Microarray Gene Expression Dataset Re-analysis Reveals Variability in Influenza Infection and Vaccination

Lavida R. K. Rogers[1,2], Gustavo de los Campos[2,3,4] and George I. Mias[2,5]*

[1] Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI, United States, [2] Institute for Quantitative Health Science and Engineering, Michigan State University, East Lansing, MI, United States, [3] Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI, United States, [4] Department of Statistics and Probability, Michigan State University, East Lansing, MI, United States, [5] Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, United States

Influenza, a communicable disease, affects thousands of people worldwide. Young children, elderly, immunocompromised individuals and pregnant women are at higher risk for being infected by the influenza virus. Our study aims to highlight differentially expressed genes in influenza disease compared to influenza vaccination, including variability due to age and sex. To accomplish our goals, we conducted a meta-analysis using publicly available microarray expression data. Our inclusion criteria included subjects with influenza, subjects who received the influenza vaccine and healthy controls. We curated 18 microarray datasets for a total of 3,481 samples (1,277 controls, 297 influenza infection, 1,907 influenza vaccination). We pre-processed the raw microarray expression data in R using packages available to pre-process Affymetrix and Illumina microarray platforms. We used a Box-Cox power transformation of the data prior to our down-stream analysis to identify differentially expressed genes. Statistical analyses were based on linear mixed effects model with all study factors and successive likelihood ratio tests (LRT) to identify differentially-expressed genes. We filtered LRT results by disease (Bonferroni adjusted $p < 0.05$) and used a two-tailed 10% quantile cutoff to identify biologically significant genes. Furthermore, we assessed age and sex effects on the disease genes by filtering for genes with a statistically significant (Bonferroni adjusted $p < 0.05$) interaction between disease and age, and disease and sex. We identified 4,889 statistically significant genes when we filtered the LRT results by disease factor, and gene enrichment analysis (gene ontology and pathways) included innate immune response, viral process, defense response to virus, Hematopoietic cell lineage and NF-kappa B signaling pathway. Our quantile filtered gene lists comprised of 978 genes each associated with influenza infection and vaccination. We also identified 907 and 48 genes with statistically significant (Bonferroni adjusted $p < 0.05$) disease-age and disease-sex interactions, respectively. Our meta-analysis approach highlights key gene signatures and their associated pathways for both influenza infection and vaccination. We also were able to identify genes with an age and sex effect. This gives potential for improving current vaccines and exploring genes that are expressed equally across ages when considering universal vaccinations for influenza.

Keywords: influenza, vaccinations, immunity, aging, meta-analysis, micro-arrays, gene-expression

1. INTRODUCTION

The influenza virus, a respiratory pathogen, is responsible for seasonal influenza (also known as the flu), influenza pandemics and high rates of morbidity and mortality worldwide (1). The influenza virus infects the upper respiratory tract by invading the epithelial cells, releasing viral RNA, replicating and spreading throughout the respiratory tract while also causing inflammation (2). Influenza is a highly contagious disease and spreads easily via contact with an infected person's nasal discharges and cough droplets (3). The main virulence factors are haemagglutinin (HA) and neuraminidase (NA) (2). These surface glycoproteins are also important for determining the sub-type of the influenza virus. The influenza virus can also reduce host gene expression through their viral proteins (4, 5). The viral proteins affect transcription and translation in the host which reduces the production of host proteins and promotes immune system evasion for the virus (4, 5). The virus interferes with host gene expression to promote viral gene expression, and this affects the immune system of the host by reducing the expression of immune components such as the major histocompatibility (MHC) molecules antigen presentation, and interferon and cytokine signaling pathways (4, 6).

Influenza is a global health burden, and as a preventative method vaccinations are offered annually. Vaccines are modified annually because the influenza virus strains change and mutate every season (7). The influenza vaccinations target the viral strains and sub-types that researchers predict would be most prevalent each flu season (3, 8). Furthermore, there are groups in the population who are considered at a higher risk for influenza infection, and they include young children, elderly, individuals who are immunocompromised, and females who are pregnant (3). The Centers for Disease Control and Prevention (CDC) has estimated, for the 2017–2018 season for influenza, 959,000 hospitalizations and over 79,000 deaths (3). 90% of the deaths during the 2017–2018 flu season were within the elderly population, while about 48,000 of the hospitalizations were in children (3). These estimates highlight that young children and especially the elderly are at higher risks for influenza and severe infections that can lead to hospitalization or death. Additionally, the CDC has recommended varying dosages for each vaccine for different age groups due to age-dependent immune responses (3, 9). Due to a decrease in efficacy of the influenza vaccines in the 65 and older population, they receive different dosages compared to younger age groups, in order to elicit a beneficial immune response (3, 9). Contrasting between changes in gene expression due to immunosenescence in healthy subjects and the age-dependent immune responses to diseases such as influenza can help our understanding of how responses to different diseases vary with age. Due to the influenza virus constantly changing and the efficacy of the vaccine being dependent on one's age, researchers have started efforts to develop a universal vaccine (10–12). The goal is for such a universal vaccine to provide protection to all influenza strains (13). One approach, is to implement the use of highly conserved influenza peptides in vaccine formulations (12, 13).
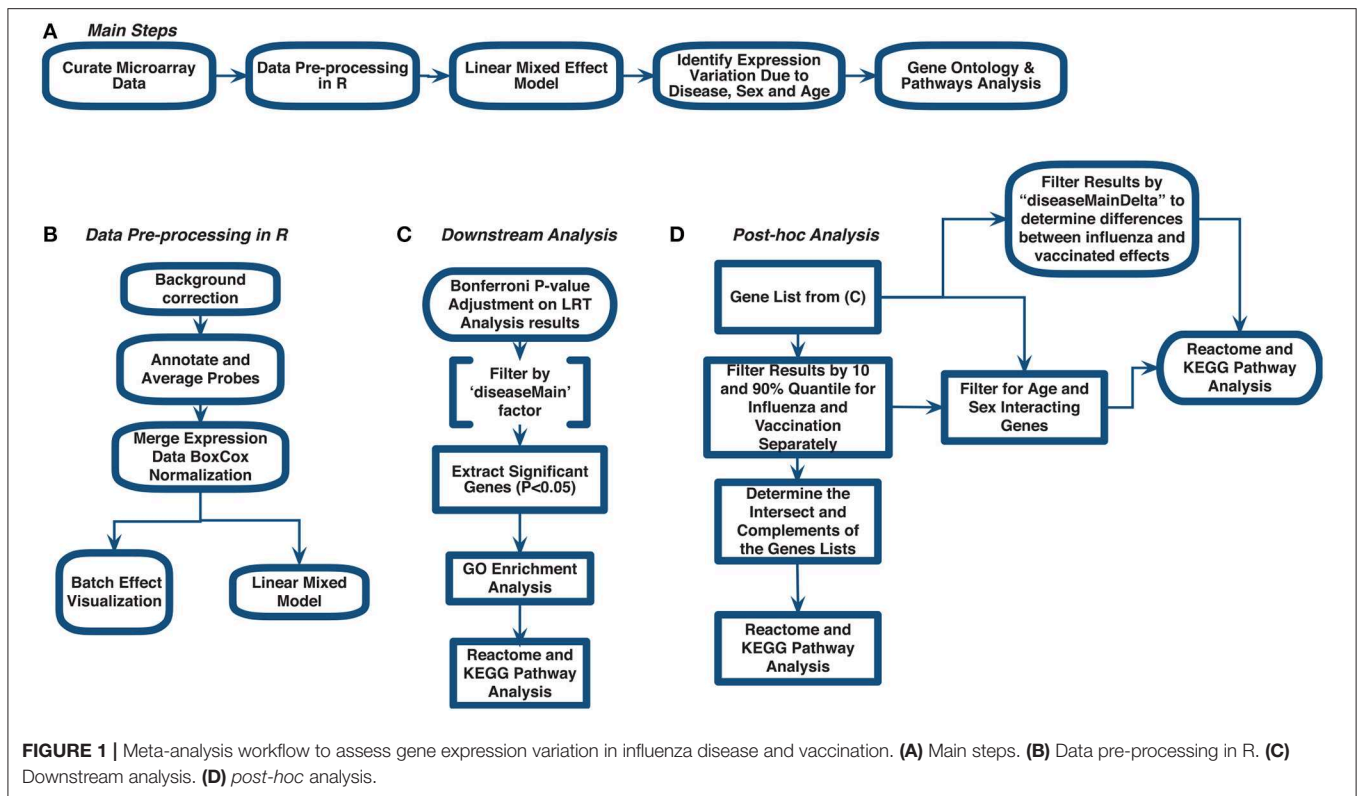
Previous studies have investigated global blood gene expression to compare influenza disease to other respiratory diseases to assess severity and pathogenesis (14). For example, influenza has been shown to induce a stronger immune response than respiratory syncytial virus by producing more respiratory cytokines (14, 15). Studies also explored responses to vaccinations to highlight gene signatures. In our meta-analysis, our aim was to combine publicly available influenza microarray data to identify the effects of disease state (control, influenza infection, and vaccination), age and sex on gene expression. We explored gene expression variation in blood for 3,481 samples (1,277 controls, 297 influenza infected, 1,907 influenza vaccinated) to identify genes and their pathways in influenza (**Figures 1**, **2**). This is to the best of our knowledge, the largest meta-analysis (18 datasets) to explore blood expression changes in influenza infection and vaccination. Our results provide gene signatures and pathways that can be targeted to improve influenza treatment and vaccinations. We also highlight disease associated genes that have interactions with age and sex, that can be used to further explore improving vaccinations, and aid efforts in identifying potential gene targets toward developing universal vaccinations to help reduce the burden of influenza.

2. METHODS

We curated 18 influenza-related microarray datasets from public database repositories (**Table 1**) to investigate changes in gene expression due to disease status, sex, and age. The 18 datasets were from Affymetrix and Illumina microarray platforms (**Table 1**). We modified and implemented the data-analysis pipeline outlined by Brooks et al. (29). To achieve our goal, after curating the datasets, we used the R programming language (30) to pre-process the raw gene expression data and to fit linear mixed effects models to determine statistically significant differentially expressed genes by factor (**Figure 1**). In addition, we identified genes that varied in expression due to disease status, sex, and age, and we also determined which gene ontology (GO) terms and pathways enrichment based on these gene sets (**Figure 1**).

2.1. Data Curation: Gene Expression Omnibus

For our meta-analysis, we focused on influenza infection and vaccination. We searched public database repositories such as Gene Expression Omnibus (GEO) (31), Array Express (AE) (32), and Immune Space (IS) (33, 34) (**Figure 2**). To begin our data search, we found datasets with the keyword "influenza" and filtered for /textitHomo sapiens (**Figure 2**). Following this filter, we then removed duplicate records. For example, there were 15 duplicate records on GEO and 16 datasets on IS overlapped with our GEO records (**Figure 2**). We further filtered the results for datasets that were published, had non-ambiguous annotation, reported the age and sex of all subjects, and used blood or peripheral blood mononuclear cells (PBMCs) as the tissue type (**Figure 2**). Based

FIGURE 1 | Meta-analysis workflow to assess gene expression variation in influenza disease and vaccination. **(A)** Main steps. **(B)** Data pre-processing in R. **(C)** Downstream analysis. **(D)** *post-hoc* analysis.
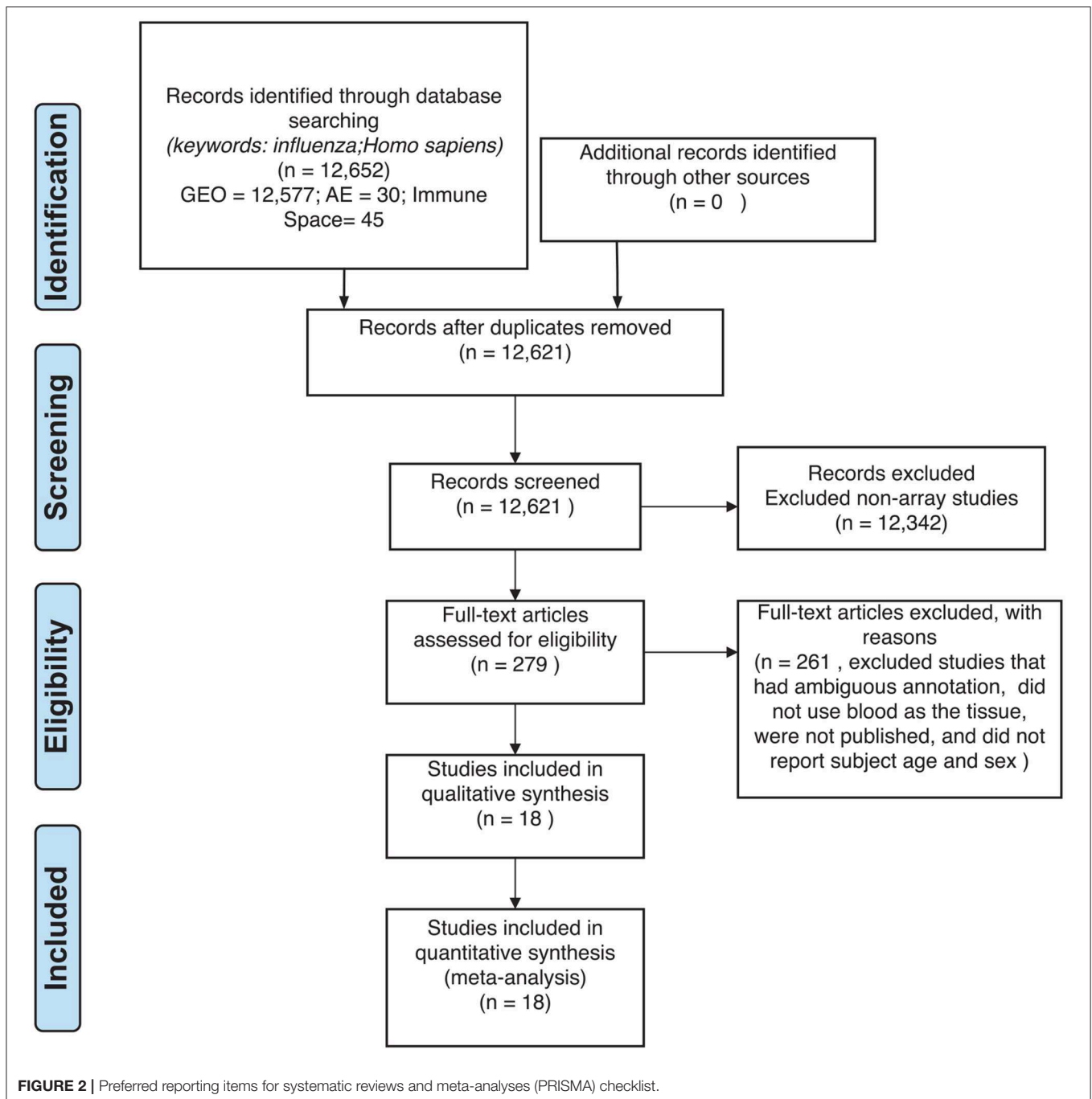
on our inclusion criteria, we identified 18 datasets on GEO to use for our meta-analysis (**Table 1** and SDF1 of online supplementary data files). For datasets such as GSE29614 (SDY64 on IS), GSE29615 (SDY269 on IS), GSE74811 (SDY270 on IS), GSE59654 (SDY404 on IS), GSE74816 (SDY1119 on IS), GSE48023 (SDY1276 on IS), 48018 (SDY1276 on IS) that did not have the ages of the subjects reported on GEO, we used the annotation from IS to gather age and sex characteristics of the samples. Additionally, we excluded 4 duplicates in GSE34205: GSM844139, GSM844141, GSM844143, and GSM844196 (which are duplicates of GSM844138, GSM844140, GSM844142, and GSM844195 datasets, respectively).

After filtering through and selecting the datasets to use in our meta-analysis, we downloaded the raw gene expression data for each dataset, and created a file per study with sample characteristics (**Table 1** and SDF1 of online supplementary data files). Our selected datasets were further filtered to remove samples that did not fit our criteria. For instance, GSE38900 and GSE34205 have samples with respiratory syncytial virus (RSV), GSE48762 contains samples who received the pneumococcal vaccine, GSE50628 has samples with rota-virus infection and patients who experience seizures, and GSE97485 has samples with acute myeloid leukemia who received the influenza vaccine. Due to this, we excluded all subjects that had a pre-existing health condition, infections other than influenza and received vaccinations other than the influenza vaccine (SDF1 of online supplementary data files).

2.2. Data Pre-processing in R and Mathematica

All raw expression files were downloaded directly from the GEO website and pre-processed in R using appropriate packages based on the type of microarray platform (**Table 1**). We carried out background correction and annotated and summarized all probes (**Figure 1B**). We used the `affy` package (35) to pre-process all of the data files for the expression data from Affymetrix Human Genome Plus 2.0 and the Affymetrix HT Human Genome U133 Plus PM. Specifically, we used the `expresso` function to pre-process the files using robust multi-array analysis (RMA) for background correction, conduct perfect-match probe correction, and to calculate expression values using "avdiff" (35). To summarize and remove replicate probes we used the `avereps` function from `limma` (36). For the Affymetrix HT Human Genome U133 Plus PM, we created our own annotation package in R using the annotation obtained from GEO (37). For the raw expression data from the Affymetrix Human Gene 1.1 ST microarray platform, we pre-processed the data using the `oligo` (38) and `affycoretools` (39) packages. To background correct the Affymetrix Human Gene 1.1 ST microarray data files we also used RMA and summarized and removed replicate probes using `avereps` function from `limma`.Our Illumina data files were pre-processed with the `limma` package. We used the NormExp Background Correction (`nec`) function from the `limma` package to remove the background of data files that reported the detection *p*-values. The (`nec`) function using the detection *p*-values when background correcting. Probes were

FIGURE 2 | Preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist.

annotated and summarized using the `aggregate` function from the `stats` package in base R (30, 36).

Following pre-processing, we merged expression data for the 18 datasets (**Table 1** and SDF1 of online supplementary data files) by matching gene symbols that were common across all datasets. We conducted a Box-Cox power transformation (40) and standardized the expression values using the functions `ApplyBoxCoxTransformExtended` and `StandardizeExtended` from the MathIOmica (version 1.2.0) package in Mathematica (41, 42)

(**Figure 1B** and SDF2 of online supplementary data files).

2.3. Linear Mixed Effects Modeling

We fitted a sequence of mixed-effects models to identify genes whose expression levels were affected by disease status (3 levels: control, influenza, vaccine) and those for which the effect of disease was modulated by either age or sex. Models were fitted using the `lmer` function of the `lme4` R-package (43). Separate models were fitted to each of the genes. Our baseline model (M0)

TABLE 1 | Demographics of curated influenza microarray datasets.

Accession number	Controls	Influenza disease	Influenza vaccine	Sex (M/F)	Age range	Platform	References
GSE38900	31	16	0	20/27	0.025–1.57	Illumina HumanHT-12 V4.0 expression beadchip	(14)
GSE107990	171	0	500	238/433	23–89	Illumina HumanHT-12 V4.0 expression beadchip	(16)
GSE111368	130	229	0	177/182	18–71	Illumina HumanHT-12 V4.0 expression beadchip	(17)
GSE27131	7	14	0	16/5	25–59	Affymetrix Human Gene 1.0 ST Array	(18)
GSE29614	9	0	18	12/15	22–46	Affymetrix Human Genome U133 Plus 2.0 Array	(19)
GSE29615	28	0	55	38/45	21–46	Affymetrix HT HG-U133+ PM Array Plate	(19)
GSE47353	117	0	175	122/170	21–62	Affymetrix Human Gene 1.0 ST Array	(20)
GSE48762	274	0	150	202/222	22–49	Illumina HumanHT-12 V3.0 expression beadchip	(21)
GSE50628	0	10	0	2//8	4–9	Affymetrix Human Genome U133 Plus 2.0 Array	(22)
GSE52005	34	0	102	62/74	0.68–14.68	Illumina HumanHT-12 V4.0 expression beadchip	(23)
GSE74816	72	0	105	59/118	21–80	Affymetrix HT HG-U133+ PM Array Plate	(24)
GSE97485	10	0	0	6/4	27–72	Affymetrix Human Gene 1.0 ST Array	(25)
GSE34205	18	28	0	24/22	0.0416–11	Affymetrix Human Genome U133 Plus 2.0 Array	(15)
GSE41080	91	0	0	37/54	20–93	Illumina HumanHT-12 V3.0 expression beadchip	(26)
GSE74811	28	0	55	23/60	21–47	Affymetrix HT HG-U133+ PM Array Plate	(24)
GSE59654	39	0	117	68/88	22–90	Illumina HumanHT-12 V4.0 expression beadchip	(27)
GSE48018	111	0	320	431/0	18.2–32.1	Illumina HumanHT-12 V3.0 expression beadchip	(28)
GSE48023	107	0	310	0/417	18.5–40.2	Illumina HumanHT-12 V4.0 expression beadchip	(28)

included the (fixed) effects of sex (M/F), age (a factor with 4 levels, (−1,3], (3,19], (19,65] and (65,100]), ethnicity (a factor with 7 levels, African-American, Caucasian, Asian, Hispanic, Middle Eastern, Other, Unclassified) and tissue (2 levels, blood and PBMCs) plus the random effects of study (18 levels, see **Table 1** for accession numbers) and of the subject (we included the subject effect because some studies had repeated measures). We first expanded this model by adding the (fixed) main effect of disease status (a factor with three levels, M1). Our next model expanded M1 by adding interactions between disease status and age (M2-DxA) and disease status by sex (M2-DxS). *P*-values for the main effects of diseases as for disease-by-sex and disease-by-age were obtained using likelihood ratio tests (LRT) between the models described above (SDF3 of online supplementary data files). LRTs were implemented using the `anova` function from base R to pairs of models. We used a sequential testing approach where: (i) we first identified genes with significant main effect of disease (this was based on a LRT between M1 and M0), (ii) among genes with significant main effect of diseases we tested the significance of DxA and DxS using a likelihood ratio test that had M1 as null hypothesis and the interaction models as alternative hypotheses. *P*-values were adjusted using Bonferroni adjustment, where for the first test (i) the number of tests was equal to the number of genes, and for the second one (ii) the number of tests was equal to the number of genes that passed the first test.

The filtering of genes based on Bonferroni-adjusted *p*-values for the main effect of disease (comparison of M1 to M0) allowed us to identify differentially expressed genes with respect to disease states (**Figure 1**). Using this gene list, we then conducted GO enrichment analysis (`GOAnalysis` function in `MathIOmica` package) and pathway enrichment analysis using Kyoto Encyclopedia of Genes and Genomes (KEGG, using the `KEGGAnalysis` functions in `MathIOmica`), and Reactome

pathway enrichment analysis [`enrichPathway` function from the `ReactomePA` package in R (44)].

2.4. Determining Gene Expression Variability Between Influenza Infection and Vaccination

We took a sequential testing approach to further analyze the identified statistically significant disease genes (SDF5 of online supplementary data files). Using this gene list, we further filtered for biological effect by using calculated estimates (which compared influenza and vaccine expression to controls) (SDF4 of online supplementary data files) and performed a two-tailed 10% quantile filter (i.e., 0.1 and 0.9 quantiles) to determine genes that were biologically significant in subjects who were vaccinated with influenza vaccinated and subjects infected with influenza disease. The biologically significant gene lists for the vaccinated and influenza subjects were further examined to identify genes in common, and genes only in the influenza list, and only in the vaccinated list (**Figure 1**). We performed GO and pathway enrichment analysis on these genes. Lastly, we filtered the disease (see SDF1 of online supplementary data files) statistically significant gene list for interacting genes between disease and age (age groups: (−1,3], (3,19], (19,65], (65,100]) and disease and sex.

3. RESULTS

Our data curation criteria resulted in 3,481 samples (1,277 controls, 297 influenza infection, 1,907 influenza vaccinated, 1,537 males and 1,944 females) (see SDF1 of online supplementary data files). Our 3,481 samples are from 1,147 individuals. Some studies include repeated measures (in the

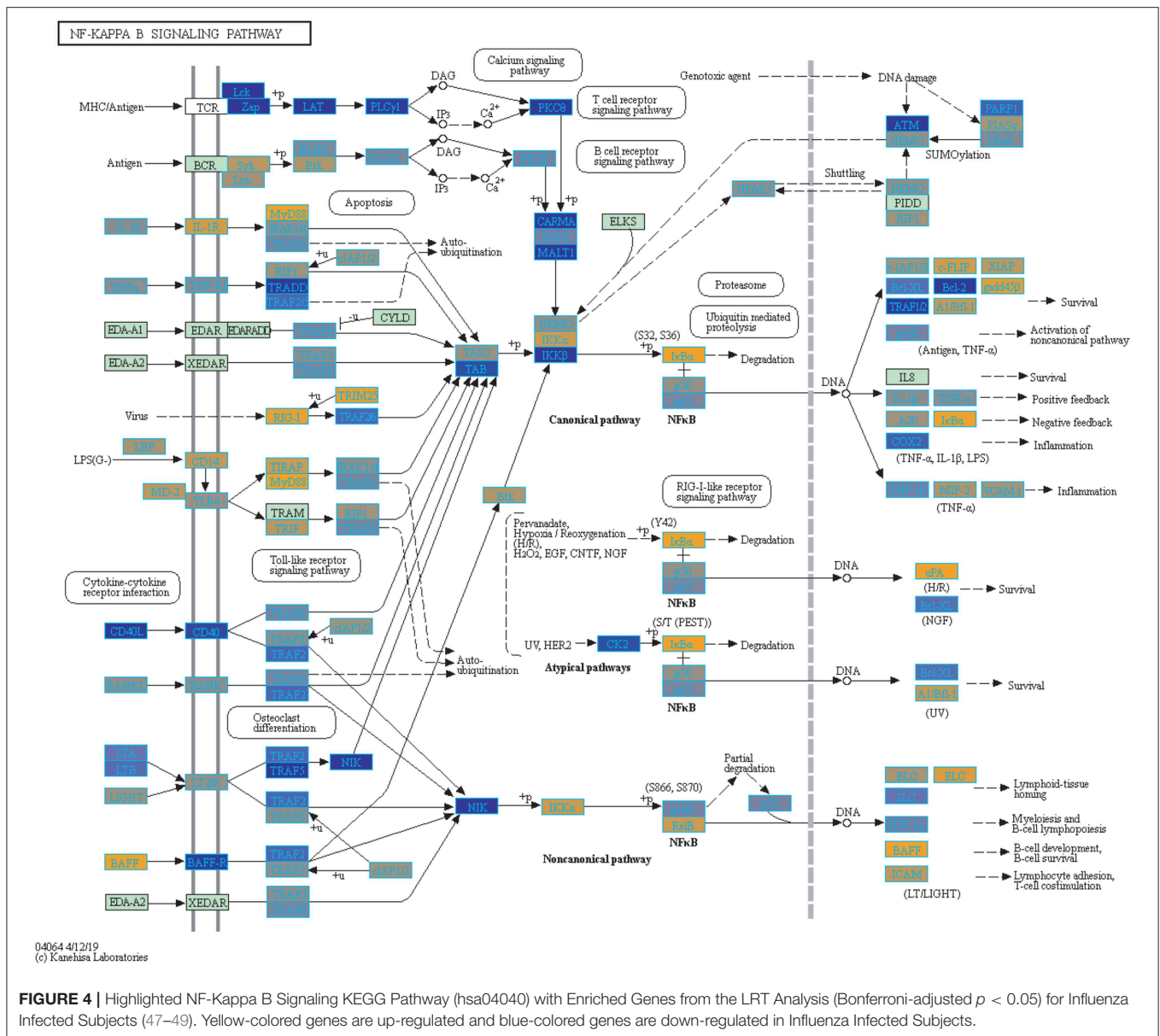**FIGURE 3** | Flowchart of gene filtering steps for influenza meta-analysis.
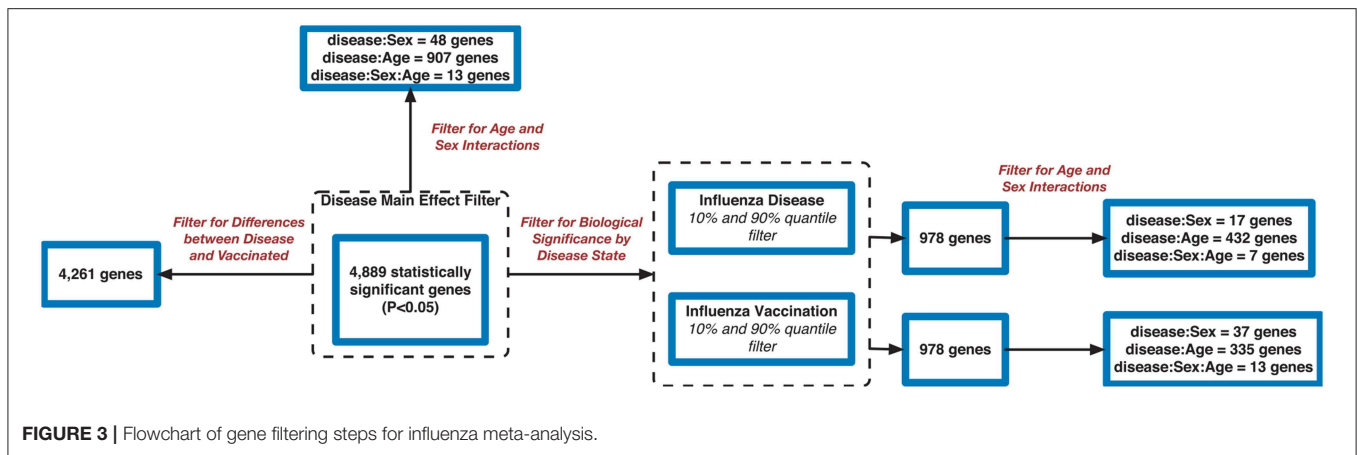
curated studies individuals were followed for several days after vaccination or infection and varying timepoints were reported as a different samples for the same subject). We included all repeated measures in our downstream analysis and accounted for them in our model. The main results are summarized below, and further discussed in the section 4.

3.1. Differentially Expressed Genes in Influenza Disease and Vaccination

Filtering our LRT analysis results by disease factor (see SDF3 of online supplementary data files) for Bonferroni adjusted p-values (< 0.05), we identified 4,889 statistically significant disease genes (see SDF5 of online supplementary data files and **Figure 3** for downstream analysis details). We performed GO enrichment analysis using BINGO in Cytoscape (version 3.7) (45, 46) and pathway enrichment analysis on the 4,889 genes (**Supplementary Figures 1–5** and see SDF6-SDF8 of online supplementary data files). We identified enriched GO terms such as: cell cycle checkpoint (51 genes), response to stimulus (987 genes), immune response (243 genes), transcription (122 genes), regulation of T-cell activation (62 genes), regulation of defense response to virus by host (8 genes) and immune system process (379 genes) (see SDF8 of online supplementary data files for full table). We found 75 enriched KEGG pathways (SDF6 of online supplementary data files). The enriched KEGG pathways include: Cell cycle (68 gene hits), Hematopoietic cell lineage (45 genes), NF-kappa B signaling pathway (46 genes), Metabolic pathways (341 genes), Primary immunodeficiency (23 genes), T cell receptor signaling pathway (44 genes), B cell receptor signaling pathway (29 genes) and also Influenza A (52 genes). We also highlighted the NF-kappa B signaling pathway and the Influenza A KEGG pathways that are relevant to disease with our calculated estimates which compared influenza infection and vaccination expression to that of healthy controls (**Figures 4–7**).

In addition, we filtered the 4,889 genes for effect size to determine biological significance of the genes (SDF4 of online supplementary data files). We used a two-tailed 10 and 90% quantile filter on the 4,889 genes to: (i)analyze the influenza disease estimates (compared expression to control) list to identify genes that are biologically significant and statistically significant (Bonferroni-adjusted $p < 0.05$) in influenza infection (ii) analyze the influenza vaccination estimates with the same filtering approach to also identify significant genes for influenza vaccination. For influenza infection our 10 and 90% quantile cut-offs for biological significance were ≤ -0.6724464 and ≥ 0.5949655, respectively. For influenza vaccination, the 10% and 90% quantile cut-offs were ≤ -0.07157763 and ≥ 0.06719048, respectively. For influenza infection we identified 978 genes of the 4,889 to be biologically significant (**Table 2** and SDF9 of online supplementary data files), and we also identified 978 genes to be biologically significant for influenza vaccination (**Table 2** and SDF10 of online supplementary data files). We then compared the two gene lists to identify the intersection (genes in common), genes only in the influenza disease list, and genes only in the influenza vaccination list (**Figure 1D** and SDF11-SDF13 of online supplementary data files). There were 334 genes in common across both lists (influenza disease and vaccination) (SDF17 of online supplementary data files) that resulted in enriched Reactome pathways such as Interferon alpha/beta signaling (14 genes), Interferon gamma signaling (12 genes), Antiviral mechanism by IFN-stimulated genes (9 genes), and Cell Cycle Checkpoints (17 genes) (SDF20 of online supplementary data files). There were 644 genes that were only in influenza infection list (SDF18 of online supplementary data files) that resulted in enriched Reactome pathways including: Neutrophil degranulation (45 genes), Cell Cycle Checkpoints (27 genes), Amplification of signal from the kinetochores (13 genes), Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal (13 genes) and Mitotic Spindle Checkpoint (14 genes) (SDF21 of online supplementary data files). Also, we identified another 644 genes that were only in the biologically significant list for the vaccinated subjects (SDF19 of online supplementary data files). Enriched Reactome pathway analysis on these genes resulted in pathways such as Interferon Signaling (24 genes), Antigen processing-Cross presentation (14 genes), ER-Phagosome pathway (12 genes), Binding and Uptake of Ligands by Scavenger Receptors (8 genes) and Class I MHC mediated antigen processing & presentation (30 gene) (SDF22 of online supplementary data files).

FIGURE 4 | Highlighted NF-Kappa B Signaling KEGG Pathway (hsa04040) with Enriched Genes from the LRT Analysis (Bonferroni-adjusted $p < 0.05$) for Influenza Infected Subjects (47–49). Yellow-colored genes are up-regulated and blue-colored genes are down-regulated in Influenza Infected Subjects.
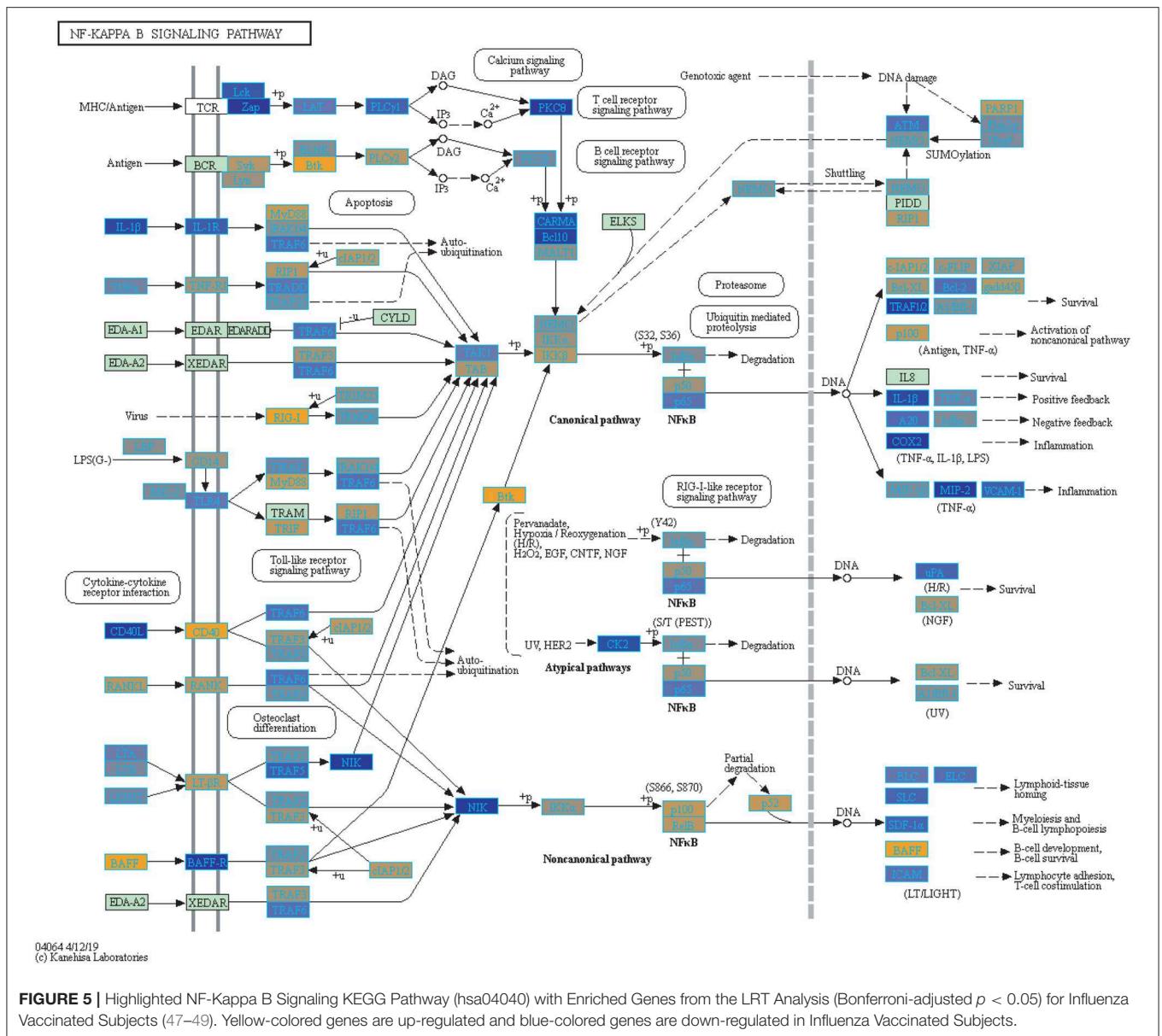
We also explored the 4,889 genes to identify how many genes were different in gene expression when looking at influenza infected subjects compared to influenza vaccinated subjects. Of the 4,889 genes, 4,261 genes showed statistically significant differences between vaccination and infection with influenza (**Figure 3** and SDF25–SDF27 of online Supplementary Data Files).

3.2. Age and Sex Effect on Gene Expression in Influenza

Using the 4,889 genes disease significant genes from above, we Bonferroni-adjusted the p-values for both the age and sex factors. We then further filtered the 4,889 list by the age factor p-values (Bonferroni-adjusted $p < 0.05$) to identify statistically significant interacting genes between disease state and age (DxA).

We also repeated this approach for the sex factor interaction with disease (DxS). Of the 4,889 statistically significant (Bonferroni-adjusted $p < 0.05$) disease genes, 907 of them had a statistically significant interaction with disease and age (SDF28 of online supplementary data files). KEGG enrichment, our results include: Cytokine-cytokine receptor interaction (34 genes), T cell receptor signaling pathway (19 genes), Natural killer cell mediated cytotoxicity (19 genes), Intestinal immune network for IgA production (11 gene hits), Hematopoietic cell lineage (14 genes), Primary immunodeficiency (8 genes), NF-kappa B signaling pathway (13 genes), and Influenza A (16 genes) (SDF30 of online supplementary data files, **Table 3**). We also looked at the biologically significant gene lists for influenza infection and vaccination (based on effect as discussed above) to determine which of these genes also had a significant interaction with

FIGURE 5 | Highlighted NF-Kappa B Signaling KEGG Pathway (hsa04040) with Enriched Genes from the LRT Analysis (Bonferroni-adjusted $p < 0.05$) for Influenza Vaccinated Subjects (47–49). Yellow-colored genes are up-regulated and blue-colored genes are down-regulated in Influenza Vaccinated Subjects.

disease and age. Of the 978 in the influenza infection biologically significant list, 432 had a statistically significant (Bonferroni-adjusted $p < 0.05$ for disease and age factor) interaction with disease and age (**Figure 3** and SDF32 of online Supplementary Data Files). In the biologically significant gene list for influenza vaccinated subjects 335 genes also had a statistically significant (Bonferroni-adjusted $p < 0.05$ for disease and age factor) interaction with disease and age (**Figure 3** and SDF35 of online Supplementary Data Files).

Furthermore, we explored differences in gene expression (based on mean differences across groups) in subjects with influenza infection, influenza vaccination and controls across the 4 age groups: $(-1,3]$, $(3,19]$, $(19,65]$, $(65,100]$ using the gene lists of identified disease:age interacting genes. First we calculated the mean expression for control subjects younger

than 3 (age group: $(-1,3]$). This served as our baseline for all comparisons to influenza infection and vaccination. We calculated the difference in means for the subjects within the other age groups only focusing on the healthy subjects and used the younger than 3 as our baseline to find the difference in means (**Figure 8**). We also calculated the difference in mean expression for all influenza infected subjects and used the influenza infected subjects younger than 3 as the baseline for comparisons of relative expression (**Figure 9A**). In addition, we calculated the difference in means by comparing influenza infected samples to the control baseline (younger than 3) (**Figure 9B**). We repeated the above steps with our vaccinated subjects to explore how expression changes with age and disease state (**Figure 10**). We also plotted the difference in means comparing influenza vaccinated subjects to influenza infected subjects to
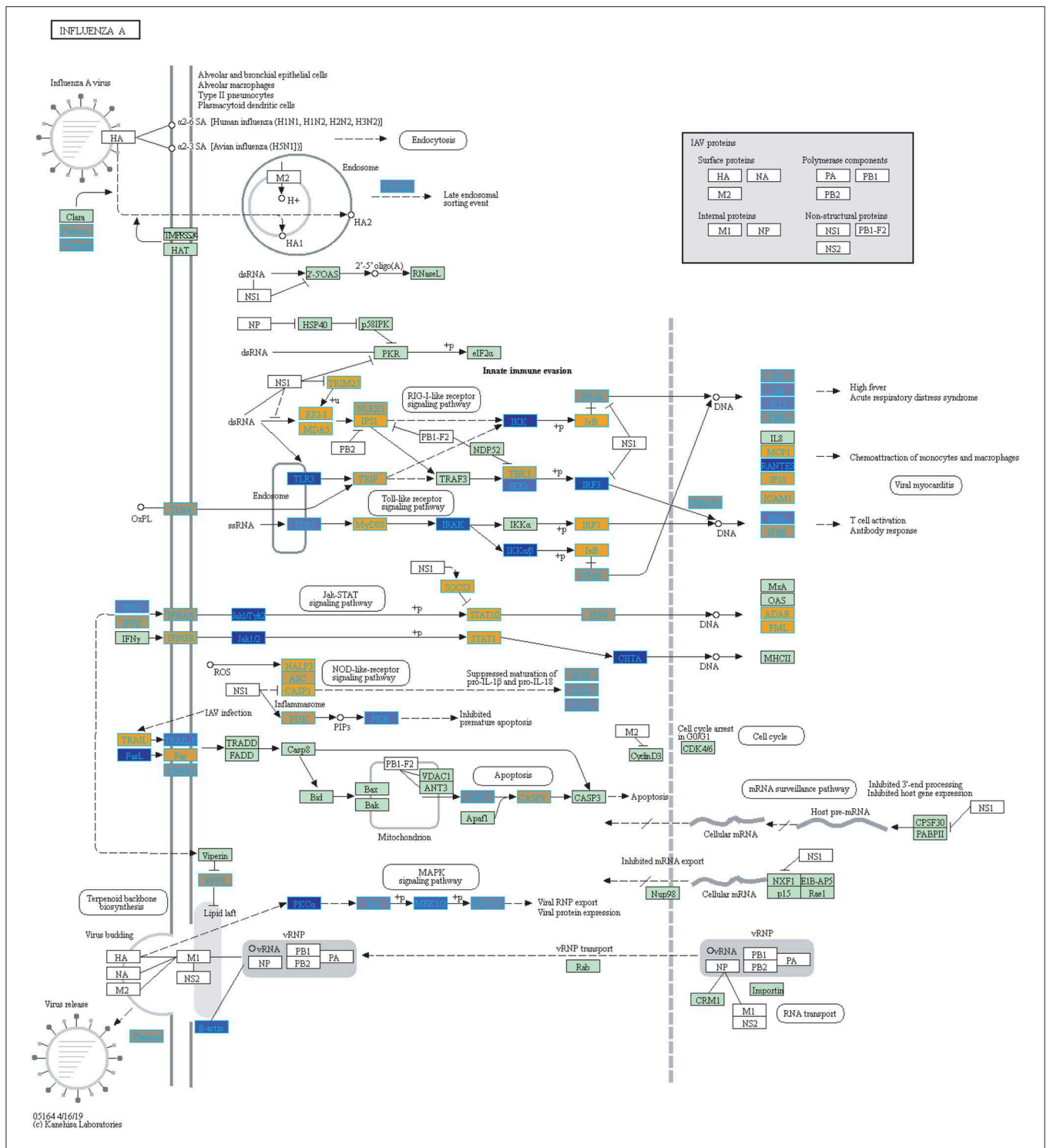
FIGURE 6 | Highlighted Influenza A KEGG Pathway (hsa05164) with Enriched Genes from the LRT analysis (Bonferroni-adjusted $p < 0.05$) for Influenza Infected Subjects (47–49). Yellow-colored genes are up-regulated and blue-colored genes are down-regulated in Influenza Infected Subjects.

highlight temporal patterns of the 907 interacting (disease:age) genes (**Figure 11**).

We also filtered our gene lists (statistically significant disease genes and the biologically significant for influenza disease and vaccination gene lists) for genes with a statistically significant

disease interaction with sex (**Figure 3**). We identified 48 of the 4,889 disease genes (Bonferroni-adjusted $p < 0.05$ for disease and sex factor) that interacted with disease and sex (**Figure 3** and SDF29 of online supplementary data files). In the influenza infected biologically significant gene list there were 17 genes
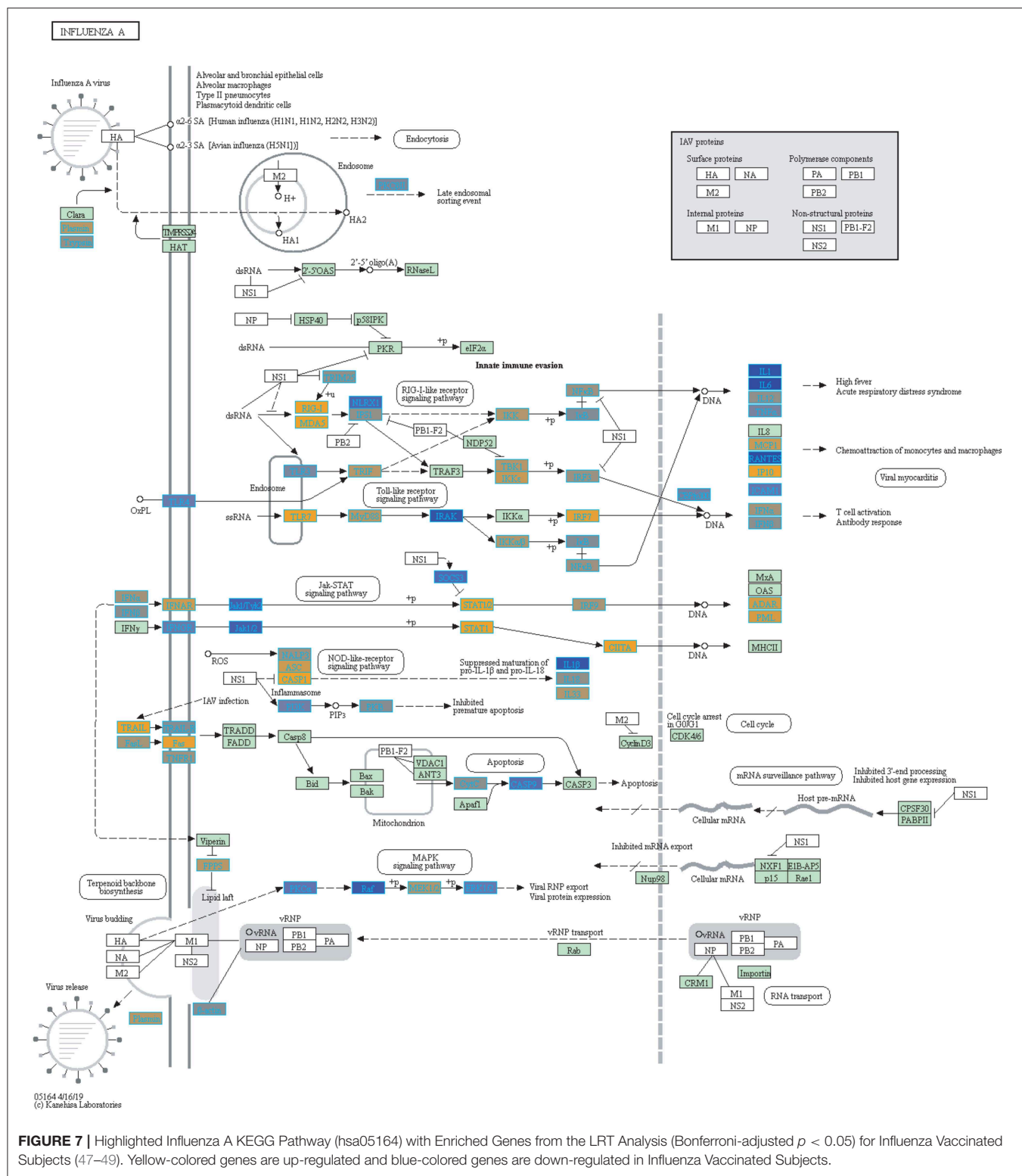
FIGURE 7 | Highlighted Influenza A KEGG Pathway (hsa05164) with Enriched Genes from the LRT Analysis (Bonferroni-adjusted $p < 0.05$) for Influenza Vaccinated Subjects (47–49). Yellow-colored genes are up-regulated and blue-colored genes are down-regulated in Influenza Vaccinated Subjects.

that interacted with disease and sex (Bonferroni-adjusted $p < 0.05$ for disease and sex factor), and 7 genes had an interaction with disease, sex and age (Bonferroni-adjusted $p < 0.05$ for disease, sex and age factor) (**Figure 3** and see also SDF33 and

SDF34 of online supplementary data files). We did not find any statistically significant enrichment in pathways for these genes. As for the biologically significant influenza vaccination genes, 37 of them were associated with disease and sex interactions

TABLE 2 | Top 10 up- and down- regulated differentially expressed genes from the influenza infected and influenza vaccination biologically significant gene lists (based on estimates).

INFLUENZA INFECTION			
Down-regulated		Up-regulated	
Gene	Difference of means	Gene	Difference of means
NELL2	−1.687	UGCG	2.005
UBASH3A	−1.583	CD177	1.875
ABCB1	−1.513	OTOF	1.844
PID1	−1.457	HP	1.625
CACNA2D3	−1.428	SSH1	1.491
PTGDR	−1.423	DTL	1.431
CD40LG	−1.392	GPR84	1.428
PTGDR2	−1.390	HJURP	1.420
TLE2	−1.379	CDC45	1.395
NCR3	−1.357	SLC1A3	1.390
INFLUENZA VACCINATION			
Down-regulated		Up-regulated	
Gene	Difference of means	Gene	Difference of means
TOP1MT	−0.179	GBP1	0.354
ARNTL	−0.176	MYOF	0.347
DIDO1	−0.172	STAT1	0.284
PDE4D	−0.169	PSTPIP2	0.281
TMX4	−0.168	SAMD9L	0.276
ZNF589	−0.166	OAS3	0.269
SLC37A3	−0.165	WARS	0.263
GNB5	−0.165	BATF2	0.263
ENO2	−0.162	ANKRD22	0.256
AP3M2	−0.159	C1QB	0.255

(Bonferroni-adjusted $p < 0.05$ for disease and sex factor), and 13 genes had associated interactions with disease, sex and age (Bonferroni-adjusted $p < 0.05$ for disease, sex and age factor) (**Figure 3** and see also SDF36 and SDF37 of online supplementary data files). We also did not find any enriched pathways for these genes.

4. DISCUSSION

Every year there is a new vaccine available to reduce the amount of influenza cases worldwide. The influenza virus is constantly changing and researchers have to predict the most common strains that will affect the population each season. During the flu season, the majority of hospitalizations and deaths from influenza are within the elderly population [3]. Young children are also at high risk for severe infections of influenza due to their underdeveloped immune system [50]. Current vaccine development methods, though effective are also flawed. In some cases, the influenza strains can mutate after the strains for the vaccine have been selected for the upcoming flu season, which then reduces the effectiveness of the vaccine [51]. Exploring

how gene expression varies in influenza infection, vaccination, and comparison of the differences may highlight prospective biomarkers/gene signatures for improving vaccinations. In addition, because of the observed age-dependency in influenza infection, investigating gene expression temporal patterns across various ages can also provide insight on how genes change due to underdevelopment and immunosenescence.
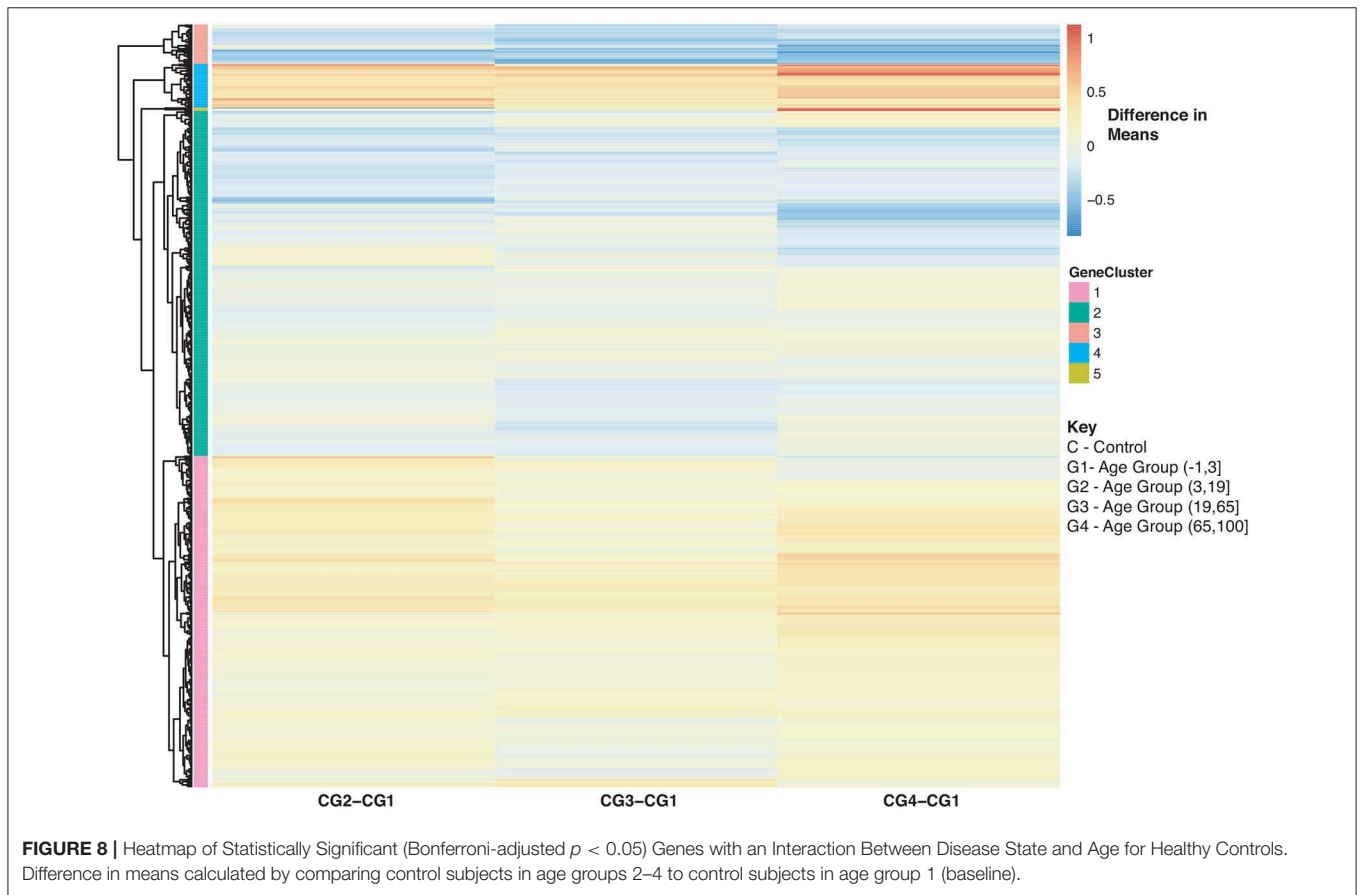
We identified 18 microarray expression datasets that passed our inclusion criteria for a meta-analysis on influenza (**Table 1**). We collected the raw expression microarray data for all datasets, pre-processed them and combined by common gene names. With 3,481 samples (including repeated measures) we modeled the pre-processed expression data with a mixed effects model and carried out LRT analysis. Our LRT analyses resulted in 4,889 statistically significant (Bonferroni-adjusted $p < 0.05$) disease genes (see SDF5 of online supplementary data files). These results include CD177 which plays a role in innate immune response by regulating chemotaxis of neutrophils [52, 53], BCL11B which regulates T-cell differentiation [53, 54], HMGB1 protein has been shown to promote viral replication [55] and plays a role in inflammation [53], TPP2 plays a role in major histocompatibility complex (MHC) presentation and TANK is involved in NF-kappa B signaling.
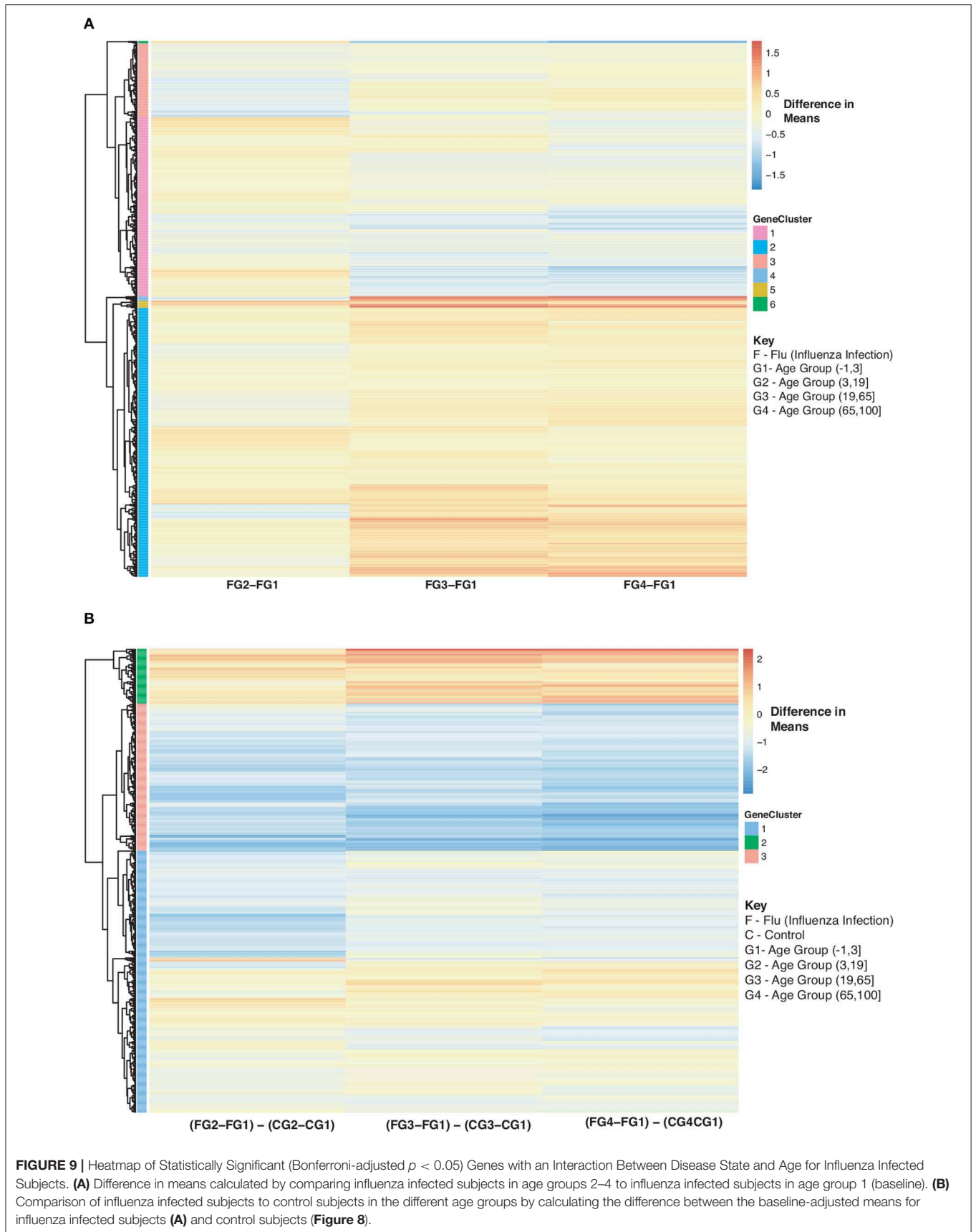
We highlighted the KEGG NF-kappa B signaling pathway using the estimates from influenza infection and vaccination (**Figures 4**, **5**). The NF-kappa B pathway is activated during influenza infection which up-regulates antiviral genes [56] and can regulate viral synthesis [57]. Previous studies have also reported that the influenza virus is capable of regulating antiviral activity by NF-kappa B and promote replication in hosts [57]. In the NF-kappa B pathway, we observed similar expression patterns between disease and vaccinated subjects, including down regulation of genes involved in MHC/Antigen presentation for both physiological states. There are also some differences in gene expression observed such as CD40 and PARP1 up-regulated in vaccinated samples. CD40 has previously been shown to regulate immune response and promotes protection against the virus [58, 59] while PARP1 has been highlighted as a host factor that can regulate the polymerase activity of influenza [60]. In **Figure 5**, the genes in our vaccine list in the RIG-I-like receptor signaling pathway are down-regulated, compared to influenza infected subjects (**Figure 4**). The RIG-I-like receptors have been previously shown to be involved in sensing viral RNA and regulating an antiviral immune response [61]. Other genes such as ICAM which is involved in lymphocyte adhesion and T-cell costimulation, and BLC and ELC involved in lymphoid tissue homing are all down-regulated in vaccinated subjects compared to infected subjects (**Figures 4**, **5**). We also highlighted expression of genes in the Influenza A KEGG pathway for influenza infected and influenza vaccinated (**Figures 6**, **7**). Although there are similarities in both figures, some key differences in expression are observed in genes connected with high fever (IL-1 and IL-6). Studies have shown elevated levels of IL-1β and IL-6 following infection with influenza A [17, 62].
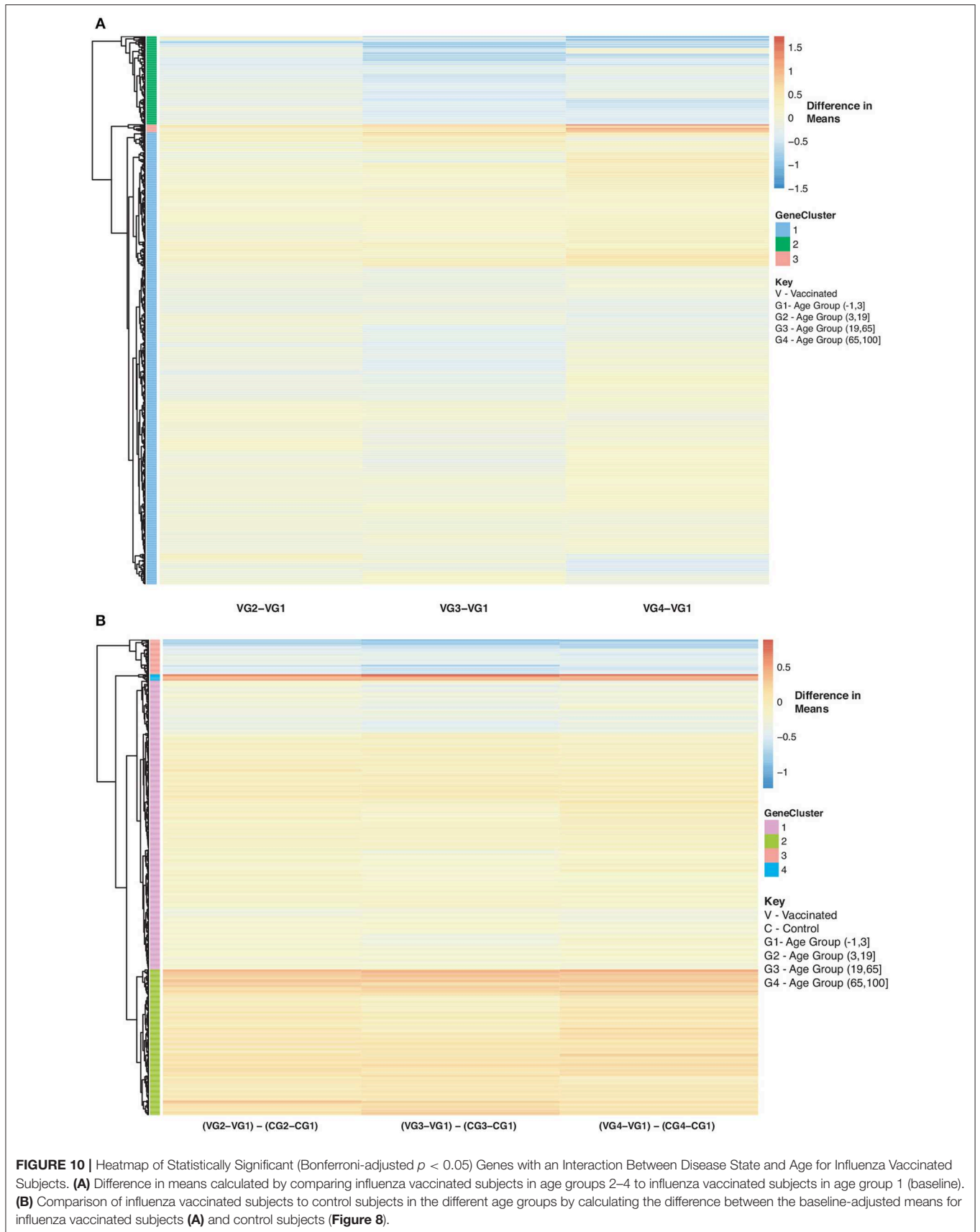
Additionally, we compared our biologically significant gene list for influenza infection (978 genes) to Dunning et al. who identified whole blood RNA signatures in hospitalized adults
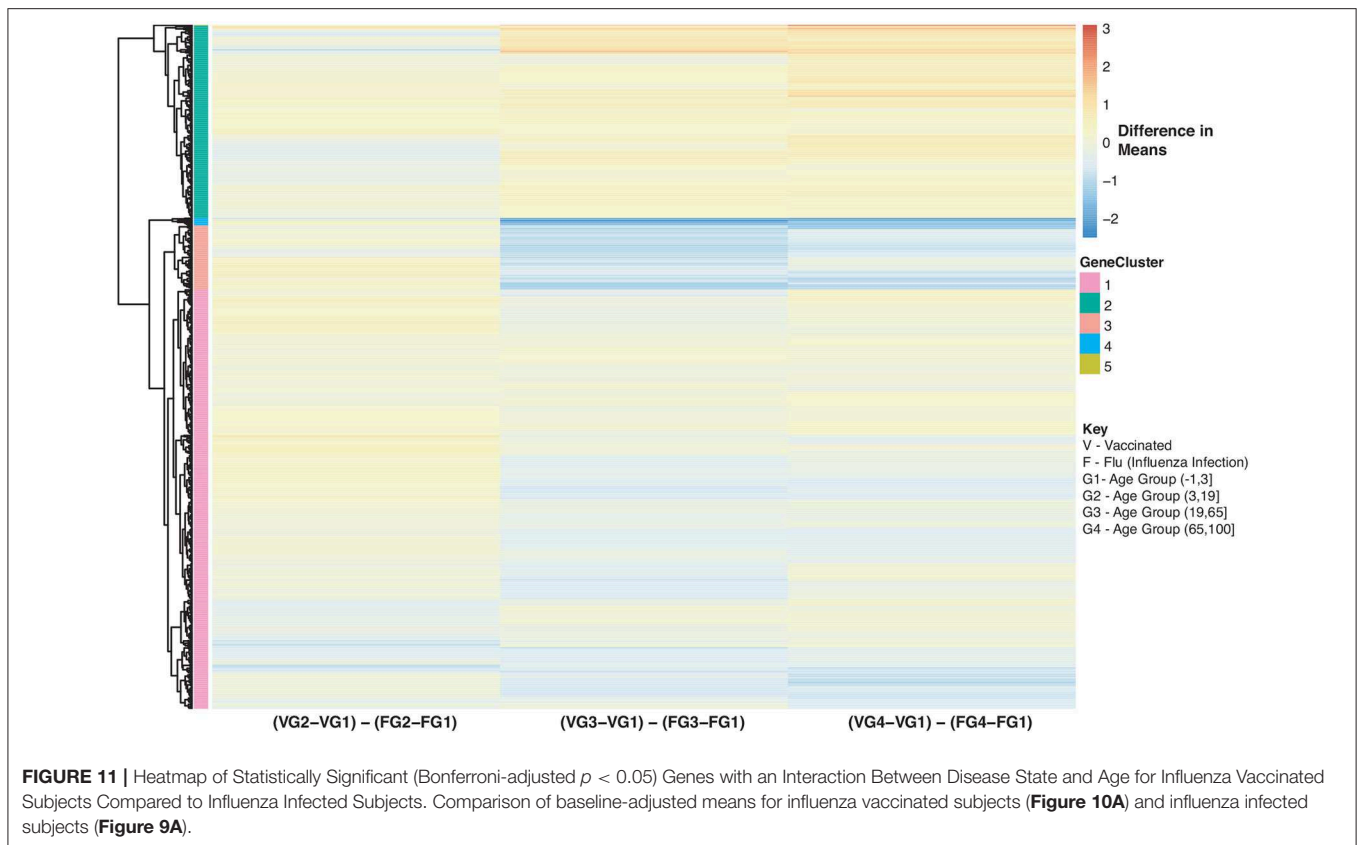
TABLE 3 | Enriched KEGG pathways from statistically significant genes with an interaction between disease status and age.

KEGG ID	KEGG pathway	Gene count	p-value	adjusted p-value
path:hsa04060	Cytokine-cytokine receptor interaction	34	5.5E-09	1.4E-06
path:hsa04660	T cell receptor signaling pathway	19	6.9E-08	8.7E-06
path:hsa04650	Natural killer cell mediated cytotoxicity	19	3.8E-06	3.2E-04
path:hsa04672	Intestinal immune network for IgA production	11	5.9E-06	3.8E-04
path:hsa04640	Hematopoietic cell lineage	14	1.8E-05	9.1E-04
path:hsa05340	Primary immunodeficiency	8	1.3E-04	4.8E-03
path:hsa04064	NF-kappa B signaling pathway	13	1.4E-04	4.8E-03
path:hsa04622	RIG-I-like receptor signaling pathway	11	1.6E-04	4.8E-03
path:hsa04068	FoxO signaling pathway	16	1.7E-04	4.8E-03
path:hsa05166	HTLV-I infection	23	6.4E-04	1.5E-02
path:hsa05162	Measles	15	6.4E-04	1.5E-02
path:hsa04062	Chemokine signaling pathway	18	9.8E-04	2.1E-02
path:hsa05330	Allograft rejection	7	1.1E-03	2.2E-02
path:hsa04380	Osteoclast differentiation	14	1.4E-03	2.6E-02
path:hsa05320	Autoimmune thyroid disease	8	1.9E-03	3.2E-02
path:hsa04110	Cell cycle	13	2.3E-03	3.6E-02
path:hsa04010	MAPK signaling pathway	21	2.8E-03	4.1E-02
path:hsa04630	Jak-STAT signaling pathway	15	2.9E-03	4.1E-02
path:hsa05164	Influenza A	16	3.3E-03	4.4E-02

FIGURE 8 | Heatmap of Statistically Significant (Bonferroni-adjusted $p < 0.05$) Genes with an Interaction Between Disease State and Age for Healthy Controls. Difference in means calculated by comparing control subjects in age groups 2–4 to control subjects in age group 1 (baseline).

FIGURE 9 | Heatmap of Statistically Significant (Bonferroni-adjusted $p < 0.05$) Genes with an Interaction Between Disease State and Age for Influenza Infected Subjects. **(A)** Difference in means calculated by comparing influenza infected subjects in age groups 2–4 to influenza infected subjects in age group 1 (baseline). **(B)** Comparison of influenza infected subjects to control subjects in the different age groups by calculating the difference between the baseline-adjusted means for influenza infected subjects **(A)** and control subjects (**Figure 8**).

FIGURE 10 | Heatmap of Statistically Significant (Bonferroni-adjusted $p < 0.05$) Genes with an Interaction Between Disease State and Age for Influenza Vaccinated Subjects. **(A)** Difference in means calculated by comparing influenza vaccinated subjects in age groups 2–4 to influenza vaccinated subjects in age group 1 (baseline). **(B)** Comparison of influenza vaccinated subjects to control subjects in the different age groups by calculating the difference between the baseline-adjusted means for influenza vaccinated subjects **(A)** and control subjects (**Figure 8**).

FIGURE 11 | Heatmap of Statistically Significant (Bonferroni-adjusted $p < 0.05$) Genes with an Interaction Between Disease State and Age for Influenza Vaccinated Subjects Compared to Influenza Infected Subjects. Comparison of baseline-adjusted means for influenza vaccinated subjects (**Figure 10A**) and influenza infected subjects (**Figure 9A**).

with influenza. Their findings indicated that genes involved in interferon-related pathways were activated at the start of the infection and by day 4 had started to decrease with a shift in inflammatory and neutrophil related pathways (17). Our findings also indicate enrichment for neutrophil related pathways in the case of influenza infection. Dunning et al. list a top 25 gene set (controls vs. influenza subjects), from which 22 genes overlap with our findings (978 genes, see SDF9 of online supplementary data files). 5 of our top 10 up-regulated gene list overlap with the Dunning et al. 25-gene set (**Table 2**), namely UGCG, CD177, OTOF, HP and SSH1.

The mechanistic role of host expressed genes in influenza has been investigated using knockdown experiments including genes such as IRF7, LAMP3, and DPF2 which in our study were differentially expressed in either influenza, vaccine or both. IRF7 and LAMP3 are members of our influenza and vaccination biologically significant gene lists, while DPF2 was statistically significant with respect to disease state. IRF7 is involved in regulation of type 1 interferon immune responses to DNA and RNA viruses (53, 63). Deletion of IRF7 in mice resulted in increased host susceptibility to H1N1 and in mortality compared to wild type mice (64). The knockdown of IRF7 in canine kidney cells also resulted in increased viral load compared to the wild type (65). Furthermore, a study by Ciancanell et al. (66) reported compound heterozygous null mutations in IRF7 in a single healthy child that had life-threatening influenza infection, and whose cells had low expression of type I and III

IFNs. LAMP3 knockdown on the other hand led to reduced production of viral nucleoproteins and inhibited viral replication (67). Through a knockdown study, DPF2 was identified as a host factor that promotes expression of viral proteins and host immune system evasion (68). Knocking down DPF2 caused a decrease in the expression of viral proteins (68). Animal models that further explore the mechanisms of the gene lists reported here via knockdown experiments may help to fully characterize the mechanistic role that individual genes play in influenza infection. This in turn will help identify gene targets for the design of more effective vaccines.

Furthermore, our identified biologically significant gene lists for influenza infection and vaccination (using a 2-tailed 10% quantile filter on expression estimates of effect size compared to healthy control) have 334 genes in common, with 644 genes being unique to influenza infection and 644 being unique to influenza vaccination (SDF17-SDF19 of online supplementary data files). Following pathway enrichment, we observed that the genes that are unique to each disease state (influenza infected and vaccinated) are involved in different processes. For example, the biologically significant genes only in influenza infected samples were enriched in pathways such as neutrophil degranulation and cell cycle checkpoints (SDF21 of online supplementary data files). Neutrophil degranulation is a defensive process neutrophils undergo to protect the host against invading pathogens. On the other hand, pathways involved in interferon signaling

and antigen processing were enriched for the genes only in the vaccinated gene list (SDF22 of online supplementary data files). This indicates that with the actual infection the body undergoes different processes to that induced by vaccination (**Supplementary Figures 3, 4** and see SDF23 and SDF24 of online supplementary data files).

The 48 genes for which we identified a statistically significant interaction between disease and sex are highlighted in SDF29 of the online supplementary data files. Sex-specific gene expression has been previously observed in influenza. Studies have observed females exhibited a stronger immune response to influenza vaccine compared to males within the first day (69). Another study suggested that males have a stronger immune response to influenza infection (70). These findings indicate sex and influenza effects are still to be explored and our gene list may offer new candidates to be investigated for their role in influenza.

With regards to aging and influenza, the statistics of the disease burden indicates specific age groups are at higher risk for infection (3). This is in part due to immune system development and deterioration. For example, B and T cell function diminishes with age (50, 71). In our analysis, we identified 907 disease-associated genes with a statistically significant interaction with age that were also enriched in immune related KEGG pathways (SDF30 of online supplementary data files). **Figure 8** compares the mean differences of healthy subjects to the baseline (healthy children younger than 3). There are 4 major groups (**Figure 8** and see SDF47 of online supplementary data files): with reference to **Figure 8**, genes in Cluster 1 were up-regulated compared to the baseline for all age comparisons, Cluster 2 and 3 genes were generally down-regulated compared to the baseline, and Cluster 4 genes are up-regulated and increase with age. Genes in Cluster 1 and 2 are involved in Reactome pathways such as cytokine signaling, interferon signaling and the immune system. Cluster 3 genes are involved in Reactome pathways such as interferon signaling and cell cycle while Cluster 4 genes are involved in cellular senescence, signaling by interleukins and immune system.

In **Figure 9**, we further explored changes in gene expression across age groups due to influenza infection of our 907 disease:age interacting genes. **Figure 9A** is compares influenza infected subjects in age groups 2,3 and 4 to the baseline (infection subjects under 3). In **Figure 9A** there are three major groups (cluster numbering with respect to **Figure 9A**): Cluster 1 (gradual decrease with age), 2 (genes up-regulated with increase in age), and 3 (gradual down-regulation with age). Genes in Cluster 1 are in Reactome pathways such as cytokine signaling, interferon signaling, antiviral mechanism by IFN-stimulated genes and chemokine receptors. Cluster 2 genes are involved in regulatory T lymphocytes, transcription, protein repair, and interleukin-2 signaling while Cluster 3 genes are involved in gene transcription (see SDF48 of online supplementary data files). **Figure 9B** instead compares influenza infected subjects to controls by looking at difference in means. There are three groups of expression patterns (cluster numbering with respect to **Figure 9B**): Cluster 1 shows a gradual increase with age, in Cluster 2 expression intensifies with age and in Cluster 3 genes are down-regulated compared to the control subjects younger than 3 (see SDF49 of

online supplementary data files). Genes in Clusters 1 and 2 were not in any enriched Reactome pathways but are associated with transcription and signaling pathways. Genes in Cluster 3 were in Reactome pathways that include cytokine signaling, interferon signaling, antiviral mechanism by IFN-stimulated genes and chemokine receptors.

As for the vaccinated subjects with respect to **Figure 10A**, we observe a gradual decrease in gene expression for gene Cluster 2 and a gradual increase in expression for genes in Cluster 1 and 3 compared to the baseline (young vaccinated subjects under age 3) (SDF50 of online supplementary data files). Genes in Cluster 1 were not enriched in pathways while genes in Cluster 2 were enriched in Reactome pathways that include interferon and cytokine signaling, antiviral mechanism and response. Genes in Cluster 3 were enriched in Reactome pathways that include interferon and cytokine signaling, cellular senescence and immune system. When we compared vaccinated subjects to control subjects across ages we observed 3 main trends, with respect to **Figure 10B**: Cluster 1 (pathways include antiviral mechanisms, interferon and cytokine signaling) 2 (pathways such as immune response and cell migration, immunological synapse and chemokine receptors) and 4 (mitochondrial translation) genes are all up-regulated in vaccinated subjects with Cluster 3 (pathways include interferon and cytokine signaling and immune system) genes being down-regulated (**Figure 10B** and SDF51 of online supplementary data files).

We also compared influenza vaccinated subjects to influenza infected subjects to explore changes in expression with age. There are 3 major groups (cluster numbering with respect to Figure: Cluster 1 and 3 genes show a gradual decrease in expression with age in vaccinated subjects compared to influenza infected subjects. Genes in Cluster 2 show a gradual increase in expression with age in influenza vaccinated subjects. **Supplementary Figures 6–8** also explore temporal patterns with age.

Our heatmaps display temporal patterns with age in response to influenza infection and vaccination. These genes that are associated with disease and age interactions, are all involved in immune-related pathways. Exploring how gene expression changes with age in immune related genes can help further characterize the disease and improve treatments. For example, age, preexisting health conditions and influenza history (previous infection or vaccination) are all factors that can affect the efficacy of the vaccine (72). There is an on-going effort to improve efficacy of vaccination in the elderly population. Studies have suggested that antibody titers decline drastically in older adults from seroconversion to day 180 after vaccination (72). The decay of antibody titers also highlight the importance of determining the right time and how many times one should be vaccinated. Vaccines for the elderly population have been modified to increase the dosage and use adjuvants to increase immunogenicity (73). and Ramsay et al., also showed that vaccination during the current influenza season provides stronger protection than vaccinations from previous seasons (74).

The vaccine type also plays a role in immunogenicity within hosts. For example, Nakaya et al., were able to detect

larger antibody titers and plasmblasts generated in the trivalent inactivated vaccine (TIV) compared to the live attenuated vaccine (LAIV), and differentially expressed genes mostly related to interferon signaling (19). LAIV responses in young children are higher than in adults. For instance, LAIV when compared to inactivated vaccines induced smaller concentrations of antibodies in response to HA in adults (75). Previous findings have shown the benefit of taking a systems biology approach to assess gene expression responses to vaccinations (24, 72, 76). Our findings not only identify genes that are different between controls compared to infected and vaccinated subjects, but with our methodology we were also able to assess differences between the influenza infected and vaccinated subjects while still investigating disease genes that interact with age and sex. Our temporal patterns with age for each disease state helps to clarify how age might be playing a role.

Understanding how a vaccine affects the host is a critical step toward the design of more effective vaccines (3, 77). Our list of differentially expressed genes with respect to disease included activation of interferon pathways, NFκB, cell cycle pathways, as well as T and B cell receptor pathways that can inform on the adaptive response both in vaccine and disease. The observed differences in gene expression can be the consequence of both immune cell responses as well as changes in the composition of the immune cell sub populations present in the blood. Knowledge of what pathways are affected by vaccination can also help to identify targets for evaluating both the efficacy of the vaccine, and the variability in symptom severity, since the gene changes detected correspond to host immune responses to challenge (78).

In our study we were able to differentiate between genes activated in either influenza vaccination, disease or both. 334 genes changed in both influenza vaccine and disease response. Many of these genes are part of pathways involved in interferon responses and in cell cycle, and associated with host-pathogen interactions. Thus, these genes may have a role in symptom severity. However, we also identified 644 genes (involved in neutrophil degranulation and cell cycle processes) with changing expression only in influenza infection, but not in the vaccinated cohorts, thus providing evidence of differences in vaccine and disease responses. Genes which are not activated under vaccination but are differentially expressed under influenza may provide new vaccine targets that could improve efficacy, and may address variability in responses to disease post vaccination. Likewise, 644 genes were only activated in the vaccinated cohort, but not in influenza disease. Many of these genes (and the involved pathways) may also affect vaccine efficacy (including interferon signaling and antigen processing-cross presentation). While in this study we cannot infer whether the effects to efficacy are negative or positive, the different responses in vaccination should be monitored further, both for efficacy considerations, as well as for minimizing potential adverse effects.

Finally, our study identified genes that had a statistically significant interaction between disease state and age. The immune response pathways involving these genes should be considered when evaluating vaccine efficacy with respect to age, including not only for dosage, but also formulations (e.g., different gene targets by age or, conversely, formulations focused on targets with constant levels of expression, regardless of age that may provide a better baseline universal vaccine response).

Newer technologies such as RNA-sequencing (RNA-seq) can provide a more comprehensive view of gene expression as compared to gene expression microarrays (that have fixed gene targets). However, the number of available RNA-seq data matching our curation criteria was limited. Several recent investigations are now utilizing RNA-seq to investigate influenza infection and vaccination. For instance, Henn et al. identified 742 differentially expressed genes common in three of the five subjects they profiled, showing large variability of response across subjects 10 days following vaccination in B cell enriched cell expression following vaccination. Though the sample sizes are rather different for comparisons, this list overlaps 84 of our 978 vaccination results (79). Using purified individual cell types to assess post vaccination immune responses to trivalent inactivated influenza vaccine Hoek et al. identified 36 genes differentially expressed in PBMCs common in the subjects 7 days post vaccination (with 5 overlapping our findings)(80). New single-cell RNA-seq approaches may further probe such variability to deconvolve the cell-type-specific gene expression changes in influenza.

Several recent studies investigated a cohort (or subsets thereof) of 159 older adults post influenza A/H1N1 vaccination, including combined analysis of the various findings from the different analyses on this cohort (81–86): (i) Haralambieva et al. identified associations with influenza A/H1N1-specific memory B cell ELISPOT response after vaccination, where they estimated association of transcriptional changes with peak memory B cell ELISPOT response [Day 28 Day 0 with respect to vaccination (**Table S2** therein)], overlapping with 232 of our 978 vaccination biologically significant findings for blood (82). (ii) Ovsyannikova et al. evaluated hemagglutination inhibition (HAI) titer as vaccine response associations to gene expression levels post vaccination, and although they identified various genes related to HAI response in comparisons of responders/non-responders, these were at FDR> 0.9. (83). However, a "biology-to-gene" analysis by Ovsyannikova et al. did identify 13 gene-sets that may explain the odds of HAI response with models. (iii) Kennedy et al. evaluated immununosenescent signatures, identifying ROBO1 expression correlation with age, which in our results did not have a statistically significant age-disease interaction following Bonferroni correction ($p < 2 \times 10^{-5}$)(81). (iv) Zimmermann et al. followed up analyzing the same cohort, to identify how transcriptomic and other omics changes related to the dependence on immune cell subpopulations, and reported correlation of PBMC composition with overall variability in gene expression (84). (v) Voigt et al. (85) used a data driven approach to identify different clusters enriched for genes involved in specific immune cell types (85). (vi) Voigt et al. (86) also identified sex-specific signatures in B-cell ELISPOT responses, overlapping 10 genes in our findings ($p < 0.01$) (86). In such comparisons, we should note that the scope of these studies was different than the current investigation, which aimed to increase power substantially through pooling multiple studies with smaller sample sizes, and allowing the detection of multiple gene signatures following

strict multiple hypothesis correction. We anticipate that as more RNA-seq data are generated a more direct comparison will be possible.

Other RNA-seq work has included cellular culture *in vitro* applications: Cao et al. studied global transcriptome of H5N1 in A549 and 293 T cells (87). Tan et al. identified global transcriptome changes in human nasal epithelial cells (*in vitro* model with cells from seven donors) due to H3N2 influenza infection (88). Zhang et al. compared host mRNA and miRNA transcriptomes induced by influenza A H5N1 in human monocyte-derived macrophages post infection (89). In single-cell RNA-seq investigations, Russel et al. studied single-cell influenza transcriptomics for innate immunity using IFN reporter variants of the A549 human lung epithelial cell line (90). In a recent single cell sequencing investigation, Steuerman et al. profiled cells derived from lungs from *in vivo* influenza in C57BL/6J mice, to characterize host and viral transcriptomes simultaneously, and identify various immune cell types involved that had cell-specific transcriptional responses in influenza (91). While several RNA-seq studies have been carried out, sample sizes have been small, making comparisons with previous microarray work limited. The potential of single-cell work cannot be understated as it has the ability to differentiate cell-specific responses, instead of the aggregate blood/PBMC approaches. We anticipate that the rapid reduction in cost and further development of single-cell RNA-seq technology to lead to larger influenza disease and vaccination studies becoming available, that will help elucidate the variability in host gene expression observed in influenza.

As we have previously observed (29), meta-analyses using microarray expression data have multiple limitations: Our findings are limited to only genes that have been annotated and are existing probes on the arrays, and also have to be consistently utilized across array platforms. Hence, we are unable to probe global gene expression, and are limited to mRNA profiling. These can be expanded in future studies using RNA-seq data, and the newer single-cell sequencing approaches that would allow cell-specific information to be discerned, which is important in evaluating immune responses and the interplay between various cell types. Taking a similar approach to our microarray dataset analysis using RNA-seq data will promote the discovery of novel genes by being able to explore the entire transcriptome. Additionally, we are limited by the varying annotations of the available public datasets, and can only explore characteristics that are uniformly reported. For example, we did not have virus strain information for all samples or vaccine details so we were unable to include such info in our analysis. In addition to this, our study is unbalanced (particularly with respect to disease state,

where a limited number of influenza infection samples were available: 3,481 samples (1,277 controls, 297 influenza infection, 1,907 influenza vaccinated, 1,537 males and 1,944 female). We additionally used repeated measures, which we accounted for in our mixed effects model.

Despite the limitations introduced by using microarray data, our study identified gene candidates by factor (disease status, age, and sex) that can be examined further to understand their role in influenza infection and vaccination. We also highlighted 907 genes that have an age-effect on gene expression. These genes can be further explored to determine their role in influenza infection and how they can be further analyzed for their role in implementing effective universal vaccines regardless of age. All these considerations are of paramount importance in designing the next generation of vaccines, as we move forward toward a universal influenza vaccine.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed for this study can be found in Figshares online digital repository at https://doi.org/10.6084/m9.figshare.8636498.

AUTHOR CONTRIBUTIONS

LR, GC, and GM listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.02616/full#supplementary-material

Our supplemental figures and tables are provided in the accompanying supplementary data pdf file. All of our datasets/results from our meta-analysis pipeline have been uploaded to FigShare as online supplemental data, as described above (the corresponding online file names begin with a prefix "SDF" and are enumerated as also referred to in the manuscript).

REFERENCES

1. Taubenberger JK, Morens DM. The pathology of influenza virus infections. *Annu Rev Pathmechdis Mech Dis.* (2008) 3:499–522. doi: 10.1146/annurev.pathmechdis.3.121806.154316
2. Fukuyama S, Kawaoka Y. The pathogenesis of influenza virus infections: the contributions of virus and host factors. *Curr Opin Immunol.* (2011) 23:481–6. doi: 10.1016/j.coi.2011.07.016

3. Centers for Disease Control and Prevention. *Seasonal Influenza (Flu).* (2018). Available online at: https://www.cdc.gov/flu/about/index.html (accessed June 26, 2019).
4. Rivas H, Schmaling S, Gaglia M. Shutoff of host gene expression in influenza A virus and herpesviruses: similar mechanisms and common themes. *Viruses.* (2016) 8:102. doi: 10.3390/v8040102
5. Zhai Y, Franco LM, Atmar RL, Quarles JM, Arden N, Bucasas KL, et al. Host transcriptional response to influenza and other acute respiratory viral

infections–a prospective cohort study. *PLoS Pathog.* (2015) 11:e1004869. doi: 10.1371/journal.ppat.1004869

6. Krug RM. Functions of the influenza A virus NS1 protein in antiviral defense. *Curr Opin Virol.* (2015) 12:1–6. doi: 10.1016/j.coviro.2015.01.007

7. Lam HC, Bi X, Sreevatsan S, Boley D. Evolution and vaccination of influenza virus. *J Comput Biol.* (2017) 24:787–98. doi: 10.1089/cmb.2017.0025

8. Paessler S, Veljkovic V. Using electronic biology based platform to predict flu vaccine efficacy for 2018/2019. *F1000Research.* (2018) 7:298. doi: 10.12688/f1000research.14140.2

9. Grohskopf LA, Sokolow LZ, Broder KR, Walter EB, Fry AM, Jernigan DB. Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices United States, 2018–19 influenza season. *MMWR Recommendations Rep.* (2018) 67:1. doi: 10.15585/mmwr.rr6703a1

10. Cincy Tech. *Blue Water Vaccines, Developing Universal Flu Vaccine, Closes $7 Million Led by CincyTech.* (2019). Available online at: https://www.eurekalert.org/pub_releases/2019-07/c-bwv070519.php (accessed July 09, 2019).

11. Business Wire. *Vaxart Enters into Research Collaboration with Janssen to Evaluate Oral Universal Influenza Vaccine.* (2019). Available online at: https://www.businesswire.com/news/home/20190709005100/en/ (accessed July 09, 2019).

12. Lorenzo MMG, Fenton MJ. Immunobiology of influenza vaccines. *Chest.* (2013) 143:502–10. doi: 10.1378/chest.12-1711

13. Soema PC, Kompier R, Amorij JP, Kersten GF. Current and next generation influenza vaccines: formulation and production strategies. *Eur J Pharmaceut Biopharmaceut.* (2015) 94:251–63. doi: 10.1016/j.ejpb.2015.05.023

14. Mejias A, Dimo B, Suarez NM, Garcia C, Suarez-Arrabal MC, Jartti T, et al. Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Med.* (2013) 10:e1001549. doi: 10.1371/journal.pmed.1001549

15. Ioannidis I, McNally B, Willette M, Peeples ME, Chaussabel D, Durbin JE, et al. Plasticity and virus specificity of the airway epithelial cell immune response during respiratory virus infection. *J Virol.* (2012) 86:5422–36. doi: 10.1128/JVI.06757-11

16. Narang V, Lu Y, Tan C, Camous XFN, Nyunt SZ, Carre C, et al. Influenza vaccine-induced antibody responses are not impaired by frailty in the community-dwelling elderly with natural influenza exposure. *Front Immunol.* (2018) 9:2465. doi: 10.3389/fimmu.2018.02465

17. Dunning J, Blankley S, Hoang LT, Cox M, Graham CM, James PL, et al. Progression of whole-blood transcriptional signatures from interferon-induced to neutrophil-associated patterns in severe influenza. *Nat Immunol.* (2018) 19:625–35. doi: 10.1038/s41590-018-0111-5

18. Berdal JE, Mollnes TE, Wæhre T, Olstad OK, Halvorsen B, Ueland T, et al. Excessive innate immune response and mutant D222G/N in severe A (H1N1) pandemic influenza. *J Infect.* (2011) 63:308–16. doi: 10.1016/j.jinf.2011.07.004

19. Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, et al. Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol.* (2011) 12:786. doi: 10.1038/ni.2067

20. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell.* (2014) 157:499–513. doi: 10.1016/j.cell.2014.03.031

21. Obermoser G, Presnell S, Domico K, Xu H, Wang Y, Anguiano E, et al. Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. *Immunity.* (2013) 38:831–44. doi: 10.1016/j.immuni.2012.12.008

22. Tsuge M, Oka T, Yamashita N, Saito Y, Fujii Y, Nagaoka Y, et al. Gene expression analysis in children with complex seizures due to influenza A (H1N1) pdm09 or rotavirus gastroenteritis. *J Neurovirol.* (2014) 20:73–84. doi: 10.1007/s13365-013-0231-5

23. Cao RG, Suarez NM, Obermoser G, Lopez SM, Flano E, Mertz SE, et al. Differences in antibody responses between trivalent inactivated influenza vaccine and live attenuated influenza vaccine correlate with the kinetics and magnitude of interferon signaling in children. *J Infect Dis.* (2014) 210:224–33. doi: 10.1093/infdis/jiu079

24. Nakaya HI, Hagan T, Duraisingham SS, Lee EK, Kwissa M, Rouphael N, et al. Systems analysis of immunity to influenza vaccination across multiple years

and in diverse populations reveals shared molecular signatures. *Immunity.* (2015) 43:1186–98. doi: 10.1016/j.immuni.2015.11.012

25. Goswami M, Prince G, Biancotto A, Moir S, Kardava L, Santich BH, et al. Impaired B cell immunity in acute myeloid leukemia patients after chemotherapy. *J Transl Med.* (2017) 15:155. doi: 10.1186/s12967-017-1252-2

26. Furman D, Jojic V, Kidd B, Shen-Orr S, Price J, Jarrell J, et al. Apoptosis and other immune biomarkers predict influenza vaccine responsiveness. *Mol Syst Biol.* (2013) 9:659. doi: 10.1038/msb.2013.15

27. Thakar J, Mohanty S, West AP, Joshi SR, Ueda I, Wilson J, et al. Aging-dependent alterations in gene expression and a mitochondrial signature of responsiveness to human influenza vaccination. *Aging.* (2015) 7:38. doi: 10.18632/aging.100720

28. Franco LM, Bucasas KL, Wells JM, Niño D, Wang X, Zapata GE, et al. Integrative genomic analysis of the human immune response to influenza vaccination. *Elife.* (2013) 2:e00299. doi: 10.7554/eLife.00299

29. Brooks LR, Mias GI. Data-driven analysis of age, sex, and tissue effects on gene expression variability in Alzheimer's disease. *Front Neurosci.* (2019) 13:392. doi: 10.3389/fnins.2019.00392

30. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna (2018). Available online at: https://www.R-project.org/ (accessed June 30, 2016).

31. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* (2002) 30:207–10. doi: 10.1093/nar/30.1.207

32. Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, et al. ArrayExpressa public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* (2003) 31:68–71. doi: 10.1093/nar/gkg091

33. Brusic V, Gottardo R, Kleinstein SH, Davis MM, Hafler DA, Quill H, et al. Computational resources for high-dimensional immune analysis from the Human Immunology Project Consortium. *Nat Biotechnol.* (2014) 32:146. doi: 10.1038/nbt.2777

34. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data.* (2018) 5:180015. doi: 10.1038/sdata.2018.15

35. Gautier L, Cope L, Bolstad BM, Irizarry RA. affyanalysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* (2004) 20:307–15. doi: 10.1093/bioinformatics/btg405

36. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007

37. Carlson M PH. *AnnotationForge: Tools for Building SQLite-Based Annotation Data Packages.* (2019). Available online at: https://bioconductor.org/packages/release/bioc/html/AnnotationForge.html/ (accessed March 30, 2019).

38. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics.* (2010) 26:2363–7. doi: 10.1093/bioinformatics/btq431

39. MacDonald JW. *affycoretools: Functions Useful for Those Doing Repetitive Analyses with Affymetrix GeneChips.* (2018). Available online at: https://www.bioconductor.org/packages/release/bioc/html/affycoretools.html/ (accessed March 30, 2019).

40. Sakia R. The Box-Cox transformation technique: a review. *Statistician.* (1992) 41:169–78. doi: 10.2307/2348250

41. Mias G. Chapter 4: Databases: E-Utilities and UCSC Genome Browser. In: *Mathematica for Bioinformatics: A Wolfram Language Approach to Omics.* Cham: Springer International Publishing (2018). p. 133–170.

42. Mias GI, Yusufaly T, Roushangar R, Brooks LR, Singh VV, Christou C. MathIOmica: an integrative platform for dynamic omics. *Sci Rep.* (2016) 6:37237. doi: 10.1038/srep37237

43. Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, et al. *Package lme4.* (2019). Available online at: https://cran.r-project.org/web/packages/lme4/lme4.pdf (accessed June 09, 2019).

44. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst.* (2016) 12:477–9. doi: 10.1039/C5MB00663E

45. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* (2003) 13:2498–504. doi: 10.1101/gr.1239303

46. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics.* (2005) 21:3448–9. doi: 10.1093/bioinformatics/bti551

47. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27

48. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* (2016) 44:D457–62. doi: 10.1093/nar/gkv1070

49. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* (2017) 45:D353–61. doi: 10.1093/nar/gkw1092

50. Brooks LR, Mias GI. *Streptococcus pneumoniae*'s virulence and host immunity: aging, diagnostics and prevention. *Front Immunol.* (2018) 9:1366. doi: 10.3389/fimmu.2018.01366

51. Paules CI, Sullivan SG, Subbarao K, Fauci AS. Chasing seasonal influenza The need for a universal influenza vaccine. *N Engl J Med.* (2018) 378:7–9. doi: 10.1056/NEJMp1714916

52. Wang JX, Nigrovic P. CD177 participates in a novel mechanism for regulating neutrophil recruitment (P3093). *J Immunol.* (2013) 190(1 Suppl):43.9. Available online at: https://www.jimmunol.org/content/190/1_Supplement/43.9 (accessed October 25, 2019).

53. 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) 54:1.30.1–33. doi: 10.1002/cpbi.5

54. Yu Y, Wang C, Clare S, Wang J, Lee SC, Brandt C, et al. The transcription factor Bcl11b is specifically expressed in group 2 innate lymphoid cells and is essential for their development. *J Exp Med.* (2015) 212:865–74. doi: 10.1084/jem.20142318

55. Moisy D, Avilov SV, Jacob Y, Laoide BM, Ge X, Baudin F, et al. HMGB1 protein binds to influenza virus nucleoprotein and promotes viral replication. *J Virol.* (2012) 86:9122–33. doi: 10.1128/JVI.00789-12

56. Ludwig S, Planz O. Influenza viruses and the NF-κB signaling pathway–towards a novel concept of antiviral therapy. *Biol Chem.* (2008) 389:1307–12. doi: 10.1515/BC.2008.148

57. Kumar N, Xin Zt, Liang Y, Ly H, Liang Y. NF-κB signaling differentially regulates influenza virus RNA synthesis. *J Virol.* (2008) 82:9880–9. doi: 10.1128/JVI.00909-08

58. Hashem AM, Gravel C, Chen Z, Yi Y, Tocchi M, Jaentschke B, et al. CD40 ligand preferentially modulates immune response and enhances protection against influenza virus. *J Immunol.* (2014) 193:722–34. doi: 10.4049/jimmunol.1300093

59. Hatziioti C, Heath AW. CD40-mediated enhancement of immune responses against three forms of influenza vaccine. *Immunology.* (2007) 122:98–106. doi: 10.1111/j.1365-2567.2007.02617.x

60. Westera L, Jennings AM, Maamary J, Schwemmle M, García-Sastre A, Bortz E. Poly-ADP ribosyl polymerase 1 (PARP1) regulates influenza A virus polymerase. *Adv Virol.* (2019) 2019:8512363. doi: 10.1155/2019/8512363

61. Loo YM, Gale M Jr. Immune signaling by RIG-I-like receptors. *Immunity.* (2011) 34:680–92. doi: 10.1016/j.immuni.2011.05.003

62. Chiaretti A, Pulitanò S, Barone G, Ferrara P, Romano V, Capozzi D, et al. IL-1β and IL-6 upregulation in children with H1N1 influenza virus infection. *Mediat Inflamm.* (2013) 2013:495848. doi: 10.1155/2013/495848

63. Jefferies CA. Regulating IRFs in IFN driven disease. *Front Immunol.* (2019) 10:325. doi: 10.3389/fimmu.2019.00325

64. Wilk E, Pandey AK, Leist SR, Hatesuer B, Preusse M, Pommerenke C, et al. RNAseq expression analysis of resistant and susceptible mice after influenza A virus infection identifies novel genes associated with virus replication and important for host resistance to infection. *BMC Genomics.* (2015) 16:655. doi: 10.1186/s12864-015-1867-8

65. Hamamoto I, Takaku H, Tashiro M, Yamamoto N. High yield production of influenza virus in Madin Darby canine kidney (MDCK) cells with stable knockdown of IRF7. *PLoS ONE.* (2013) 8:e59892. doi: 10.1371/journal.pone.0059892

66. Ciancanelli MJ, Huang SX, Luthra P, Garner H, Itan Y, Volpi S, et al. Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. *Science.* (2015) 348:448–53. doi: 10.1126/science.aaa1578

67. Zhou Z, Xue Q, Wan Y, Yang Y, Wang J, Hung T. Lysosome-associated membrane glycoprotein 3 is involved in influenza A virus replication in human lung epithelial (A549) cells. *Virol J.* (2011) 8:384. doi: 10.1186/1743-422X-8-384

68. Shin D, Lee J, Park JH, Min JY. Double plant homeodomain fingers 2 (DPF2) promotes the immune escape of influenza virus by suppressing beta interferon production. *J Virol.* (2017) 91:e02260–16. doi: 10.1128/JVI.02260-16

69. Wen F, Guo J, Li Z, Huang S. Sex-specific patterns of gene expression following influenza vaccination. *Sci Rep.* (2018) 8:13517. doi: 10.1038/s41598-018-31999-x

70. Avitsur R, Mays JW, Sheridan JF. Sex differences in the response to influenza virus infection: modulation by stress. *Hormones Behav.* (2011) 59:257–64. doi: 10.1016/j.yhbeh.2010.12.002

71. Bahadoran A, Lee SH, Wang SM, Manikam R, Rajarajeswaran J, Raju CS, et al. Immune responses to influenza virus and its correlation to age and inherited factors. *Front Microbiol.* (2016) 7:1841. doi: 10.3389/fmicb.2016.01841

72. Young B, Sadarangani S, Jiang L, Wilder-Smith A, Chen MIC. Duration of influenza vaccine effectiveness: a systematic review, meta-analysis, and meta-regression of test-negative design case-control studies. *J Infect Dis.* (2017) 217:731–41. doi: 10.1093/infdis/jix632

73. Castrucci MR. Factors affecting immune responses to the influenza vaccine. *Human Vaccines Immunotherapeut.* (2018) 14:637–46. doi: 10.1080/21645515.2017.1338547

74. Ramsay LC, Buchan SA, Stirling RG, Cowling BJ, Feng S, Kwong JC, et al. The impact of repeated vaccination on influenza vaccine effectiveness: a systematic review and meta-analysis. *BMC Med.* (2019) 17:9. doi: 10.1186/s12916-018-1239-8

75. Hoft DF, Lottenbach KR, Blazevic A, Turan A, Blevins TP, Pacatte TP, et al. Comparisons of the humoral and cellular immune responses induced by live attenuated influenza vaccine and inactivated influenza vaccine in adults. *Clin Vaccine Immunol.* (2017) 24:e00414–16. doi: 10.1128/CVI.00414-16

76. Bucasas KL, Franco LM, Shaw CA, Bray MS, Wells JM, Niño D, et al. Early patterns of gene expression correlate with the humoral immune response to influenza vaccination in humans. *J Infect Dis.* (2011) 203:921–9. doi: 10.1093/infdis/jiq156

77. Haaheim LR, Madhun AS, Cox R. Pandemic influenza vaccines–The challenges. *Viruses.* (2009) 1:1089–109. doi: 10.3390/v1031089

78. Barton AJ, Hill J, Pollard AJ, Blohmke CJ. Transcriptomics in human challenge models. *Front Immunol.* (2017) 8:1839. doi: 10.3389/fimmu.2017.01839

79. Henn AD, Wu S, Qiu X, Ruda M, Stover M, Yang H, et al. High-resolution temporal response patterns to influenza vaccine reveal a distinct human plasma cell gene signature. *Sci Rep.* (2013) 3:2327. doi: 10.1038/srep02327

80. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. *PLoS ONE.* (2015) 10:e0118528. doi: 10.1371/journal.pone.0118528

81. Kennedy RB, Ovsyannikova IG, Haralambieva IH, Oberg AL, Zimmermann MT, Grill DE, et al. Immunosenescence-related transcriptomic and immunologic changes in older individuals following influenza vaccination. *Front Immunol.* (2016) 7:450. doi: 10.3389/fimmu.2016.00450

82. Haralambieva IH, Ovsyannikova IG, Kennedy RB, Zimmermann MT, Grill DE, Oberg AL, et al. Transcriptional signatures of influenza A/H1N1-specific IgG memory-like B cell response in older individuals. *Vaccine.* (2016) 34:3993–4002. doi: 10.1016/j.vaccine.2016.06.034

83. Ovsyannikova IG, Oberg AL, Kennedy RB, Zimmermann MT, Haralambieva IH, Goergen KM, et al. Gene signatures related to HAI response following influenza A/H1N1 vaccine in older individuals. *Heliyon.* (2016) 2:e00098. doi: 10.1016/j.heliyon.2016.e00098

84. Zimmermann MT, Kennedy RB, Grill DE, Oberg AL, Goergen KM, Ovsyannikova IG, et al. Integration of immune cell populations, mRNA-seq, and CpG methylation to better predict humoral immunity to influenza vaccination: dependence of mRNA-seq/CpG methylation on immune cell populations. *Front Immunol.* (2017) 8:445. doi: 10.3389/fimmu.2017.00445

85. Voigt EA, Grill DE, Zimmermann MT, Simon WL, Ovsyannikova IG, Kennedy RB, et al. Transcriptomic signatures of cellular and humoral immune responses in older adults after seasonal influenza vaccination identified by data-driven clustering. *Sci Rep.* (2018) 8:739. doi: 10.1038/s41598-017-17735-x

86. Voigt EA, Ovsyannikova IG, Kennedy RB, Grill DE, Goergen KM, Schaid DJ, et al. Sex differences in older adults' immune responses to seasonal influenza vaccination. *Front Immunol.* (2019) 10:180. doi: 10.3389/fimmu.2019.00180

87. Cao Y, Zhang K, Liu L, Li W, Zhu B, Zhang S, et al. Global transcriptome analysis of H5N1 influenza virus-infected human cells. *Hereditas.* (2019) 156:10. doi: 10.1186/s41065-019-0085-9

88. Tan KS, Andiappan AK, Lee B, Yan Y, Liu J, Tang SA, et al. RNA sequencing of H3N2 influenza virus-infected human nasal epithelial cells from multiple subjects reveals molecular pathways associated with tissue injury and complications. *Cells.* (2019) 8:986. doi: 10.3390/cells8090986

89. Zhang N, Bao YJ, Tong AH, Zuyderduyn S, Bader GD, Malik Peiris JS, et al. Whole transcriptome analysis reveals differential gene expression profile reflecting macrophage polarization in response to influenza A H5N1 virus infection. *BMC Med Genomics.* (2018) 11:20. doi: 10.1186/s12920-018-0335-0

90. Russell AB, Elshina E, Kowalsky JR, Te Velthuis AJW, Bloom JD. Single-cell virus sequencing of influenza infections that trigger innate immunity. *J Virol.* (2019) 93:e00500-19. doi: 10.1128/JVI.00500-19

91. Steuerman Y, Cohen M, Peshes-Yaloz N, Valadarsky L, Cohn O, David E, et al. Dissection of influenza infection *in vivo* by single-cell RNA sequencing. *Cell Syst.* (2018) 6:679–91.e4. doi: 10.1016/j.cels.2018.05.008

Conflict of Interest: GM has consulted for Colgate-Palmolive.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.