Pre-vaccination transcriptomic profiles of immune responders to the MUC1 peptide vaccine for colon cancer prevention

Cheryl M. Cameron[1†], Vineet Raghu[2,3†], Brian Richardson[1,4†], Leah L. Zagore[4], Banumathi Tamilselvan[1], Jackelyn Golden[4], Michael Cartwright[4], Robert E. Schoen[5], Olivera J. Finn[6*], Panayiotis V. Benos[7,8*], Mark J. Cameron[4*]

[1] Department of Nutrition, Case Western Reserve University, Cleveland, OH
[2] Department of Computer Science, University of Pittsburgh, Pittsburgh, PA
[3] Massachusetts General Hospital, Harvard Medical School, Cambridge, MA
[4] Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH
[5] Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh, Pittsburgh, PA
[6] Department of Immunology, University of Pittsburgh, Pittsburgh, PA
[7] Department of Epidemiology, University of Florida, Gainesville, FL
[8] Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA

† Contributed equally
* Co-corresponding authors

Keywords: colon cancer, colorectal adenoma, cancer vaccine, transcriptomics, MUC1, serological response

Abstract

Self-antigens abnormally expressed on tumors, such as MUC1, have been targeted by therapeutic cancer vaccines. We recently assessed in two clinical trials in a preventative setting whether immunity induced with a MUC1 peptide vaccine could reduce high colon cancer risk in individuals with a history of premalignant colon adenomas. In both trials, there were immune responders and non-responders to the vaccine. Here we used PBMC pre-vaccination and 2 weeks after the first vaccine of responders and non-responders selected from both trials to identify early biomarkers of immune response involved in long-term memory generation and prevention of adenoma recurrence. We performed flow cytometry, phosflow, and differential gene expression analyses on PBMCs collected from MUC1 vaccine responders and non-responders pre-vaccination and two weeks after the first of three vaccine doses. MUC1 vaccine responders had higher frequencies of CD4 cells pre-vaccination, increased expression of CD40L on CD8 and CD4 T-cells, and a greater increase in ICOS expression on CD8 T-cells. Differential gene expression analysis revealed that iCOSL, PI3K AKT MTOR, and B-cell signaling pathways are activated early in response to the MUC1 vaccine. We identified six specific transcripts involved in elevated antigen presentation, B-cell activation, and NF-κB1 activation that were directly linked to finding antibody response at week 12. Finally, a model using these transcripts was able to predict non-responders with accuracy. These findings suggest that individuals who can be predicted to respond to the MUC1 vaccine, and potentially other vaccines, have greater readiness in all immune compartments to present and respond to antigens. Predictive biomarkers of MUC1 vaccine response may lead to more effective vaccines tailored to individuals with high risk for cancer but with varying immune fitness.

Introduction

Self-antigens abnormally expressed in tumors, known as non-viral cancer-associated antigens, have been extensively tested over the last three decades as antigens in therapeutic cancer vaccines (1-3). In preclinical studies, an immune response to these antigens can prevent cancer growth without causing toxicity. In humans, preexisting immunity to some such antigens correlates with better disease outcome or reduced risk of cancer recurrence (4). Nevertheless, therapeutic vaccines utilizing these antigens have had low immunogenicity and no clinical efficacy. This has been attributed to the presence of many immunosuppressive influences in the tumor microenvironment (5, 6). MUC1 is a cancer-associated antigen that has been effective as a vaccine in preclinical animal models but showed limited immunogenicity and efficacy as a therapeutic vaccine in clinical trials in colon, breast, pancreas, prostate and lung cancer (7-11). Hypothesizing that the major difference between the outcome of the vaccine in preclinical models and clinical trials is the high level of immune suppression in cancer patients, we began to develop models and MUC1 vaccines for cancer prevention in patients at risk; before immune suppression develops. As MUC1 is expressed on early premalignant lesions as well as cancer, we chose to study immunogenicity, safety and potential efficacy of this vaccine in the preventative setting in individuals with a history of colonic polyps that increases their risk of colon cancer (12).

From 2008 to 2012, we conducted a single arm trial (NCT-007773097) (13) in 41 individuals. Forty-three percent (43%) of vaccinated participants responded to the vaccine as measured by production of anti-MUC1 IgG at week 12 post vaccination (vaccine responders), and 57% did not respond (vaccine non-responders). From 2015-2020, we

conducted the second study, a randomized, double-blind placebo-controlled multi-center efficacy trial of the same MUC1 vaccine in the setting of newly diagnosed advanced adenomas in 110 individuals (NCT-02134925) (14). Twenty-seven percent (27%) of the vaccinated participants responded to the vaccine. In addition to the immune response, in this trial we evaluated adenoma recurrence by follow-up colonoscopy ≥1 year from the start of vaccination. In vaccine responders, adenoma recurrence was reduced by 38% compared to non-responders and placebo controls. Predictable factors such as gender, age, and HLA-type were not significantly different between vaccine responders and non-responders. It became important to understand why some individuals mounted a potentially protective immune response, while others did not, having the same diagnosis.

In this study, we analyzed PBMC samples collected from both trials at baseline (pre-vaccination) and 2 weeks post-first of 3 vaccines (week 0, week 2 and week 10) from vaccine responders and non-responders and identified comprehensive gene and pathway biomarkers related to vaccine response. We discovered that several key T- and B-cell cellular proliferation and stress pathways were enriched in responders, while oxidative phosphorylation and DNA damage response and repair pathways were enriched in non-responders. Responders had higher frequencies of CD4 cells at baseline, with higher activation and/or costimulatory signaling in CD8 and CD4 T-cells from baseline to week 2 in CD8 T-cells. Phosflow analysis revealed enhanced phosphorylation of B-cell signaling molecules and T-cell help targets in responders at baseline and a significant increase in NFκB phosphorylation in B-cells at week 2. Lastly, we applied graphical modeling approaches (15-17) to this data and built a regression model to

discriminate future responders and non-responders via their predicted and actual IgG response at week 12.

Materials and Methods

PBMC collection

PBMC samples from patients with a history of, or with newly diagnosed, advanced colonic adenoma and at high risk for colon cancer were collected as part of two clinical trials of a MUC1 vaccine registered at clinicaltrials.gov (NCT-007773097, NCT-02134925) (13, 14). The ethics committee/IRB of the following institutions gave ethical approval for this work: Mayo Clinic, Rochester MN; Kansas City Veterans Affairs Medical Center, Kansas City, KS; University of Pittsburgh Medical Center, Pittsburgh PA; University of Puerto Rico, San Juan PR; Thomas Jefferson University Hospital, Philadelphia PA; and Massachusetts General Hospital, Boston MA. All participants provided written informed consent. Blood samples were processed within 24 hours by the same individual, using the same protocol. Heparinized blood was layered on lymphocyte separation medium (MPbio) and centrifuged at 800 g for 10 min with lowest acceleration and deceleration speed. PBMC were collected from the interphase, washed twice, resuspended in 80% human serum and 20% DMSO, and stored in liquid nitrogen.

RNA-Seq

PBMC samples were thawed, pelleted, and lysed in 350 uL of RLT with beta-mercaptoethanol. RNA was isolated using the RNeasy Mini kit (Qiagen). RNA quality was assessed with the Fragment Analyzer (Agilent) and its Standard Sensitivity RNA kit. Total RNA was normalized to 100 ng prior to random hexamer priming and libraries generated

by the TruSeq Stranded Total RNA – Globin kit (Illumina). The resulting libraries were assessed on the Fragment Analyzer (Agilent) with the High Sense Large Fragment kit and quantified using a Qubit 3.0 fluorometer (Life Technologies. Medium depth sequencing (>30 million reads per sample) was performed with a HiSeq 2500 (Illumina) on a High Output, 125 base pair, Paired End run.

Bioinformatic analysis

Raw demultiplexed fastq paired end read files were trimmed of adapters and filtered using the program skewer (18), discarding those with an average phred quality score <30 or a length <36. Trimmed reads were aligned to human reference genome GRCh38 using HISAT2(19) and sorted using SAMtools (20). Aligned reads were counted and assigned to gene meta-features using the program featureCounts (21) as part of the Subread package. Quality control, normalization and analysis were performed in R, using an in-house pipeline utilizing the limma-trend method for differential gene expression testing and the GSVA (22) library for gene set sample enrichment. Final differential gene expression lists were filtered to remove non-coding RNAs as well as LOC features. The datasets for this study can be found in the Gene Expression Omnibus (GEO) public database with the accession number pending.

Flow cytometric analysis

For immune cell phenotyping and assessment of intracellular levels of bcl2, and phosphorylation of STAT3, erk1/2, NF-kB and MTORC targets, cells were first stained with Live/Dead Aqua (Invitrogen) followed by cocktails of monoclonal antibodies

recognizing the following cell surface markers: CD4, CD8, CD45RA, CD27, CCR7, CD152, CD86, CD275, CD11c, CD56, CD16, CD19, CD3, HLA-DR, CD14, CD40, and CD11b. Cells were washed, fixed and permeabilized, then stained with antibodies specific for the following intracellular proteins: NFkB p65, erk1/2 (pT202/pY204), STAT3 (pY705), Akt1, pS6 (S235/236 & S240) and p4E-BP1 (T36/46). Cells were washed and fixed and events were collected on a BD ARIA-SORP instrument. A 15-minute incubation at 37C with recombinant human IL-6 (100ng/mL) (BD Pharmingen) was performed to induce NFκB signaling. After washing, cells were resuspended in staining buffer and sorted on an ARIA-SORP. Data was analyzed using FlowJo software (TreeStar).

Predictive model for post-vaccination immune response

A detailed explanation of computational model development and evaluation can be found in the supplemental methods. Briefly, a LASSO logistic regression (23) was used to develop a prediction model for a binary outcome of response defined by the clinical trial endpoint (≥2-fold increase in IgG from baseline to week 12), using transcriptomic data measured two-weeks post-vaccination (Week 2 data). A Mixed Graphical Models (MGM) algorithm was used to infer a unidirectional graphical model followed by FCI-MAX to determine direction. All statistical analysis was performed in R.

Statistics

Unless otherwise indicated, the Student's t-test was used, with $p \leq 0.05$ chosen as the level of significance.

Results

Vaccine responders and non-responders show differential gene expression in PBMCs pre-vaccination

Next-generation RNA-seq analysis was performed on PBMCs from 46 participants of the two trials, conducted in the same setting of advanced adenoma and with the identical vaccine and vaccination protocol. The vaccine, composed of 100μg of MUC1 peptide plus the polyICLC adjuvant Hiltonol®, was administered at week 0, 2, 10 and 52. We assayed PBMC samples collected at the time of the first injection (baseline) and 2 weeks later, at the time of the second injection in order to be able to define preexisting (at baseline) and early post-vaccination (week 2) signatures of response to MUC1 vaccination. Vaccine responders (R, n=13) in both trials were defined as having anti-MUC1 IgG levels at week 12 (after all three injections) at least two-fold higher than baseline. For some of our analyses we also classified responders by antibody levels into high responders (HR, anti-MUC1 IgG OD450 at 1:80 plasma dilution ≥0.4), low responders (LR, OD450 at 1:80 plasma dilution <0.4), and non-responders (NR) (no difference from baseline).

RNA-seq performed on PBMCs collected immediately pre-vaccination (baseline) revealed a total of 2,321 genes that were differentially expressed between all responders and all non-responders at baseline (Fig 1A). Within these genes, 1,337 showed increased transcript levels and 984 showed decreased levels. Among the top 50 differentially expressed genes by p-value (Fig 1B), 47 genes were upregulated in responders compared to non-responders. The upregulated genes were involved in transcriptional and epigenetic regulation, including multiple subunits of the SWI/SNF chromatin remodeling complex (ARID1A, ARID1B, and SMARCC2), as well as NCOA6, a multifunctional

transcriptional coactivator and component of the Set1-like H3K4-methyltransferase complex ASCOM, all of which have been shown to play a role in the pathogenesis of cancer (24). Additional cancer-relevant transcriptional regulators higher in responders before vaccination include SP2, NFYC, AKNA, MYPOP, and ZNF652. CUX1 is a subunit of the NF-muNR repressor that binds to the matrix attachment regions of the immunoglobulin heavy chain enhancer and the TCR enhancer. The epigenetic regulator HCFC1 tethers Set and Sin3 histone modifying complexes together and was also higher in responders at baseline (25). We observed an increase in TRAPPC9, an activator of NFκB, and FAM168A, which is involved in the PI3K/AKT/NFκB signaling pathway. RRAGA was one of 3 downregulated genes among the top differentially expressed genes in responders and plays a role in regulating the mTORC1 complex (26). The presence of transcriptional and epigenetic regulators within the list of upregulated genes could explain the large number of significant changes in steady-state RNA levels we observed and suggests a differing global transcriptional program between responders and non-responders before vaccination. As these upstream regulators have the potential to broadly remodel the transcriptome, they may represent potent therapeutic targets.

Vaccine responders and non-responders show differences in gene expression in PBMCs at week 2 post-vaccination

At two weeks post-first injection, we found 1,887 genes differentially expressed in responders vs. non-responders, 934 genes upregulated and 953 genes downregulated (Fig 1C). The top 50 genes arranged by p-value are shown in Figure 1D. Among the upregulated genes are several cancer-related transcriptional regulators including PPARD

(27) and HMGA1 (28), key regulators of lipid pathways (29), transcription elongation factor SPT6 (SUPT6H), SOAT2, an enzyme involved in lipoprotein and cholesterol regulation, and GPD1, an enzyme that plays a key role in lipid metabolism, are also upregulated in responders. Among the downregulated cancer-related genes we found five involved in mitosis and G2/M DNA replication checkpoint, the kinesin-like proteins KIF11 and KIF15 (30), centromere/centrosome proteins CENPF and CEP55 (31), and the cell cycle regulator protein DLGAP5 (32). Importantly, CEP55 and DLGAP5 are key predictors of antibody response in our graphical model discussed below.

We then performed double contrast analysis of the genes that were significantly differentially changed from baseline to week 2 in the responders vs. non-responders (Fig 1E). The top 50 genes by p-value (Fig 1F) are enriched in immune-related genes. IFNL1 is upregulated in contrast to CD38 and IL12RB2, which are downregulated in responders. Again, selective upregulation of transcriptional and epigenetic regulators in responders is evident; examples include PRDM5, ZNF230, ZCCHC9, ZKSCAN4, and the epigenetic regulator ALKBH3, which demethylates DNA and RNA in cancer cells (33).

ICOS/ICOSL signaling is differentially associated with response to the MUC1 vaccine

We performed gene set variation analysis (GSVA) of the baseline gene expression data to identify biological pathways regulating the response to vaccination. Vaccine responders displayed significant upregulation of genes involved in the ICOS-ICOSL pathway in T-helper Cells signaling pathway, with increased expression of multiple genes both at baseline and week 2 post-vaccination (Fig 2A and 2B, respectively). At baseline,

responders expressed higher levels of ICOSL, IL2RB, and CD4 coreceptor genes, while at week 2 post-vaccination higher levels of the downstream NFKB pathway genes including NFKB2, RELB and RELA were evident.

By flow cytometry, we determined that higher frequencies of CD4 T-cells were present in responders at baseline, potentially related to enhanced ICOSL signaling (Fig 2C). We also detected differences in expression levels of key proteins in this pathway. Higher levels of expression of CD40L were evident in CD4 and CD8 cells (Fig 2D and 2E respectively). Greater increases in ICOS expression were detected post-vaccination inCD8 T-cells of responders. (Fig 2F).

mTOR signaling is upregulated in responders to MUC1 vaccination

As many components of the ICOS/ICOSL pathway were significantly higher in the responders vs. non-responders, we focused on differences in the mTOR signaling pathway which lies directly downstream of ICOS/ICOSL engagement. Top enriched pathways in high responders vs. non-responder comparisons included PI3K/AKT/MTOR signaling, WNT/beta-catenin signaling and hedgehog signaling (Fig 3A). In contrast, the Myc targets V1 pathway and DNA repair were negatively associated in high responders.

To validate enhanced mTOR signaling in responders, we developed a phosflow-based panel of antibodies (see Materials and Methods) to measure the phosphorylation levels of RPS6, a commonly used readout of mTORC1 activity. There was a greater increase in RPS6 phosphorylation in CD4 and CD8 T-cells, B-cells (CD19), and monocytes (CD14) of responders, an observation validating our finding at the phosphoprotein level (Fig 3B).

We also performed intracellular staining targeting the phosphorylated AKT1 kinase upstream of the MTORC1 signaling complex. Similarly, we found a greater increase in AKT1 phosphorylation in responders compared to non-responders.

B-cell signaling and enhanced antigen presentation signatures are positively associated with response to MUC1 vaccination

As ICOS/ICOSL-mediated signaling promotes fitness of the T lymphocyte compartment, we hypothesized that signaling from the T-cells to the B-cell and APC compartment was also differentially induced. Pathway enrichment analysis of the transcriptomic data revealed significant enrichment of B-Cell Receptor Signaling and PI3K Signaling in B Lymphocytes pathways at baseline (Fig 4A and 4B) and NFκB and CD40 Signaling pathways at baseline and at week 2 post-vaccination (Fig 4C and 4D). CD40 receptor engagement on the surface of antigen presenting cells, such as B-cells, leads to activation of NFκB signaling and enhanced cellular survival and function. Notably, multiple signaling component genes (MAP kinases and Jak3) are significantly upregulated in responders at baseline, followed by upregulation of additional signaling molecules at week two (TRAF1, TRAF3, NFKB1, NFKB2, RELA and RELB*).*

We validated increased expression of CD40 and HLA-DR on B-cells (CD19+) in responders (Fig 4E) using flow cytometric analyses. We used an intracellular phosflow panel to detect phosphorylation of the p65 subunit of NFκB. We found increased IL6-induced NFκB signaling via phosphoryation of p65 in T-cells, HLA-DR+ non-B/non-DC APCs and B-cells of responders (Fig 4F), and B-cells of high responders expressing significantly higher levels of HLADR compared to non-responders (Fig 4E).

Finally, we performed gene set variation analysis using the Nakaya et al. vaccine immunogenicity pathway database (34) and determined that a plasmacytoid dendritic cell (DC) signature was already enriched in the PBMCs of responders at baseline (Fig 5A) and further enriched at week 2 post-vaccination (Fig 5B). Vaccine responders showed increased HLA-DR levels on DCs (CD3-, CD19-, HLA-DR+, CD11c+) at baseline (Fig 5C). Responders also showed a greater relative change in CD86 and CD40 expression from baseline to two weeks (Fig 5D-E). Altogether, these results indicate that additional signatures of enhanced antigen presentation are associated with enhance response to MUC1 vaccination.

Six differentially expressed transcripts 2 weeks post-vaccination predict week 12 IgG response to the MUC1 vaccine

Based on evidence of key differences in cell populations and molecular pathways between responders and non-responders at baseline and post-vaccination, we hypothesized that some differentially expressed genes may be useful for patient selection and outcome prediction. We tested this hypothesis by applying LASSO regression and MGM-FCI-MAX (17), a graphical modeling algorithm, to 7,968 transcripts meeting a minimal variance threshold. We first performed a cross-validation experiment (see Supplemental Methods) to assess the ability to predict antibody response to the vaccine at week 12, using the transcriptomic signatures at week 2 post-vaccination. Our model achieved an area under the receiver operating characteristic curve (AUROC) value of 0.741 to predict response vs. non-response (Fig 6A). At a predicted probability threshold of 0.5, the model achieved a sensitivity of 91.7% (22 predicted responders / 24 true

297 responders) and a specificity of 36.8% (7 predicted non-responders / 19 true non-
298 responders). Predicted response odds were correlated with the magnitude of antibody
299 titer at week 12 (R^2 = 0.209, p<0.001) (Fig 6B), and with the ratio of IgG titer at week 12
300 versus baseline (R^2 = 0.147, p=0.015).

301 Next, we used graphical models to determine the variables directly linked to week 12
302 antibody titer and distinguish them from simple correlates. We produced a full model with
303 all genes selected in the previous cross-validation experiments, which were learned using
304 the entire week 2 dataset (Fig 6C). Finally, we identified 6 genes that are directly linked
305 to week 12 antibody titer: RP11.81H14.2, CEP55, and TNFSF14 (negatively associated)
306 and C22orf29, DDX12P, and HLA-DQA2 (positively associated) (Fig 6D). The role of
307 these transcripts and how they may contribute to the induction of immune response and
308 vaccine efficacy (discussed below) warrants further investigation.

309 **Discussion**

310 Therapeutic cancer vaccines, tested in numerous clinical trials over several decades,
311 failed to realize the promise generated by the discovery of tumor antigens capable of
312 eliciting humoral and cellular immunity. In most cases, vaccines administered after
313 primary tumor removal failed to boost anti-tumor immunity and prevent tumor recurrence.
314 Ultimately, a greater understanding of the many immunosuppressive forces in the tumor
315 microenvironment helped to explain the reduced efficacy of therapeutic vaccines. These
316 discoveries support preventative vaccines as an alternative approach to cancer
317 vaccination to reduce cancer risk and incidence as they could be applied in the absence
318 of cancer and cancer-induced immunosuppression. The two clinical trials from which we

319 derived the PBMCs, applied this preventative approach by vaccinating individuals without
320 cancer but at high risk for colon cancer due to advanced colonic adenoma diagnosis (13,
321 14). We expected a vaccine response from most individuals, measured by the production
322 of anti-MUC1 IgG antibodies. It came as a surprise, therefore, that the majority did not
323 respond. However, those who responded had high levels of anti-MUC1 antibodies and
324 established a long-lasting memory response, showing that the vaccine was capable of
325 inducing immunity and response was determined by the individuals receiving the vaccine.
326 Mechanisms underlying this variable response to an apparently efficacious vaccine were
327 not clear.

328 To address this major knowledge gap with an unbiassed approach, we performed RNA-
329 seq on total PBMCs from participants in the two MUC1 peptide vaccine trials to identify
330 genes and pathways differentially regulated at baseline as well as post-vaccination in the
331 participants that responded versus those that failed to respond to the vaccine. The
332 analysis revealed that vaccine responders at baseline exhibited an enrichment of key
333 pathways governing survival and proliferation in immune cells, such as mTOR and NFKB
334 signaling, as well as increased frequencies of CD4 and CD8 T-cells. There were more
335 memory CD8 T-cells at baseline and week 2 (post MUC1 vaccine) in responders (Fig 2A
336 and 2E). Responders also had more CD4 T-cells at baseline and higher frequencies of
337 memory CD4 T-cells at week 2 (Fig 2A, 2C, 2D). Other immune compartments appeared
338 to differ in favor of responders with higher levels of BCL2 expression at baseline in CD14+
339 and greater increases in BCL2 expression at week 2 post-vaccination (data not shown).
340 The control of DC longevity by the regulation of BCL2 directly impacts immune responses.

Higher levels of BCL2 suggest enhanced survival in the myeloid compartment and consequently better antigen presentation.

The differentially expressed genes and pathways that we have identified in vaccine responders and non-responders pre-vaccination and two weeks post-vaccination are top candidates for early biomarkers of vaccine immunogenicity at week 12. Among the six genes directly linked to MUC1 antibody production at week 12, some have already been identified as diagnostic biomarkers (CEP55, TNFSF14) while the others merit deeper investigation as they hold the potential to enhance our understanding of vaccine response. Overexpression of CEP55 has been observed in numerous cancer cell types, including premalignant lesions of the colon (35), and is a known correlate of poor prognosis (36). Notably, a CEP55 peptide vaccine was proposed for breast and colorectal carcinoma immunotherapy as CEP55 is involved in the PI3K/Akt signaling pathway. TNFSF14, also known as LIGHT, functions as a co-stimulatory factor for the activation of lymphoid cells and modulates T-cell proliferation (37, 38). HLA-DQA2 codes for the alpha chain of the HLA-DQ complex and is primarily involved in antigen presentation (37). Interestingly, HLA-DQ phenotypes have been linked with non-responsiveness to hepatitis B vaccination (39). DDX12P is an m^6A-associated prognostic pseudogene, correlated with favorable outcomes in patients with head and neck squamous cell carcinoma (40). Furthermore, expression patterns of DDX12P were correlated with anti-tumor response and may regulate immune-involved genes through miRNA targeting. RP11-81H14.2 (LINC02384) is a long intergenic non-coding RNA primarily expressed in TH1 cells (41). Little is known about the function of LINC02384; however, it has been proposed to act as a competitive endogenous RNA of IL2RA and IL7R by reducing available shared

regulatory miRNAs (42). C22orf29 (also known as RTL10) may have the capacity to induce apoptosis in a BH3 domain-dependent manner, presumably by engaging the Bcl2 family regulatory network to modulate the intrinsic apoptotic signaling pathway (43). The identification of known diagnostic biomarkers and immunotherapy targets within our predictive genes lends credence to the graphical models utilized in this study.

Given the cancer immunoprevention potential of the MUC1 peptide vaccine response, characterized by a reduction of adenoma recurrence (14), the differentially expressed genes and regulated pathways we identified hold promise as therapeutic targets for vaccine non-responders. While these observations were made on responders and non-responders to the MUC1 vaccine, it is likely that a number of these differentially enriched genes and pathways play a role in other vaccine responses. Many vaccines do not elicit a response in all recipients, such as the yearly flu vaccine which varies in effectiveness between 40% and 60% (44). The selected adjuvant for the MUC1 vaccine, polyICLC, excels at activating dendritic cells to promote type I (innate) immunity (45). Alternative adjuvants may need to be considered for non-responders to the MUC1 peptide adjuvanted with polyICLC. While efforts are often made to improve the vaccine, it also may be important to consider an individual's incoming immune history to respond to the vaccine. Indeed, numerous research studies have demonstrated a correlation between the immune status prior to vaccination and the subsequent antibody response (46, 47). Overall, individuals that responded to the MUC1 vaccine showed a greater readiness in all the immune compartments to present and respond to antigen. The ability to profile individuals as potential responders or non-responders can aid in the selection of those who benefit most from a particular vaccine. At the same time, understanding the barriers

387 to response in non-responders can inform the development of better vaccine designs
388 suitable for specific immune genotypes and phenotypes.

Authors' Disclosures

R.E. Schoen reports support from Freenome, Exact Sciences, and Immunovia during the conduct of the study; in addition, R.E. Schoen has a patent for Anti-MUC1 Binding Agent and Uses Thereof pending. O.J. Finn reports personal fees from PDS Biotech, Invectys, Immodulon, and Ardigen outside the submitted work.

Funding

This research was supported by NCI and NHLBI funding to OJF (R35CA210039), PVB (R01HL159805), and MJC (P30CA043703 Sub-Project 9164). VR was supported by a fellowship through the T32CA082084 grant.

Acknowledgements

We appreciate the tremendous efforts of the team members that conducted the two clinical trials and provided samples for this study, particularly Lisa Boardman, Marcia Cruz-Correa, Ajay Bansal, David Kastenberg, Chin Hur, Lynda Dzubinski, Sharon Kaufman, Luz M Rodriguez, Ellen Richmond, Asad Umar, Eva Szabo, Andres Salazar, John McKolanis, Pamela Beatty, Reetesh Pai, Aatur Singhi, Camille Jacqueline, Riyue Bao, Brenda Diergaarde, Ryan McMurray, Carrie Strand, Nathan Foster, David Zahrieh, and Paul Limburg. We are also indebted to all the trial participants for their commitment to our cancer prevention mission. We thank the Genomics Core at the Lerner Research Institute of Cleveland Clinic and the Genomics and Applied Functional Genomics Cores at Case Western Reserve University for their technical and analytical support.

Figure Legends

Figure 1: Differential gene expression pre- and post-MUC1 vaccination in responders and non-responders. Differentially expressed genes (DEGs) are shown for each time point and contrast in the left column, top 50 DEGs are shown in the right column. Group status is indicated in the row above the heatmap with responders (R) in dark blue, and non-responders (NR) in light blue. Z-scored normalized gene expression for each gene is displayed horizontally across all samples (diverging color scale legend on the upper right of each heatmap). Log2 fold-change and p values are indicated in the purple and green vertical columns, respectively. Unsupervised clustering of the samples is indicated by the black dendrogram at the top of heatmap, while clustering of the genes is indicated at the far left. Heatmaps showing all DEGs pre-vaccination (Baseline/Week 0) (A), top 50 DEGs pre-vaccination (Baseline/Week 0) (B), DEGs at Week 2 post-vaccination (C), top 50 DEGs at Week 2 post-vaccination (D), all genes demonstrating longitudinal changes at Week 2 vs. Week 0 (Delta Wk2-Wk0) (E), top 50 genes demonstrating longitudinal changes at Week 2 vs. Week 0 (Delta Wk2-Wk0) (F) in PBMCs from responders versus non-responders.

Figure 2: Differentially expressed T-cell fitness signatures in PBMCs from responders and non-responders to MUC1 vaccination is associated with CD4 frequencies and expression of multiple regulators of T-cell help. (A,B) iCOSL signaling pathway-related genes are associated with response to MUC1 vaccination at Baseline (A) and Week 2 post-vaccination (B) in MUC1 vaccine responders and non-responders. Heatmaps are organized as in Figure 1. (C) Violin plots of CD4+ T-cell

436 frequencies as determined by flow cytometry. (D-F) Violin plots of CD40L expression on
437 CD4 T-cells (D), CD8 T-cells (E), and the change in ICOS levels on CD8 T-cells (Delta
438 Week 2 vs. Week 0/Baseline) (F) measured by geometric mean fluorescence intensity
439 (MFI).
440
441 **Figure 3: The mTOR signaling pathway is upregulated in enhanced response to**
442 **MUC1 vaccination. (A)** Heatmap showing top differentially enriched pathways from the
443 Hallmark Gene Set (MSigDB) in the PBMCs from high responders (HR) vs. non-
444 responders (NR) at Baseline. Group status is indicated in the row above the heatmap as
445 follows: high responders (HR) - pink, while non-responders (NR) - light blue. Z-scored
446 normalized pathway enrichment log 2 fold-change and p values are displayed as in Figure
447 1. Unsupervised clustering of the samples is indicated at the top of the heatmap, while
448 clustering of the pathways is displayed on the far left. **(B)** Violin plots showing the level of
449 S6 ribosomal protein phosphorylation in the indicated cell subsets (CD4, CD8, CD19/B-
450 cells and CD14/monocytes). **(C)** Violin plots showing the level of AKT1 phosphorylation
451 in the indicated cell subsets (CD4, CD8, CD19/B-cells and CD14/monocytes). For all
452 violin plots, geometric mean fluorescence intensity (MFI) is shown on the y-axis.
453 Responders (R) and non-responders (NR) are designated by dark blue and light blue
454 respectively.
455
456 **Figure 4: B-cell signaling and NFκB signaling signatures are associated with**
457 **response to MUC1 vaccination. (A,B)** Heatmaps showing enrichment of B-Cell
458 Receptor Signaling pathways in the PBMCs from Responders (R) vs. non-responders

(NR) at Baseline (A) and high responders (HR) vs. non-responders (NR) at Baseline (B). **(C,D)** Differential gene expression from the CD40 Signaling pathway from PBMCs from R vs. NR at Baseline (C) and Week 2 (D). Heat maps are organized as in Figure 1. **(E)** Violin plots showing expression levels of CD40 and HLA-DR on B-cells (CD19+). **(F)** Violin plots of NFκB complex p65 subunit phosphorylation in T-cells (CD3+), non-DC/non-B antigen presenting cells (CD11C-HLADR+), and B-cells (CD19+). For violin plots, geometric mean fluorescence intensity (MFI) is shown on the y-axis.

Figure 5: Signatures of enhanced antigen presentation are evident in participants with an enhanced response to MUC1 vaccination. Heatmaps showing enrichment of dendritic cell (DC) specific genes in the PBMCs from Responders (R) vs. non-responders (NR) at Baseline (A) and Week 2 post-vaccination (B). Heatmaps are organized as in Figure 1. Violin plots of HLA-DR expression on DCs at Baseline (Week 0) (C), as well relative change in expression of CD86 (D) and CD40 (E) in these cells. For all violin plots, geometric mean fluorescence intensity (MFI) is shown on the y-axis.

Figure 6: Graphical models of response from transcriptomic data measured two-weeks post-vaccination. (A) Receiver Operating Characteristic curve of response (≥2-fold increase in IgG) using week 2 transcriptome signature. (B) Correlation of predicted response odds with the magnitude of antibody titer at week 12 (C) Full model showing all neighbors and second neighbors of week 12 antibody titer levels (Week 12 IgG), (D) reduced model showing only direct causes of week 12 antibody titer. Color of edge denotes a positive vs negative correlation, and size denotes edge stability.

Supplemental Material

Table 1: Detailed information on antibody panels used for flow cytometric analysis.

Supplemental Methods:

Causal Modeling with MGM-FCI-MAX

A Probabilistic Graphical Model (PGM) represents the joint distribution of variables in a dataset as a graph where each node corresponds to a variable and an edge between two nodes, A and B, corresponds to a conditional dependence between A and B given the rest of the variables in the data (48). PGM's come in two types: directed graphical models rely on additional assumptions to infer cause and effect direction between variables, while undirected graphical models indicate only conditional dependence.

MGM-FCI-MAX is a new method to learn a directed model (17). The algorithm begins by inferring an undirected graphical model using the Mixed Graphical Models (MGM) algorithm (16) and then uses FCI-MAX to determine causal direction. MGM models categorical variables as multinomial and continuous variables as Gaussian with a mean given by a linear regression on all other variables. The full joint distribution of the model is given in Equation 1 below. Here, x_s is the sth of p continuous y_j is the jth of q categorical variables. β_{st} is the linear interaction term between two continuous variables, and α_s is the continuous node potential. ρ_{sj} is the edge potential function between continuous and categorical variables, and it takes on one value for each category of the variable y_j. Finally, ϕ_{rj} is the potential function between two categorical variables with a unique value for all combinations of categories of the variables y_r and y_j.

$$p(x,y;\theta) \propto exp\left(\sum_{s=1}^{p}\sum_{t=1}^{p} -\frac{1}{2}\beta_{st}x_s x_t + \sum_{s=1}^{p}\alpha_s x_s + \sum_{s=1}^{p}\sum_{j=1}^{q}\rho_{sj}(y_j)x_s + \sum_{j=1}^{q}\sum_{r=1}^{q}\phi_{rj}(y_r,y_j)\right) \quad (1)$$

The pseudolikelihood approach was used to optimize the model parameters (49). The pseudolikelihood is the product of the conditional distributions of each variable, and

it is a consistent estimator of the goodness of fit of the model to the data. To ensure a sparse graph, edges are penalized via the method proposed in (16) with separate penalty parameters for each edge type: (CC = Continuous-Continuous, CD = Continuous-Discrete, DD= Discrete-Discrete) (Equation 2). Here, $\tilde{l}(\Theta)$ is the negative-log pseudolikelihood and the rest are penalty terms which ensure a sparse model.

$$argmin_{\Theta}\ \tilde{l}(\Theta) + \lambda_{CC}\sum_{\varphi<\omega}\left|\beta_{\varphi\omega}\right|_1 + \lambda_{CD}\sum_{\omega,\tau}\|\rho_{\omega\tau}\|_2 + \lambda_{DD}\sum_{\delta<\omega}\|\gamma_{\delta\omega}\|_F \quad (2)$$

FCI-MAX is used to determine causal directions using the undirected graph as a starting point. FCI-MAX is an extension of the Fast-Causal Inference (FCI) algorithm (50), which is a sound and complete constraint-based algorithm for learning the causal structure of a set of variables in the presence of confounding variables. FCI uses conditional independence tests to rule out unlikely cause and effect relationships. FCI-MAX improves the accuracy of FCI by performing additional tests to more accurately assign orientations, especially in datasets with small sample sizes. The output of the algorithm is a graphical causal model where there are four possible edges. An edge of the form ("A --> B") suggests that A is a cause of B and B is not a cause of A. An edge ("A <--> B") suggests that neither A nor B is a cause of the other, that is, a confounding variable causes both. An edge ("A o--> B") suggests that if there are no latent variables causing both A and B, then A is a cause of B. Finally, an edge of the form ("A o-o B") suggests that both endpoints are inconclusive. In high dimensional datasets (small sample size, large number of variables) these algorithms are less accurate in inferring causal orientations as they are in inferring the presence or absence of an edge(17).

A likelihood ratio independence test (15) suitable for mixed data was used by FCI-MAX. All three sparsity parameters for MGM ($\lambda_{CC}, \lambda_{CD}, \lambda_{DD}$) were set to the default 0.2 and

531 α = 0.1 was used for the independence test threshold for FCI-MAX. MGM-FCI-MAX was
532 run on 100 bootstrap samples of the data, and edges were included in the final model if
533 they appeared in at least 10% of bootstrapped samples.
534
535 **Computational Model Development and Evaluation**
536 LASSO logistic regression (23) was used to develop a prediction model (select
537 genes and infer logistic regression coefficients) for a binary outcome of response defined
538 by the clinical trial endpoint (≥2-fold increase in IgG from baseline to week 12), using
539 transcriptomic data measured two-weeks post-vaccination (Week 2 data). To develop and
540 simultaneously evaluate model predictions, a nested leave-one-out cross validation
541 approach was used. Iteratively, each individual sample is used as an evaluation set with
542 the remaining samples used to learn model parameters. On each training set, a LASSO
543 logistic regression was performed with an internal leave-one-out cross validation to
544 choose the optimal sparsity penalty value (λ). The predictions on the single left-out
545 sample in each round of cross validation were then used for downstream analysis.
546 The Receiver Operator Characteristic (ROC) curve was calculated, and predictive
547 accuracy of the model was measured using the area under the curve (AUC) of response
548 vs. non-response, as well as sensitivity and specificity of predicted probabilities. Feature
549 stability was measured to ensure that models remained similar across different cross-
550 validation iterations.
551 MGM-FCI-MAX was used to infer the variables directly linked to response, using
552 clinical data (age, sex, and BMI) and those genes selected by LASSO in the week 2
553 transcriptomic data in at least one of the ten folds. LASSO logistic regression was used

554 to build a predictive model of response in each cross-validation fold. All statistical analysis

555 was performed in R.

556

REFERENCES

1. Zitvogel L, Perreault C, Finn OJ, Kroemer G. Beneficial autoimmunity improves cancer prognosis. Nat Rev Clin Oncol. 2021;18(9):591-602.
2. Jacqueline C, Finn OJ. Antibodies specific for disease-associated antigens (DAA) expressed in non-malignant diseases reveal potential new tumor-associated antigens (TAA) for immunotherapy or immunoprevention. Semin Immunol. 2020;47:101394.
3. Finn OJ, Rammensee HG. Is It Possible to Develop Cancer Vaccines to Neoantigens, What Are the Major Challenges, and How Can These Be Overcome? Neoantigens: Nothing New in Spite of the Name. Cold Spring Harb Perspect Biol. 2018;10(11).
4. Gnjatic S, Bronte V, Brunet LR, Butler MO, Disis ML, Galon J, et al. Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy. J Immunother Cancer. 2017;5:44.
5. Liu J, Fu M, Wang M, Wan D, Wei Y, Wei X. Cancer vaccines as promising immuno-therapeutics: platforms and current progress. J Hematol Oncol. 2022;15(1):28.
6. Gordon B, Gadi VK. The Role of the Tumor Microenvironment in Developing Successful Therapeutic and Secondary Prophylactic Breast Cancer Vaccines. Vaccines (Basel). 2020;8(3).
7. Kimura T, McKolanis JR, Dzubinski LA, Islam K, Potter DM, Salazar AM, et al. MUC1 vaccine for individuals with advanced adenoma of the colon: a cancer immunoprevention feasibility study. Cancer Prev Res (Phila). 2013;6(1):18-26.
8. Ramanathan RK, Lee KM, McKolanis J, Hitbold E, Schraut W, Moser AJ, et al. Phase I study of a MUC1 vaccine composed of different doses of MUC1 peptide with SB-AS2 adjuvant in resected and locally advanced pancreatic cancer. Cancer Immunol Immunother. 2005;54(3):254-64.
9. Pantuck AJ, van Ophoven A, Gitlitz BJ, Tso CL, Acres B, Squiban P, et al. Phase I trial of antigen-specific gene therapy using a recombinant vaccinia virus encoding MUC-1 and IL-2 in MUC-1-positive patients with advanced prostate cancer. J Immunother. 2004;27(3):240-53.
10. Bartsch R, Singer CF, Pfeiler G, Hubalek M, Stoeger H, Pichler A, et al. Conventional versus reverse sequence of neoadjuvant epirubicin/cyclophosphamide and docetaxel: sequencing results from ABCSG-34. Br J Cancer. 2021;124(11):1795-802.
11. Schimanski CC, Kasper S, Hegewisch-Becker S, Schroder J, Overkamp F, Kullmann F, et al. Adjuvant MUC vaccination with tecemotide after resection of colorectal liver metastases: a randomized, double-blind, placebo-controlled, multicenter AIO phase II trial (LICC). Oncoimmunology. 2020;9(1):1806680.
12. Ma P, Beatty PL, McKolanis J, Brand R, Schoen RE, Finn OJ. Circulating Myeloid Derived Suppressor Cells (MDSC) That Accumulate in Premalignancy Share Phenotypic and Functional Characteristics With MDSC in Cancer. Front Immunol. 2019;10:1401.
13. Lohmueller JJ, Sato S, Popova L, Chu IM, Tucker MA, Barberena R, et al. Antibodies elicited by the first non-viral prophylactic cancer vaccine show tumor-specificity and immunotherapeutic potential. Sci Rep. 2016;6:31740.

14. Schoen RE, Boardman LA, Cruz-Correa M, Bansal A, Kastenberg D, Hur C, et al. Randomized, Double-Blind, Placebo-Controlled Trial of MUC1 Peptide Vaccine for Prevention of Recurrent Colorectal Adenoma. Clin Cancer Res. 2023;29(9):1678-88.
15. Sedgewick AJ, Buschur K, Shi I, Ramsey JD, Raghu VK, Manatakis DV, et al. Mixed graphical models for integrative causal analysis with application to chronic lung disease diagnosis and prognosis. Bioinformatics. 2019;35(7):1204-12.
16. Sedgewick AJ, Shi I, Donovan RM, Benos PV. Learning mixed graphical models with separate sparsity parameters and stability-based model selection. BMC Bioinformatics. 2016;17 Suppl 5:175.
17. Raghu VK, Ramsey JD, Morris A, Manatakis DV, Sprites P, Chrysanthis PK, et al. Comparison of strategies for scalable causal discovery of latent variable models from mixed data. Int J Data Sci Anal. 2018;6(1):33-45.
18. Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics. 2014;15:182.
19. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60.
20. Yang J, Ding X, Sun X, Tsang SY, Xue H. SAMSVM: A tool for misalignment filtration of SAM-format sequences with support vector machine. J Bioinform Comput Biol. 2015;13(6):1550025.
21. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.
22. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
23. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996;58(1):267-88.
24. Wang XX, Fu L, Li X, Wu X, Zhu Z, Fu L, Dong JT. Somatic mutations of the mixed-lineage leukemia 3 (MLL3) gene in primary breast cancers. Pathol Oncol Res. 2011;17(2):429-33.
25. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev. 2003;17(7):896-911.
26. Shen K, Valenstein ML, Gu X, Sabatini DM. Arg-78 of Nprl2 catalyzes GATOR1-stimulated GTP hydrolysis by the Rag GTPases. J Biol Chem. 2019;294(8):2970-5.
27. Pudakalakatti S, Titus M, Enriquez JS, Ramachandran S, Zacharias NM, Shureiqi I, et al. Identifying the Metabolic Signatures of PPARD-Overexpressing Gastric Tumors. Int J Mol Sci. 2022;23(3).
28. Padua D, Pinto DF, Figueira P, Pereira CF, Almeida R, Mesquita P. HMGA1 Has Predictive Value in Response to Chemotherapy in Gastric Cancer. Curr Oncol. 2021;29(1):56-67.
29. Lin Y, Lin L, Fu F, Wang C, Hu A, Xie J, et al. Quantitative proteomics reveals stage-specific protein regulation of triple negative breast cancer. Breast Cancer Res Treat. 2021;185(1):39-52.
30. Terribas E, Fernandez M, Mazuelas H, Fernandez-Rodriguez J, Biayna J, Blanco I, et al. KIF11 and KIF15 mitotic kinesins are potential therapeutic vulnerabilities for malignant peripheral nerve sheath tumors. Neurooncol Adv. 2020;2(Suppl 1):i62-i74.

31. Han J, Xie R, Yang Y, Chen D, Liu L, Wu J, Li S. CENPA is one of the potential key genes associated with the proliferation and prognosis of ovarian cancer based on integrated bioinformatics analysis and regulated by MYBL2. Transl Cancer Res. 2021;10(9):4076-86.
32. Feng Y, Li F, Yan J, Guo X, Wang F, Shi H, et al. Pan-cancer analysis and experiments with cell lines reveal that the slightly elevated expression of DLGAP5 is involved in clear cell renal cell carcinoma progression. Life Sci. 2021;287:120056.
33. Ueda Y, Ooshio I, Fusamae Y, Kitae K, Kawaguchi M, Jingushi K, et al. AlkB homolog 3-mediated tRNA demethylation promotes protein synthesis in cancer cells. Sci Rep. 2017;7:42271.
34. Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, et al. Systems biology of vaccination for seasonal influenza in humans. Nat Immunol. 2011;12(8):786-95.
35. Sakai M, Shimokawa T, Kobayashi T, Matsushima S, Yamada Y, Nakamura Y, Furukawa Y. Elevated expression of C10orf3 (chromosome 10 open reading frame 3) is involved in the growth of human colon tumor. Oncogene. 2006;25(3):480-6.
36. Jeffery J, Sinha D, Srihari S, Kalimutho M, Khanna KK. Beyond cytokinesis: the emerging roles of CEP55 in tumorigenesis. Oncogene. 2016;35(6):683-90.
37. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics. 2016;54:1 30 1-1 3.
38. Capece D, Verzella D, Fischietti M, Zazzeroni F, Alesse E. Targeting costimulatory molecules to improve antitumor immunity. J Biomed Biotechnol. 2012;2012:926321.
39. Singh R, Kaul R, Kaul A, Khan K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J Gastroenterol. 2007;13(12):1770-87.
40. Tan L, Qin Y, Xie R, Xia T, Duan X, Peng L, et al. N6-methyladenosine-associated prognostic pseudogenes contribute to predicting immunotherapy benefits and therapeutic agents in head and neck squamous cell carcinoma. Theranostics. 2022;12(17):7267-88.
41. Spurlock CF, 3rd, Tossberg JT, Guo Y, Collier SP, Crooke PS, 3rd, Aune TM. Expression and functions of long noncoding RNAs during human T helper cell differentiation. Nat Commun. 2015;6:6932.
42. Zhang C, Dang D, Cong L, Sun H, Cong X. Pivotal factors associated with the immunosuppressive tumor microenvironment and melanoma metastasis. Cancer Med. 2021;10(14):4710-20.
43. Zhang X, Weng C, Li Y, Wang X, Jiang C, Li X, et al. Human Bop is a novel BH3-only member of the Bcl-2 protein family. Protein Cell. 2012;3(10):790-801.
44. Smith DJ, Forrest S, Ackley DH, Perelson AS. Variable efficacy of repeated annual influenza vaccination. Proc Natl Acad Sci U S A. 1999;96(24):14001-6.
45. Caskey M, Lefebvre F, Filali-Mouhim A, Cameron MJ, Goulet JP, Haddad EK, et al. Synthetic double-stranded RNA induces innate immune responses similar to a live viral vaccine in humans. J Exp Med. 2011;208(12):2357-66.
46. Team H-CSP, Consortium H-I. Multicohort analysis reveals baseline transcriptional predictors of influenza vaccination responses. Sci Immunol. 2017;2(14).

694 47. Fourati S, Tomalin LE, Mule MP, Chawla DG, Gerritsen B, Rychkov D, et al. Pan-
695 vaccine analysis reveals innate immune endotypes predictive of antibody responses to
696 vaccination. Nat Immunol. 2022;23(12):1777-87.
697 48. Koller D, Friedman, N. Probabilistic graphical models: principles and techniques:
698 MIT press; 2009.
699 49. Lee JD, Hastie TJ. Learning the Structure of Mixed Graphical Models. J Comput
700 Graph Stat. 2015;24(1):230-53.
701 50. Spirtes P, Meek C, Richardson T, editors. Causal inference in the presence of
702 latent variables and selection bias. Proceedings of the Eleventh conference on
703 Uncertainty in artificial intelligence; 1995: Morgan Kaufmann Publishers Inc.
704

Figure 1 All DEGs Top 50 DEGs

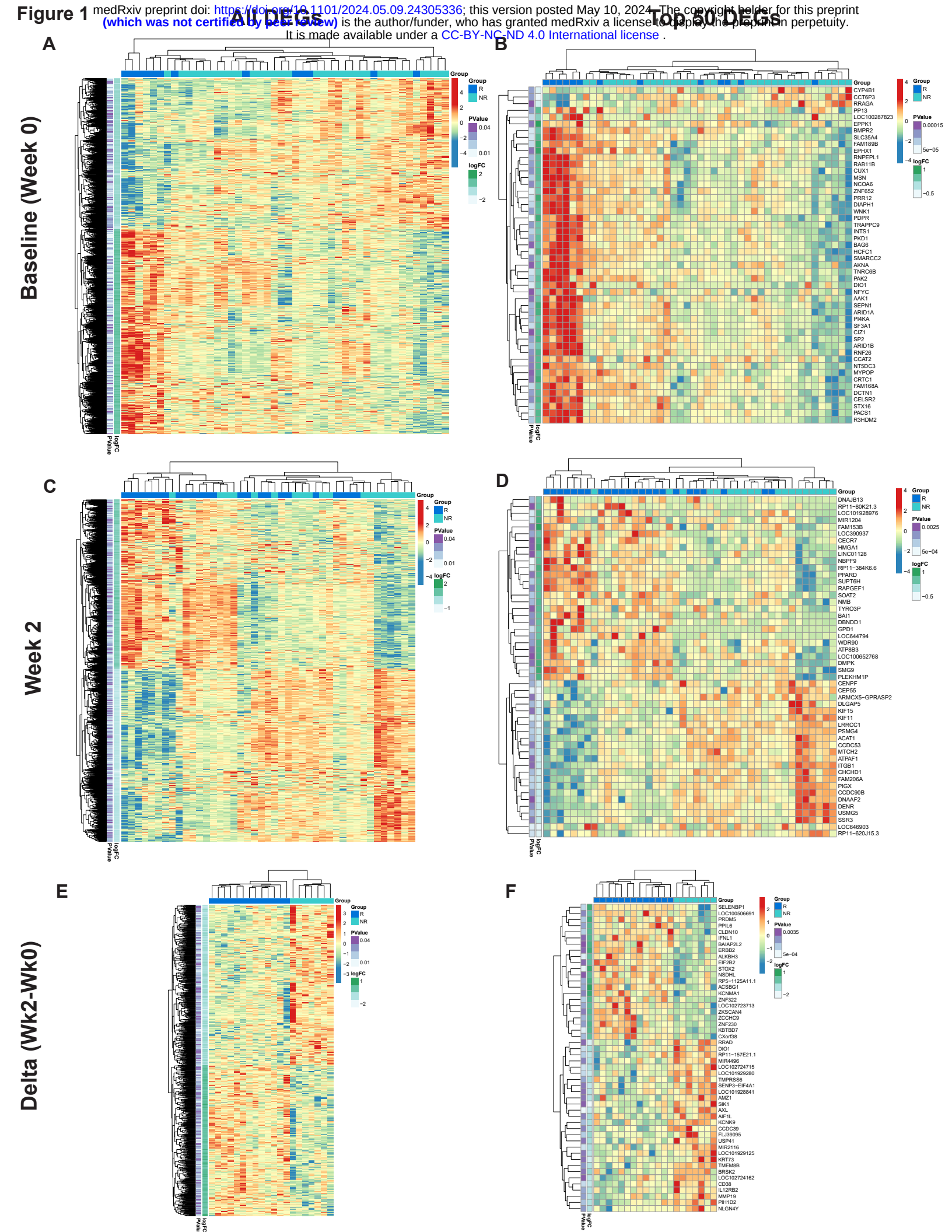
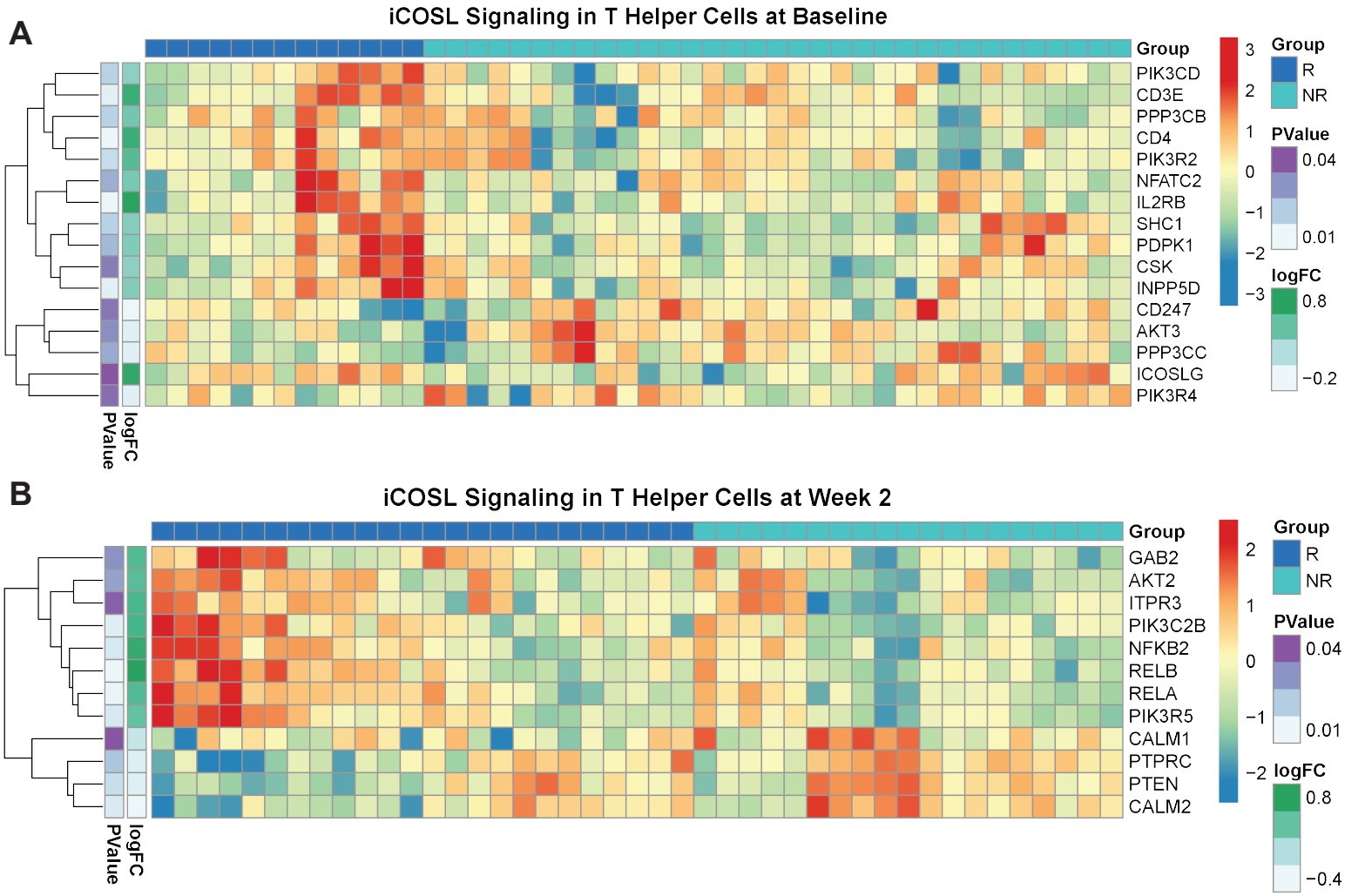
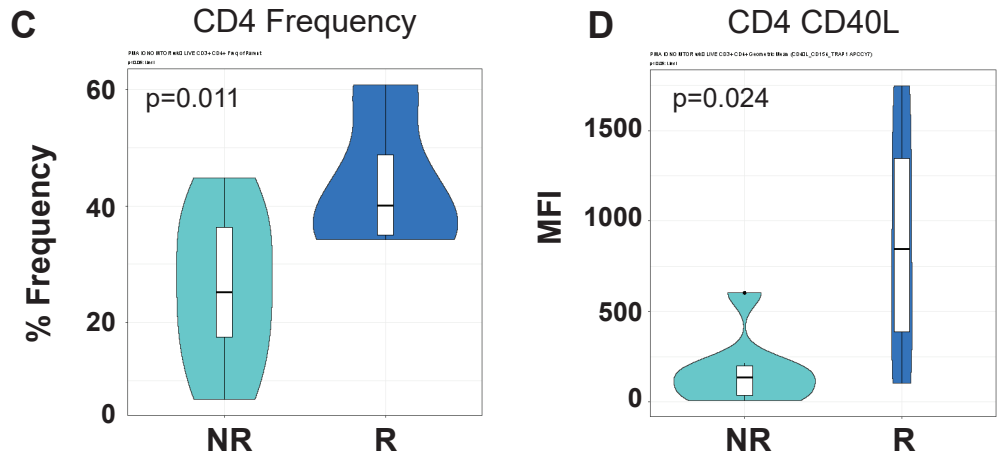


Figure 2





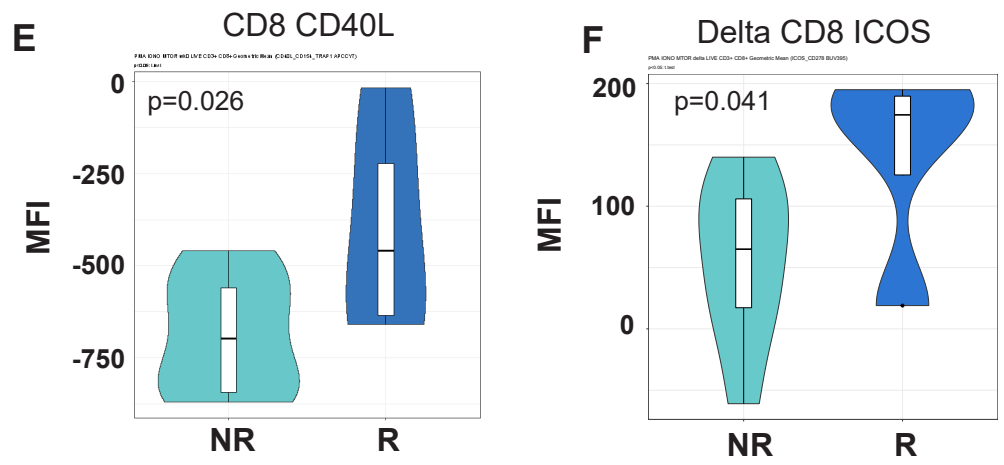


Figure 3

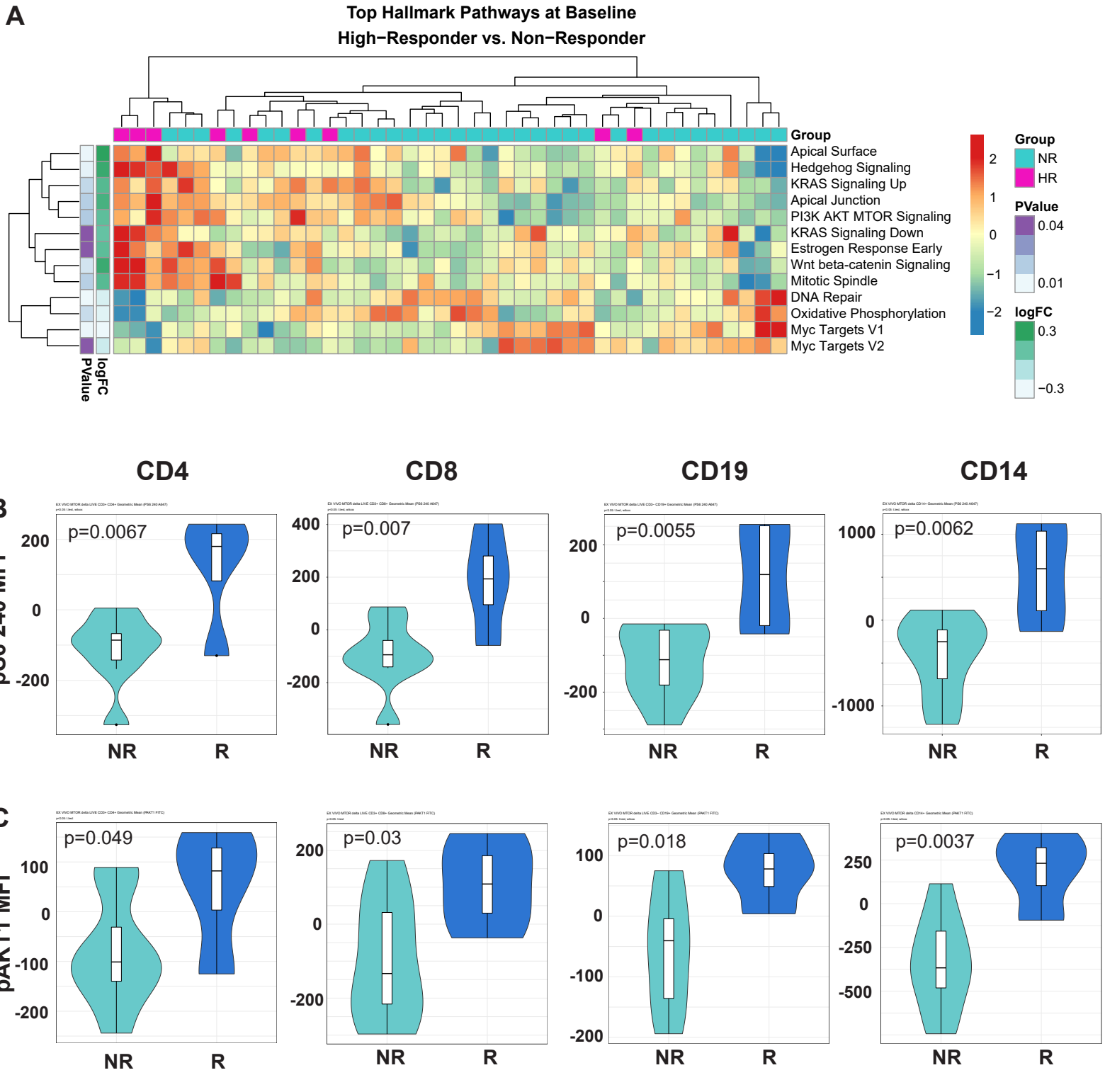


Figure 4

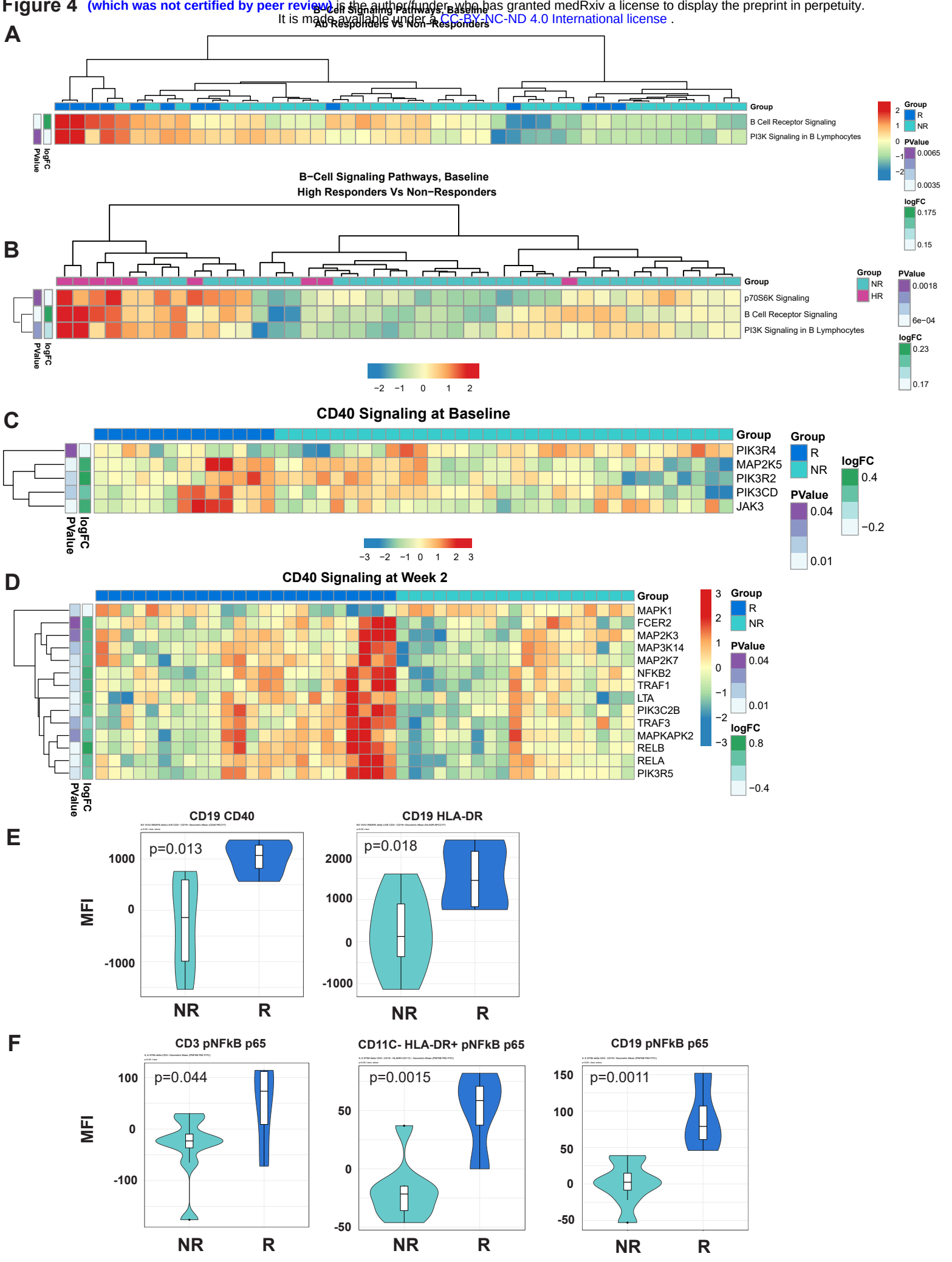


Figure 5

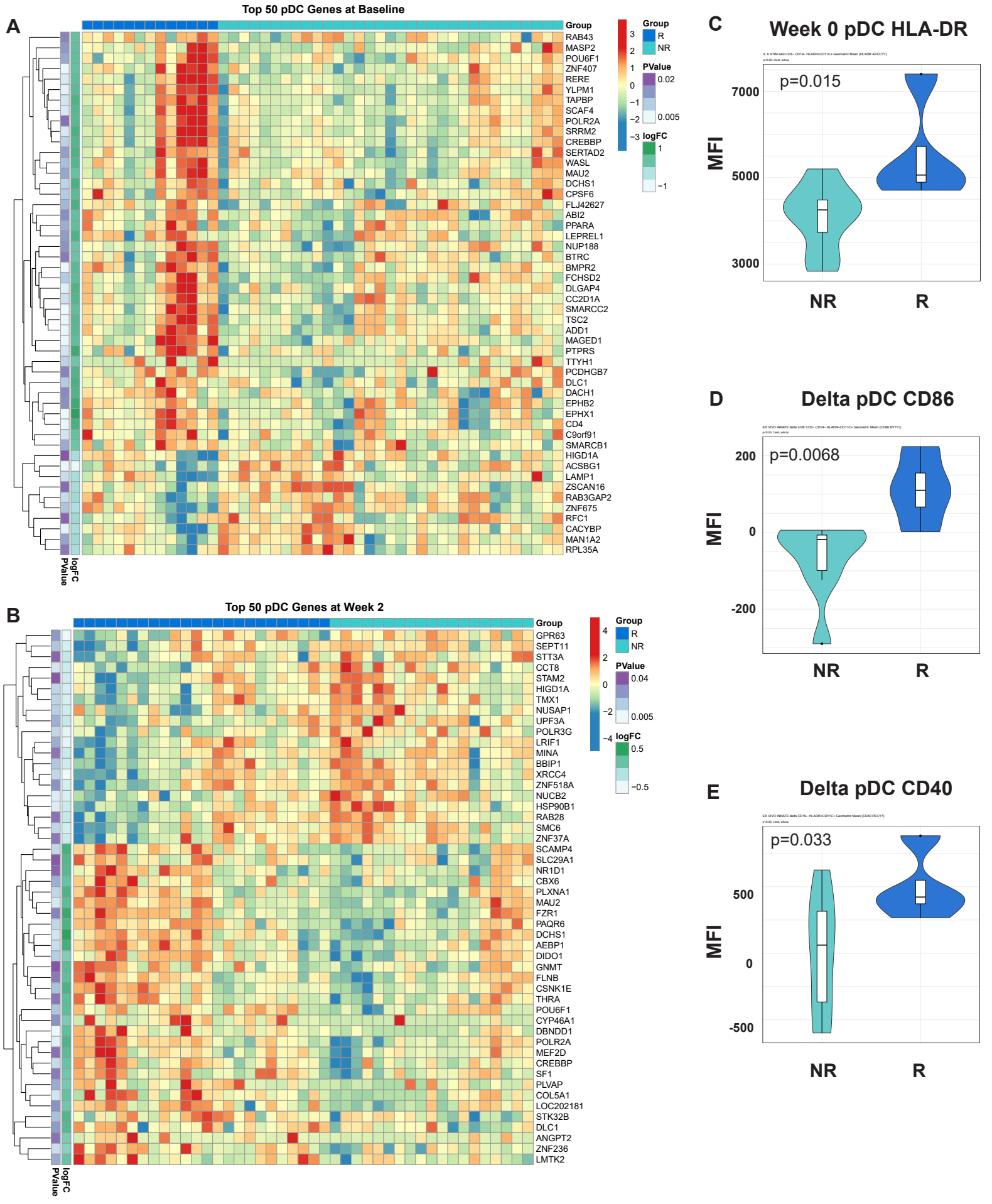
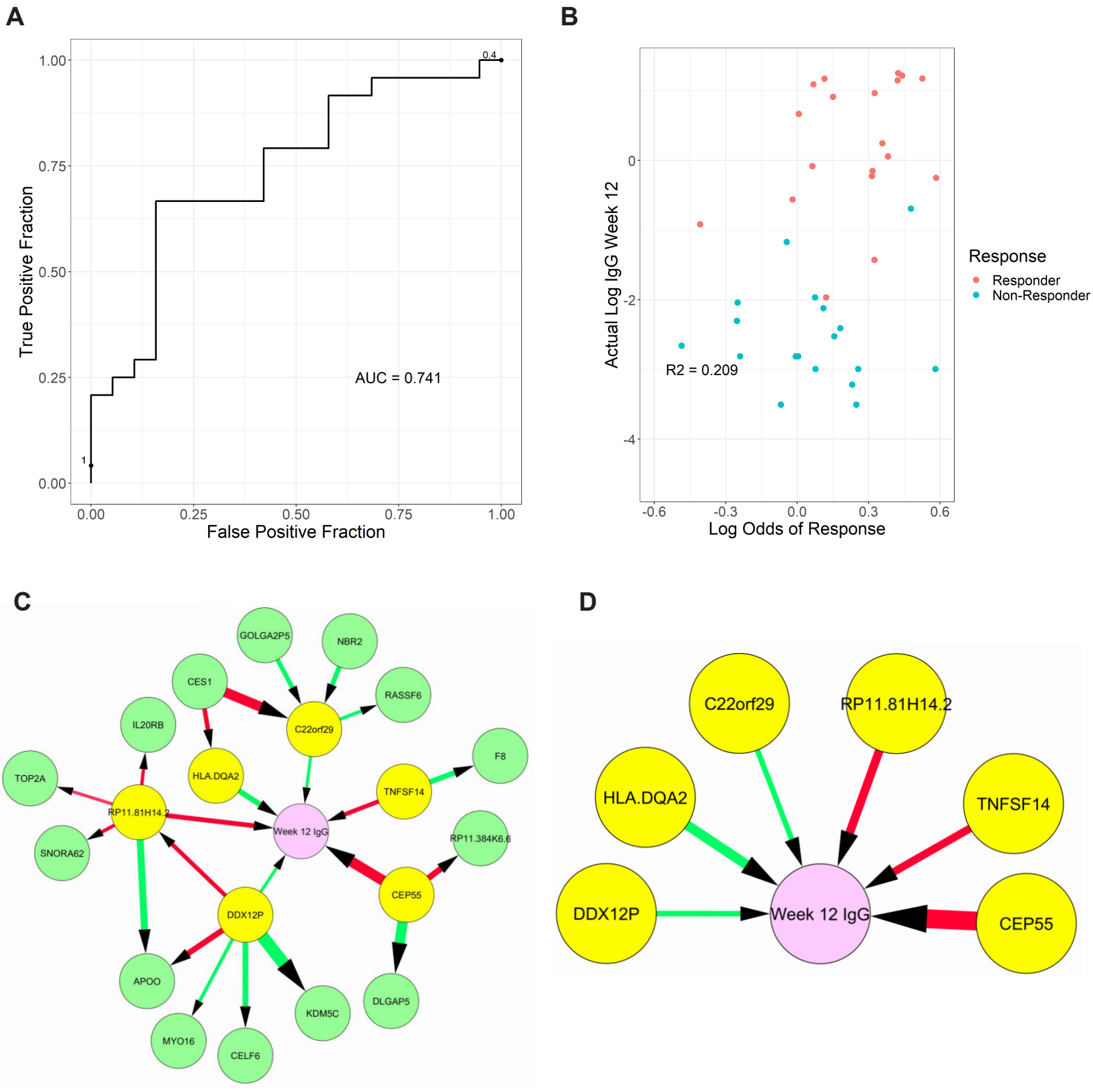


Figure 6



Supplemental Table 1

Panel 1: Innate Phenotyping

Company	Cat #	Antibody	Color
BD	560114	STAT3 (pY705)	PerCP-Cy5.5
BD	558421	anti-NF-κB p65 (pS529)	A488
BD	563931	CD152/CTLA4	BV786
Biolegend	305440	CD86	BV711
BD	743008	CD275/ICOSLG	BV650
BD	563929	CD11c	BV605
Invitrogen	L34957	Live/Dead	Amcyan
Biolegend	318340	CD56	BV510
BD	558122	CD16	PB
BD	564303	CD19	BUV737
BD	563546	CD3	BUV395
BD	335796	HLA-DR	APC Cy7
BD	557923	CD14	A700
BD	612593	Anti-ERK1/2 (pT202/pY204)	APC
Biolegend	334321	CD40	PE-Cy7
BD	563601	BCL-2	PE-CF594
BD	555388	CD11b	PE

Panel 2: MTOR Signaling

Company	Cat #	Antibody (Clone)	Color
BD	338426	CD16	PerCP-Cy5.5
BD	560048	anti-Akt1 (PKBa/Akt)	FITC
BD	564058	CD56	BV786
BD	563677	CD8	BV711
Biolegend	304135	CD45RA	BV650
Invitrogen	Q10008	CD4	Q.605
Invitrogen	L34957	Live/Dead	Amcyan
Biolegend	302242	CD19	BV510
BD	561457	pS6 (S235/S236)	V450
BD	564444	CD14	BUV737
BD	565885	Anti-ICOS (CD278)	BUV395
BD	563588	Anti-Human CD154 (CD40L)	APCeFluor780
BD	557943	CD3	A700
BD	560432	Anti-S6 pS240	A647
BD	562321	CD19	PE-CF594
BD	560684	Anti-Human CD28	PE-Cy7
BD	560285	p4E-BP1 (T36/46)	PE