent Vaccine for Prevention of Dengue Fever	NCT01224639	https://clinicaltrials.gov/ct2/show/NCT01224639
Critical commercial assays		
Paxgene Blood RNA Kit	Qiagen	Cat# 762164
RNeasy MinElute Cleanup kit	Qiagen	Cat# 74204
Qubit RNA High Sensitivity Assay kit	Thermo Fisher	Cat# Q32852
RNA ScreenTapes	Agilent	Cat# 5067-5576
Superscript II	Invitrogen	Cat# 18-064-022
nanoString custom nCounter panel (XT_GX CodeSet 384 rxn)	nanoString Technologies	Item No.116000004
nCounter Master Kit – 192 rxns (NAA-AKIT-192)	nanoString Technologies	Item No.100050
SuperScript™ VILO™ cDNA Synthesis Kit	Thermo Fisher	Cat# 11754050
TaqMan™ Fast Advanced Master Mix	Thermo Fisher	Cat# 4444963
Taqman assay Gene Expression: Assay IDs Hs00356631_g1,Hs01675197_m1,Hs00382744_m1,Hs00155468_m1,Hs01597859_m1,Hs00931718_m1,Hs00748900_s1, and Hs03405707_g1	Thermo Fisher	Cat# 4331182, and 4448490
TrueSeq RNA CD Index Plate (96 Indexes, 96 samples)	Illumina	Cat# 20019792
TruSeq Total Stranded RNA HT kit (w/Ribo-Zero Globin)	Illumina	Cat# 20020612
D1000 ScreenTapes	Agilent	Cat# 5067-5582
Agencourt AMPure XP beads system	Beckman Coulter	Cat# A63881
KAPA Library Quantification Kit Illumina	Kapa Biosystems	Cat# 79960298001
HiSeq 3000/4000 paired-end cluster kit	Illumina	Cat# PE410-1001
HiSeq 3000/4000 SBS kit	Illumina	Cat# FC410-1003
Deposited data		
RNA-seq transcript and gene data	This study	https://www.ncbi.nlm.nih.gov/geo GEO: GSE146658
Software and algorithms		
Network analysis scripts	This study	https://github.com/eykim909/TDV_paper_Scripts
FastQC v 0.11.4	Babraham Bioinformatics	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FASTX-Toolkit	Hannon Lab	http://hannonlab.cshl.edu/fastx_toolkit/
HISAT2 v 2.0.4	Kim et al. (2015)	http://ccb.jhu.edu/software/hisat2/manual.shtml
StringTie v 1.2.2	Pertea et al. (2015)	http://ccb.jhu.edu/software/stringtie/
Samtools v 0.1.19	Li et al. (2009)	http://samtools.sourceforge.net/
DESeq2	Love et al. (2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BioMart	Durinck et al. (2009)	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
Igraph	The igraph core team	https://igraph.org/
WGCNA	Langfelder and Horvath (2008)	https://cran.r-project.org/web/packages/WGCNA/index.html
GSEA v 4.0.0	Broad Institute, Inc	https://www.gsea-msigdb.org/gsea/index.jsp
Leafcutter	Li et al., 2018	https://github.com/davidaknowles/leafcutter
Picard v 2.18.15	Broad Institute, Inc	https://broadinstitute.github.io/picard/
GATK v 3.4.46	Broad Institute, Inc	https://gatk.broadinstitute.org/hc/en-us
Python	prepDE.py	https://www.python.org/
GO biological process	The Gene Ontology Consortium, 2015	http://geneontology.org/
Panther	Thomas et al. (2003)	http://www.pantherdb.org/
Reactome	Reactome 2016	https://reactome.org/
NDEx	NDEx v2.5.0	https://public.ndexbio.org/#/
STRING - Human Protein Links - High Confidence (Score ≥0.7)	Szklarczyk et al. (2017)	https://public.ndexbio.org/#/network/275bd84e-3d18-11e8-a935-0ac135e8bacf
R studio	R project	https://www.r-project.org/(version 3.5.0)
KEGG	Ogata et al. (1999)	https://www.genome.jp/kegg/annotation/
Metascape	Zhou et al., (2019)	http://metascape.org/gp/index.html#/main/step1
Cytoscape	Shannon et al. (2003)	https://cytoscape.org/(version 3.4.0)
OSLOM	Lancichinetti et al. (2011)	http://www.oslom.org/
gProfiler	Reimand et al., (2007)	http://biit.cs.ut.ee/gprofiler/
CDAPS	Singhal et al. (2020)	https://cdaps.readthedocs.io/en/stable/
admixture v1.3.0	Alexander et al. (2009)	https://dalexander.github.io/admixture/publications.html
Plink v1.9	Purcell et al. (2007)	https://www.cog-genomics.org/plink/1.9/