Integrated Drivers of Basal and Acute Immunity in Diverse Human Populations

Aisha Souquette[1,2], E. Kaitlynn Allen[1], Christine M. Oshansky[1], Li Tang[3], Sook-san Wong[4], Trushar Jeevan[4], Lei Shi[3], Stanley Pounds[3], George Elias[5,6,7,8], Guillermina Kuan[9,10], Angel Balmaseda[10,11], Raul Zapata[10,11], Kathryn Shaw-Saliba[12], Pierre Van Damme[13], Viggo Van Tendeloo[5], Juan Carlos Dib[14], Benson Ogunjimi[6,7,8,15], Richard Webby[4], Stacey Schultz-Cherry[4], Andrew Pekosz[12,16], Richard Rothman[12], Aubree Gordon[17], Paul G. Thomas[1]*

[1] Department of Immunology, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA; [2] University of Tennessee Health Science Center, Memphis, Tennessee, 38163, USA; [3] Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA; [4] Department of Infectious Disease, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA; [5] Laboratory of Experimental Hematology (LEH), Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, 2610, Belgium; [6] Center for Health Economics Research and Modeling Infectious Diseases (CHERMID), Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, 2610, Belgium; [7] Antwerp Center for Translational Immunology and Virology (ACTIV), Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, 2610, Belgium; [8] Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing (AUDACIS), University of Antwerp, Antwerp, 2610, Belgium; [9] Centro de Salud Sócrates Flores Vivas, Ministry of Health, Managua, 12014, Nicaragua; [10] Centro Nacional de Diagnóstico y Referencia, Ministry of Health, Managua, 16064, Nicaragua; [11] Sustainable Sciences Institute (Site - Managua, Nicaragua), San Francisco, California, 94102, USA; [12] Department of Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, 21209, USA; [13] Center for the Evaluation of Vaccination (CEV), Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, 2610, Belgium; [14] Department of Medicine, Universidad del Norte, Barranquilla, 081001, Colombia; [15] Department of Paediatrics, Antwerp University Hospital, Antwerp, 2610, Belgium; [16] W. Harry Feinstone Department of Molecular Microbiology and Immunology, The Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA; [17] Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, 48109, USA

*Corresponding author: paul.thomas@stjude.org

Keywords: immune variation, influenza, herpesvirus, co-infection, systems immunology, basal immunity, acute immunity, immunogenetics, immunophenotype, correlate of severity

Abstract

Prior studies have identified genetic, infectious, and biological associations with immune competence and disease severity; however, there have been few integrative analyses of these factors and study populations are often limited in demographic diversity. Utilizing samples from 1,705 individuals in 5 countries, we examined putative determinants of immunity, including: single nucleotide polymorphisms, ancestry informative markers, herpesvirus status, age, and sex. In healthy subjects, we found significant differences in cytokine levels, leukocyte phenotypes, and gene expression. Transcriptional responses also varied by cohort, and the most significant determinant was ancestry. In influenza infected subjects, we found two disease severity immunophenotypes, largely driven by age. Additionally, cytokine regression models show each determinant differentially contributes to acute immune variation, with unique and interactive, location-specific herpesvirus effects. These results provide novel insight into the scope of immune heterogeneity across diverse populations, the integrative effects of factors which drive it, and the consequences for illness outcomes.

Introduction

Functional consequences of immune heterogeneity are demonstrated by variation in vaccine efficacy, responses to therapeutic treatment, infectious disease severity, and susceptibility to autoimmunity. Understanding the landscape and drivers of immune variation would provide valuable insight into immune competence and points of intervention to improve disease outcome. Indeed, baseline immune predictors of vaccine responses and neoplastic disease prognosis have already been identified in a variety of studies.[1,2]

Immune variation, at baseline and in response to immune challenge, is dynamically influenced by unique and interactive effects of intrinsic and extrinsic host factors. Intrinsic factors are internal to the host and include variables such as age, sex, genetics, and microbiome. Extrinsic factors are external to the host and include variables such as diet, environment, and acquired infections. Each of these determinants can influence multiple facets of immunity, including, but not limited to, susceptibility to infection or immune mediated disease, the basal activation state of the immune system, pathogen detection, downstream signaling, priming of adaptive immunity, leukocyte phenotype and function, and/or response to therapeutic treatment.[3–22] For example, in a study of three independent influenza cohorts, we found that a single nucleotide polymorphism rs3448144-A in *IFITM3* is enriched in severe patients, governs transcription factor binding and promoter activity, and alters methylation of the *IFITM3* promoter in CD8 T cells, which correlates with cell frequency in nasal washes.[12] Coinciding with these results, recent studies have reported associations between rs3448144 and COVID-19 disease severity.[23,24] Importantly, the collective impact of immune modulation induced by intrinsic and extrinsic host factors can alter vaccination responses and illness outcome. However, further study is needed to clarify the quantitative effects of these variables, as conflicting conclusions have occurred across studies, possibly due to skews in age distribution, geographical location, or ancestry impacting observed results. Ultimately, this limits both our understanding of the extent of immune heterogeneity and the general applicability of conclusions drawn from single cohorts.

Here, we quantify the contribution of intrinsic and extrinsic host factors on the setpoint of baseline immunity, the transcriptional response to *in vitro* immune challenge, and the magnitude and quality of the immune response to *in vivo* influenza infection in diverse populations around the world. Immune measures include cytokines (from plasma and nasal wash), cellular phenotypes, and/or gene expression (both at baseline and post PBMC stimulation with PRR ligands). Of these, cytokines were our primary and most extensively used measurement, with data collected at the RNA and protein level. While most prior work has focused on cell phenotypes (by flow cytometry), cytokines have several advantages from both a scientific and clinical standpoint. Samples are easily obtained and stored, and analytes can quickly and precisely be measured with minimal equipment, sample input, and processing. In addition, they have a substantial dynamic range and offer broad profiling of innate and adaptive immunity. For influenza infected subjects, outcome measures include presence/absence of symptoms, symptom severity scores, viral load, and duration of influenza virus shedding. To explain variation in these measures, linear regression modeling was utilized, which allows for an easily interpretable and transparent integrative analysis of all independent variables and a quantification of their unique, interactive, and collective effects on the immune measure of interest. Collectively, the data presented here, representing 123,178,117 data points from 1,705 individuals, show age, sex, herpesviruses, and host genetics can independently and interactively influence baseline immunity and the acute immune response to influenza virus infection. Of these determinants, ancestry is the most significant factor. Importantly, these factors differentially contribute to variation in a given immune measure, and the cumulative impact of immune modulation by these factors results in unique immune profiles associated with disease severity. Our results demonstrate that simultaneous assessment of diverse populations is necessary to better understand the scope of immune variation and the factors that contribute to it.

Results

Explanatory factor variation within and across diverse populations

This study includes samples from 1,705 healthy or influenza infected subjects from 5 countries and 8 distinct populations (**Fig 1**). North American populations include the Memphis cohort and the Baltimore cohort. Latin American populations include the Nicaragua cohort and the Colombia cohorts, which include subjects from the local city of Calabazo and two indigenous populations, Seywiaka and Umandita. Samples also include a European cohort from Belgium and an Asian cohort from Taiwan. Select assays and corresponding analyses focused on our two most diverse cohorts, Memphis and Nicaragua, which have the largest sample size, availability, and variety of sample types collected. These populations also capture environmental diversity, and, collectively, represent three distinct racial/ethnic groups (White/Caucasian, Hispanic/Latino, Black/African American), two of which are highly admixed, thereby representing greater genetic diversity. Enrichment of diverse, understudied racial/ethnic groups is a key feature of our study, as recent work has shown that increasing genetic diversity, as opposed to increasing the sample size of European ancestry groups, provides substantial

121 improvement to genetic analyses, such as fine-mapping functional variants and portability of
122 polygenic risk scores.[25,26]
123
124 Before examining baseline and acute immunity, we assessed the degree of variation in
125 explanatory variables across cohorts. Here, intrinsic host factors include biological and genetic
126 variables. Biological factors include age and sex, and the distribution of each varied by cohort
127 (**Fig 2A-B**). Overall, the study was 58% female and age ranged from <1 year up to 92 years of
128 age.
129
130 Heritable contribution to immune variation was determined for a subset of healthy and infected
131 subjects with available DNA or PBMCs (n = 355, 234 Memphians and 121 Nicaraguans), using
132 genome-wide ancestry informative markers and candidate functional, immune-related SNPs.
133 Genetic based ancestry, as determined by ancestry informative markers (AIMs), accounts for
134 inherent genetic differences across diverse populations, typically from distinct geographical
135 regions. This allows for a genetic and quantitative representation of admixed populations, as
136 opposed to discrete classifications of race, which do not adequately represent populations with
137 mixed ancestral backgrounds. AIMs were analyzed utilizing a principal component analysis
138 (PCA), and the top two principal component values were included in downstream regression
139 analyses. Graphical representation of the PCA results shows distinct clusters that were similar
140 to the self-reported racial and ethnic makeup of the Memphis (75% African/Black American and
141 25% Caucasian/White American) and Nicaragua (100% Latino) cohorts, with a subset of
142 ancestrally mixed individuals falling between clusters (**Fig 2C**). Importantly, only AIMs were able
143 to distinguish between ancestral population clusters (**Supp Fig 1A-B**) and account for genetic
144 variation within a given population (**Supp Fig 1C**). This highlights the limited applicability of
145 current “Race” and “Ethnicity” categories collected in the clinical setting and applied in
146 biomedical research, and supports current arguments for the need for inclusion of more
147 appropriate factors in human studies aiming to distinguish between biological, epidemiological,
148 and social determinants of health and disease.[27,28] For the candidate SNP list, we examined
149 common SNPs in functional regions of genes encoding the 41 cytokines in our assay and in 3
150 transcription factors important in initiating antiviral immunity, namely IRF3, IRF7, and all 3
151 subunits of NFkB (detailed description of prioritization is in the Methods). While some SNPs
152 have similar allele frequencies across the Memphis and Nicaragua cohorts, several, such as
153 rs13900 in MCP1 and rs3024496 in IL10, show significant variation and opposing major alleles
154 in the two populations (**Fig 2D**, **Supp Table 1**), indicating genetic differences in immune-related
155 genes with potential functional consequences due to ancestry.
156
157 Herpesviruses were utilized as an extrinsic factor of infectious exposure because they are
158 globally prevalent, cause chronic infections that are generally well tolerated by
159 immunocompetent hosts, and have previously been associated with modulation of the immune
160 system.[29–32] The serostatus of cytomegalovirus (CMV), Epstein-Barr virus (EBV), herpes
161 simplex virus (HSV, 1 and 2 combined), human herpesvirus 6 (HHV6), and varicella-zoster virus
162 (VZV) were determined for healthy and influenza infected subjects with available plasma from
163 all eight cohorts (n = 1,532). Herpesvirus seroprevalence ranged from as low as 47.5% up to
164 97.5% (**Fig 2E**). Although the median age varies across cohorts, comparisons can be made

between cohorts of similar age ranges and/or with age adjustment and these results demonstrate how herpesvirus prevalence varies across distinct groups. For example, although the Baltimore, Taiwan, and Belgium cohorts are similar in age, Belgium has 31-35% lower prevalence of HSV but a 9-24% higher prevalence of EBV. Additionally, despite having a similar age group, the Colombian cohort has 22% higher prevalence of CMV, 31% higher prevalence of EBV, and a 32% lower prevalence of HHV6 compared to Nicaragua. Importantly, with the exception of the Belgium cohort, which has the most limited age range in participants, each cohort exhibits a positive correlation between age and the total number of herpesviruses an individual has (**Supp Fig 1D**). This is particularly important to account for in downstream analyses because influenza disease severity is age associated, very young individuals (infants and children under 5) tend to have more severe infection and tend to be herpesvirus seronegative. Highlighting the ubiquity of herpesviruses, of the 1,532 individuals in this analysis, only 10 (0.587%) were negative for CMV, HSV, EBV, and HHV6, and only 118 (6.92%) tested positive for a single herpesvirus (**Fig 2F**). These data underscore the importance of accounting for multiple herpesvirus infections when trying to assess their contributions to a given disease outcome, as most individuals have more than one. However, due to the high prevalence of HHV6 and VZV, the herpesvirus serostatuses in downstream analyses are focused on CMV, EBV, and HSV.

Distinct basal immune profiles mold response to immune challenge

To begin assessing baseline immune status, the levels of 38 plasma/serum cytokines were measured in 631 healthy subjects.[33–38] Cytokine levels were natural log (ln) transformed and visualized using a t-Distributed Stochastic Neighbor Embedding (t-SNE) plot which aids in dimensionality reduction while utilizing data from all cytokines measured. These results show marked variation in cytokine levels, with distinct clustering by cohort. Some populations exhibit tight clusters (Baltimore), whereas others exhibit dispersed clusters (Colombia) or even two distinct clusters (Memphis), suggesting the variance of cytokine profiles across individuals varies by population and, consequently, focusing on a population with minimal variation across subjects may underestimate the extent of immune variation and bias results (**Fig 3A**). Split clustering of the Memphis cohort was not due to age, sex, herpesvirus serostatus, race, season/year, assay plate, or time between assay runs (**Supp Fig 2A-G**). Moreover, in a subsequent analysis, cohort clusters and their positions relative to each other remain after removal of Memphis cohort data (**Supp Fig 2H**). Variation was also observed at the individual cytokine level for immunomodulatory cytokines, pro-inflammatory cytokines, chemokines, and growth factors (**Fig 3B**).

To examine the effects of intrinsic and extrinsic host factors on basal cytokine levels, a linear regression model (Model #1, **Fig 1**) was used to take into account the effects of: age, sex, cohort (accounting for genetics and unknown environmental factors specific to the geographically distinct populations), each herpesvirus alone, any interaction between a herpesvirus and age (due to the age association with serostatus), and any interaction between herpesviruses (since most individuals are infected with more than one). These results show herpesviruses have largely unique effects on basal cytokine levels, and the herpesvirus which caused the most effects was CMV (**Fig 3C** and **Supp Fig 3**). Each additional standard deviation

208 unit increase in age (years) is cumulatively associated with altered GRO, IFNg, Eotaxin, and
209 IL12p70, thus the larger the difference in age, the greater the magnitude of age effects (**Supp**
210 **Fig 3**). Male sex was associated with decreased TGFa and cohort differentially affected all
211 cytokines (**Supp Fig 3**). The adjusted R^2 result from a linear regression model is an estimate of
212 how much variability in the dependent variable the model can account for, and the "adjustment"
213 refers to a penalty to the R^2 value due to the incorporation of additional variables. This value
214 was highest for the IL4 model at 0.4709 (47.1%) (**Supp Fig 3**). Another output from the
215 regression models is an estimate of a coefficient for each factor which represents the change in
216 the dependent variable per unit change in the factor. Comparing coefficients across factors
217 provides insight into the magnitude of the effect of that factor. In these healthy cytokine models,
218 cohort was associated with the strongest effects on circulating cytokine levels (**Supp Fig 3**).
219 Collectively, these data suggest distinct populations have unique cytokine profiles, the effect of
220 which likely represents divergent baseline immune states, significant impacts on downstream
221 effector functions, including immune cell differentiation and functional responses to stimuli.
222
223 To further explore the potential for distinct immunophenotypes, we focused on two cohorts with
224 PBMC samples, Memphis and Nicaragua, which allowed for more comprehensive investigation
225 and statistical analysis. Flow cytometry was utilized to characterize the activation and/or
226 differentiation state of circulating lymphocytes and antigen presenting cells at baseline. Samples
227 included 28 Memphians (average age = 22, median age = 18) and 19 Nicaraguans (average
228 age = 20, median age = 15). Nicaraguans exhibited a trend towards higher regulatory T cell
229 (Treg)-like CD25+ CD4 T cells (padj = 0.071, **Supp Table 2**). To validate these findings, the
230 Treg transcription factor FOXP3 (forkhead box protein P3) was included in a follow up
231 experiment and results confirmed Nicaraguans have an increased frequency of FOXP3+ CD4
232 Tregs (p = 0.028, **Fig 3D**). Nicaraguans also showed significantly lower frequency and number
233 of activated CD4 and CD8 T cells (padj = 5.5e-4 for both subsets), but higher plasmablasts
234 (padj = 3.04e-3) and CD8 effector memory T cells (padj = 0.014), the latter being consistent with
235 increased prevalence of chronic infections (**Fig 3D**). Moreover, Memphians showed a higher
236 frequency and number of pro-inflammatory plasmacytoid dendritic cells (pDCs) (padj = 2.37e-3)
237 (**Fig 3D**).
238
239 Substantial variation was also observed between the two cohorts at the RNA level after PBMC
240 stimulation with plain media (an unstimulated control), lipopolysaccharide (a bacterial TLR4
241 ligand), Pam3CSK4 (a bacterial TLR1/2 ligand), transfected poly(I:C) (representative of RNA
242 virus infection), or transfected poly(dA:dT) (representative of DNA virus infection). For each
243 stimulation condition, samples included 94 Memphians and 94 Nicaraguans. Examples of genes
244 which varied by cohort include, but are not limited to, *IFI16* (a PRR important for herpesvirus
245 detection), *IRF1* and *IRF9* (upstream immune response transcription factors), *TBK1* (a
246 downstream immune signaling protein), *CCL4* and *CXCL2* (chemokines), *CD86* (an antigen
247 presenting cell activation marker and ligand for a co-stimulatory molecule expressed on T cells),
248 *KLRC1* (inhibitory receptor NKG2A, expressed on NK and T cells), and *IFNG* (an antiviral
249 cytokine) (**Fig 4A, Supp Table 3**). These genes represent each step of the immune response
250 and highlight the numerous ways and potential for unique immune profiles to arise in diverse
251 populations. Notably, differential gene expression was stimulation condition dependent,

252 highlighting the need for future studies to investigate immune variation under a variety of
253 immune pressures.
254
255 Coinciding with the basal cytokine results (**Fig 3A**), PCA analysis of the media (unstimulated
256 control) gene expression data shows substantial variation in the basal immune transcriptional
257 profile (**Fig 4B**). To determine which factors were driving this variation, linear regression
258 modeling was utilized to account for age, sex, herpesvirus status, and ancestry (Model #2, **Fig**
259 **1**). These results show basal immune transcriptional heterogeneity is largely driven by genetic
260 ancestry (affected 62.2% of 90 genes), followed by herpesvirus infection (22.2% of genes), and
261 demographic variables, age and sex (7.8% of genes) (**Fig 4C, Supp Fig 4A, Supp Table 3**). A
262 similar model was utilized to assess determinants of stimulated gene expression, and
263 accounted for stimulant, age, sex, herpesvirus status, ancestry, and the basal immune profile
264 (Model #3, **Fig 1**). Similar to the basal response, ancestry was the most prominent determinant,
265 contributing to statistically significant alterations in 83% of genes examined, the basal
266 transcriptional profile affected 74% of genes (on top of the effects already captured by
267 ancestry), and herpesvirus status affected 66%, while demographic variables (age, sex)
268 affected 31% (**Fig 4C, Supp Fig 4B, Supp Table 3**).
269
270 Exploring these transcriptional results further, expression quantitative trait loci (eQTL) were
271 determined to define SNPs that were associated with altered gene expression and demonstrate
272 a means by which genetic effects could contribute to the observed transcriptional heterogeneity.
273 These data were analyzed with linear regression modeling in order to account for: genotypes for
274 immune-related SNPs of interest, genetic based ancestry, age, sex, herpesvirus status, and the
275 stimulation condition (Model #4, **Fig 1**). Of the 94 prioritized candidate SNPs, 36 SNPs in 26
276 genes (24 cytokines and 2 transcription factors) were significant factors in the expression of
277 their respective genes (**Supp Fig 4C, Supp Table 4**). Notably, 13 of these genes also showed a
278 measurable heritable influence in Brodin et al. at the protein level (GCSF, IFNg, IL10, IL12p70,
279 IL13, IL15, IL6, TNFb, GROa, IL8, MCP1, TGFa, and VEGF).[7] While further study is required to
280 determine the mechanism behind the eQTLs, these data demonstrate how genetic variation can
281 significantly affect gene expression both at baseline and during immune responses. With
282 respect to immune variation, this would be of most interest when eQTL allele frequencies vary
283 across diverse populations, such as that shown for rs13900 in *MCP1*, which had a minor allele
284 frequency (MAF) of 4% in Memphis and 38% in Nicaragua (**Fig 2D, Supp Table 1**).

285 ## Divergent age-associated influenza cytokine correlates of severity

286 Results from unstimulated, baseline analyses demonstrate significant immune heterogeneity
287 across diverse populations at the RNA, protein, and cellular level. Further, these data establish
288 that basal immune variation is differentially affected by intrinsic and extrinsic host factors. The
289 cumulative effect of modulation in the basal immune state may lead to unique immune profiles
290 in response to infection. In support of this, the *in vitro* TLR stimulation results show distinct
291 transcriptional responses between the Memphis and Nicaragua cohorts. To further assess the
292 extent to which unique immunophenotypes may occur, we examined how robust immune
293 signatures are across distinct cohorts during acute viral infection *in vivo*.
294

295 We determined cytokine correlates of severity in three independent human cohorts of naturally
296 acquired influenza virus infection. Cytokine levels were measured in plasma and/or the nasal
297 wash from influenza positive subjects in the Memphis (n = 165), Baltimore (n = 222), and
298 Taiwan (n = 226) cohorts. Each cohort was independently analyzed and results were compared
299 to identify unique and shared correlates. Utilizing a subset of samples, a previous study in the
300 Memphis cohort found increased levels of nasal wash IFNa2, MCP3, IL6, and VEGF correlated
301 with total, upper respiratory (URT), and/or lower respiratory (LRT) symptom scores. [33]
302 Additionally, increased plasma MCP3, IL10, and IL6 predicted hospitalization and were
303 positively associated with symptom scores, and IP10 positively correlated with influenza viral
304 load. [33] The current Memphis analysis expands upon this dataset and includes three additional
305 seasons of cytokine data.
306
307 There are key similarities and differences amongst the cohorts included in this analysis that are
308 important to consider in interpretation of the results. First, the Memphis cohort is centered
309 around pediatric recruitment, and thus is enriched in children and young adults, though they
310 also enrolled adult contacts. In contrast, the Baltimore and Taiwan cohorts are both adult
311 studies. Additionally, the Baltimore and Memphis cohorts are similar in genetic and geographical
312 background, whereas Taiwan is distinct from the rest. Further, the Baltimore, and Taiwan
313 cohorts each have binary (presence/absence) disease outcome data whereas the Memphis
314 cohort includes both binary and continuous outcomes, the latter of which requires a different
315 statistical approach. Importantly, binary disease outcomes with a high prevalence dilute the
316 specificity of our severity measures, and thus are removed from downstream analyses. Results
317 below are focused on continuous variables or binary variables with a prevalence <40% (range
318 14-39%, additional details in Study Subjects section of Methods).
319
320 All cohorts exhibit increased levels of plasma IL6, TNFa, GCSF, and IL8 in hospitalized,
321 pneumonia, and/or supplemental oxygen patients. This suggests these cytokines are associated
322 with severe influenza disease in general, irrespective of age and geographical location (**Fig 5,**
323 **Supp Fig 5, Supp Table 5**). However, these analytes may reflect severe (respiratory) viral
324 infection more broadly, as they have also been associated with severe COVID-19.[39–42] Cytokine
325 profiles as a whole, in both the plasma and nasal wash, are distinct by age groups, possibly
326 reflecting unique pathophysiologies. The severe pediatric response is characterized by global
327 hyper-inflammatory cytokine levels, whereas the severe adult response shows largely
328 decreased cytokine levels with a selective inflammatory response of the aforementioned
329 analytes (**Fig 5, Supp Fig 5, Supp Table 5)**. Notably, although the directionality is similar
330 among significant results in the adult cohorts, only decreased IFNg was shared, and thus
331 represents the most robust correlate for this age group. All additional plasma cytokine correlates
332 of severity in these adult populations were also associated with lower cytokine levels, albeit
333 specific to one of the adult cohorts (**Fig 5**, **Supp Fig 5**, **Supp Table 5**).
334
335 These data demonstrate correlates of severity can differ across distinct populations and
336 highlight the need for future work to simultaneously examine immune signatures of health and
337 disease in diverse populations for a more comprehensive analysis and identification of robust
338 therapeutic targets. Additionally, studies which aim to identify cytokine correlates should utilize a

339 broad panel, as focusing on a few select analytes is insufficient to characterize the overall
340 cytokine profile. These results also suggest that a “one size fits all” therapeutic approach may
341 not be efficacious, as distinct immune profiles may exist by extremes of age and likely require
342 different treatments. Multiple severity associated immunotypes were also recently observed in a
343 study of SARS-CoV-2 cellular responses, wherein 3 distinct cellular immunophenotypes were
344 identified for severe COVID-19.[43]

345 Age specific herpesvirus associations with influenza severity

346 In addition to immune correlates of disease severity, independent studies have identified
347 associations between age, sex, and CMV and the outcome of vaccination or infection with
348 influenza in humans.[33,44,45] To expand upon these findings, we examined disease outcome
349 utilizing a linear regression model which accounts for age, sex, CMV, EBV, and HSV serostatus,
350 as very few individuals are only infected with CMV (**Fig 2E-F**, **Supp Fig 1**).
351
352 In Memphis, the pediatric cohort, CMV serostatus was the only significant variable and it was
353 associated with improved outcome. Patients with a positive serostatus are enriched in the non-
354 severe subjects, and show a decrease of 29.1 ± 10.93 ($p = 0.0103$) in peak lower respiratory
355 symptom severity score (range = 0 to 176) (**Fig 6A, Supp Table 6**). Conversely, in the adult
356 Baltimore cohort, a positive CMV serostatus was associated with increased risk of pneumonia
357 (OR = 2.904, $p = 0.0457$) or hospitalization (OR = 4.966, $p = 5.72e-5$) (**Fig 6A, Supp Table 6**).
358
359 Aside from symptom severity, the Memphis cohort included two additional measures of the
360 quality of the immune response to influenza infection - influenza viral load was determined using
361 qRT-PCR (quantitative reverse transcription polymerase chain reaction) and functional antibody
362 titers were measured using microneutralization assays. To analyze these data, linear regression
363 modeling was utilized to assess the effect of biological (age, sex) and infectious (herpesvirus
364 serostatus) factors, while accounting for study day (when applicable). Neither CMV, EBV, nor
365 HSV had any effect on influenza viral load ($p = 0.66$, 0.39, and 0.69, respectively) (**Fig 6B,**
366 **Supp Table 6**). However, a positive HSV serostatus is associated with a decrease in duration of
367 influenza virus shedding by approximately 3.14 ± 1.42 days ($p = 0.0304$, **Fig 6B, Supp Table**
368 **6**). These results suggest that HSV co-infection may alter viral clearance mechanisms, such that
369 it decreases the days an individual sheds influenza virus. Although, there was no detected effect
370 on influenza viral load, one challenge with observational human studies is unknown true day of
371 infection and the viral load upon initial infection.
372
373 Analysis of microneutralization titers show that a positive CMV serostatus is associated with an
374 increase in functional antibody titers by approximately 144 ± 54.50 ($p = 0.0086$) (**Supp Table 6**).
375 Additionally, male sex is associated with an increase in antibody titer of 199 ± 60.70 ($p =$
376 0.00112) and titer decreases by 105.04 ± 30.60 per standard deviation unit change in age ($p =$
377 0.00067). These results are consistent with previous observations by Furman et al. which show
378 that CMV seropositive young adults exhibit augmented antibody responses to Fluzone
379 vaccination compared to CMV seronegative.[44]
380

381 Collectively, these data suggest a dichotomous, age-specific role of CMV on disease severity.
382 This is consistent with prior work which has shown CMV improves vaccine responses in young
383 populations, but increases frailty in the elderly.[44,46–48] Initial benefits may be mediated by trained
384 innate immunity, in which constant surveillance of chronic herpesvirus infection (in order to
385 mitigate reactivation) necessitates altered epigenetic and transcriptional profiles which confer
386 augmented response to restimulation.[49] However, the cumulative effects of CMV on immune
387 aging over time, such as through inflammaging and memory inflation, may eventually exceed
388 beneficial effects and lead to impaired, immunosenescent responses and, subsequently, poorer
389 outcome.[50–52] Future study is required to delineate the extent to which trained innate immunity,
390 CMV associated immunosenescence, or an alternative mechanism is the cause of these
391 differential influences on heterologous immunity.

392 ## Differential contributions of host genetics and herpesviruses to acute immune variation

393 The correlates of severity analysis demonstrates immune heterogeneity in acute immune
394 responses, and thus variation in putative therapeutic targets. Moreover, it identified biological
395 and infectious associations with influenza disease outcome. Having observed biological and
396 infectious contributions to basal cytokine levels, we sought to determine whether the association
397 between disease severity and these factors may be due, at least in part, to their effects on
398 cytokines, particularly those that were identified as a correlate of severity. These results will also
399 provide insight into whether such analyses would be useful in the identification of optimum
400 therapeutic targets by determining the extent to which there are differential contributions of
401 intrinsic and extrinsic host factors on cytokines of interest.
402
403 To achieve this, plasma and nasal wash samples were used from 83 subjects with naturally
404 acquired influenza infection in the Memphis cohort from time points throughout the duration of
405 infection. A total of 38 and 41 cytokines were measured in plasma (n = 237) and nasal wash (n
406 = 325), respectively. Since eQTLs were identified for a subset of analytes, this analysis begins
407 with an assessment of the effects of age, sex, and herpesviruses on cytokine levels, while
408 accounting for influenza infection related variables that can contribute to the immune state
409 (Models #5-6, **Fig 1**). This is followed by a second regression analysis, focused on cytokines
410 with an eQTL identified in the baseline analyses, which expands upon the previous model and
411 incorporates genetics by adding AIMs PCA values and SNP genotypes (Models #7-8, **Fig 1**).
412
413 For all analytes, a linear regression model was used to take into account the effects of: age,
414 sex, study day, influenza viral load, each herpesvirus alone, any interaction between a
415 herpesvirus and age, and any interaction between herpesviruses (Models #5-6, **Fig 1**). These
416 data show each herpesvirus has its own unique signature of immune modulation that is location
417 specific. In plasma (**Fig 6C, Supp Fig 6**), CMV alters the most cytokine levels independently,
418 which coincides with its primary site of latency in cells of the myeloid lineage. Furthermore,
419 herpesvirus interactions are most prevalent in the plasma, versus in the nasal wash, where
420 independent herpesvirus effects are more common. In the nasal wash (**Fig 6F, Supp Fig 7**), the
421 herpesvirus with the most independent effects is EBV, which predominantly infects the
422 nasopharynx region. In the plasma, the adjusted R^2 value reached 0.358 (or 35.8%) for IL10,
423 and herpesvirus interactions are generally associated with the strongest effects (i.e.

424 magnitudes) on cytokine levels as compared to the effects of independent herpesviruses or
425 other variables (i.e. age, influenza viral load) (**Supp Fig 6)**. In the nasal wash, the adjusted R^2
426 value reached 0.332 (or 33.2%) for RANTES, and three factors accounted for most effects on
427 cytokine levels - viral load, age, and EBV, with the strongest effects from the latter two (**Supp**
428 **Fig 7**).
429
430 For the 23 cytokines which had eQTLs and protein level data, a second regression model was
431 performed for plasma and nasal wash, which included each factor previously described (age,
432 sex, herpesviruses, study day, viral load), in addition to the AIMs principal components 1 and 2
433 (genetic based ancestry) and the genotype for each SNP that significantly correlated with the
434 respective cytokine's gene expression (Models #7-8, **Fig 1**). GCSF and its SNPs which trended
435 toward a correlation were also included because fewer samples had detectable gene
436 expression (**Supp Table 7**). In the plasma, when there was an effect from incorporating
437 genetics into the cytokine regression models, it generally increased the amount of variability
438 accounted for by the model; however, the nasal wash models generally decreased. This is likely
439 due, at least in part, to the fact that eQTL status was determined from stimulated PBMCs as
440 opposed to nasal wash cells, which are at the site of infection, have a distinct milieu, and are
441 exposed to the external environment. Eight out of 24 cytokine models had a significant genetic
442 component and increased variability accounted for, two of which solely included a significant
443 AIMs PC without a significant association with a curated SNP (**Supp Table 7**). These cytokines
444 include chemokines (Eotaxin, MDC, MIP1a), immunomodulatory IL10, NK and memory T cell
445 proliferation inducing IL15, and growth factors (FGF2, GCSF, TGFa). The biggest change in
446 adjusted R^2, upon inclusion of genetic factors, occurred in Eotaxin with a 43.35% increase in the
447 variability accounted for by the model, for a final adjusted R^2 value of 0.5568 (**Fig 7A**).
448 Significant factors in the Eotaxin model include: sex, influenza viral load, SNP rs5027581, CMV
449 and its age interaction, the CMV and HSV interaction, the EBV and CMV interaction, and the
450 interaction of EBV, CMV, and HSV. The next highest adjusted R^2 was 0.3915 (39.15%) for
451 plasma IL10 and significant factors include age, sex, viral load, SNP rs3024496, the HSV and
452 age interaction, and the EBV and CMV interaction (**Fig 7B**).

453 ## Validation of the model: rs3024496 genotype modulates IL10 expression

454 To follow up on genetic associations from our analyses, we selected two associated SNPs with
455 potentially functional consequences for *in vitro* assays to assess effects on promoter activity.
456 SNP rs3024496 in *IL10* was chosen because IL10 is a correlate of severe influenza infection,
457 and its genotype prevalence varied across diverse populations (Memphis: 39% TT, 47% CT,
458 14% CC; Nicaragua: 67% TT, 32% CT, 1% CC). Of note, two SNPs within the gene encoding
459 IL10 were identified as eQTLs (rs3024496, rs3024498). However, including both SNPs in the
460 IL10 regression model served as a conditional SNP analysis and suggested there was only one
461 independent signal resulting from rs3024496. SNP rs2304204 in *IRF3* was also selected
462 because it is a transcription factor upstream of antiviral responses, with varying allele
463 prevalences across cohorts (Memphis: 40% GG, 40% AG, 20% AA; Nicaragua: 11% GG, 45%
464 AG, 44% AA). Additionally, an independent analysis using the GTEx Portal confirmed our
465 transcriptional results and showed rs2304204 is an eQTL in tissues such as the thyroid (p =
466 1.1e-19) and lung (p = 0.000027).

For each target, two luciferase reporter assay plasmids were created, which only differed in sequence at the SNP site (**Fig 7C**, detailed in Methods). Unstimulated A549s exhibited higher luciferase expression when the C allele was present for IL10-rs3024496 and when the G allele was present for IRF3-rs2304204 (**Fig 7D**). Results were consistent 24 hours post infection with influenza A virus and showed a higher magnitude of differential activity (**Fig 7E**). Coupled with the eQTL findings, these results confirm rs3024496 is a functional SNP directly affecting IL10 promoter activity and coincide with the association predicted by the regression model - the CC genotype was associated with higher plasma IL10 than the CT and TT genotypes which had coefficient estimates of -0.3315 and -0.6830, respectively, when compared to the reference CC genotype (**Fig 7B**). SNP rs2304204 in *IRF3* was also verified as a functional SNP with coinciding directionality. This finding, coupled with the varying allele prevalence across diverse populations, demonstrates a mechanism by which broad transcriptional changes could occur, both at baseline and in response to immune challenge, as was observed in this study (**Fig 4**). Together, these results demonstrate the integrative manner in which biological, infectious, and genetic factors can influence variation in cytokines at the protein level during naturally acquired influenza infection. The observation that these effects occur in proteins identified as correlates of severity, highlights how variation in these factors can lead to distinct immunophenotypes during acute infection and the potential impact on infectious disease outcome. Collectively, this underscores the need for future studies to search for correlates of protection or severity across diverse populations and to consider the factors which contribute to their variation in order to identify the number of severe immune profiles, the most broadly applicable treatment, and practical therapeutic options (i.e. those more affected by factors we can control or target).

Discussion

Studies of human immune variation thus far suggest genetics and age contribute to 20-40% and 5% of immune variance, respectively, leaving 55-75% possibly due to sex or environmental influences.[53] However, studies to date often focus on a single ancestral group, geographical location, immunological measure, and/or age range, and are sometimes limited in inclusion of both sexes.[4,6–8,20,21,54–56] Often, this is in an attempt to limit additional variance attributed to determinant categories that are not the focus of a given study. While statistically beneficial and still informative, it is impossible to ascertain the scope of immune variation when studies are focused on specific subsets of the human population and this limits the applicability of the conclusions that are drawn. Moreover, this can obscure phenotypes, leading to false positive or negative associations, and inconsistent results across studies that ultimately present a barrier to translating results from the laboratory to effective, broadly applicable therapeutic treatments. Lastly, comparisons of relative importance can be better understood with simultaneous assessment of multiple determinant categories, and with increasing technologies for multiparameter assessment, this approach is becoming more tractable, as demonstrated by this study.

The demographics, immune measures, and analytical approach make this study unique from prior work. First, the populations included are diverse in ancestral background, geographical location, age, and sex. Second, immune measures were determined at baseline and during

510 immune challenge (both *in vitro* and *in vivo*), and include cell phenotypes, cytokines (both
511 circulatory and at the site of infection), and gene expression in response to PRR ligands
512 representative of both bacterial and viral infection. Third, the analytical approach utilized linear
513 regression modeling which allowed us to simultaneously account for and quantify the unique
514 and interactive effects of biological, infectious, and genetic factors. Fourth, our genetic approach
515 was based on ancestry informative markers and a targeted list of candidate immune-related
516 SNPs. This allowed for increased power compared to GWAS studies by limiting the impact of
517 multiple comparisons and provided clearly defined genetic factors that can be further explored in
518 mechanistic studies that aim to identify causal factors compared to heritability scores which are
519 unable to provide direct functional insight.
520
521 Investigation of baseline immunity shows differences in cellular phenotypes, cytokine levels, and
522 the basal transcriptional profile. Heterogeneity in baseline gene expression was primarily
523 influenced by genetic ancestry, followed by (in order of importance) herpesvirus infection, age,
524 and sex. Activation and acquisition of effector function is dependent on both the
525 microenvironment and differentiation state of leukocytes, and variation in these facets of basal
526 immunity were reflected in response to pathogen-like challenge. Stimulation with bacterial or
527 viral ligands showed differential transcriptional responses across the Nicaragua and Memphis
528 cohorts in genes important for each aspect of acute immunity. Differences between cohorts
529 were often dependent on the stimulatory conditions, suggesting that disparate
530 immunophenotypes can be context/infection dependent and further investigation of immune
531 variation under a wide spectrum of immune pressures is needed. Furthermore, the
532 transcriptional response to stimulation is largely affected by genetic ancestry. Remaining
533 determinants, in order of impact, include the basal transcriptional profile, herpesvirus status, and
534 demographic variables, age and sex. Together, these results highlight the substantial impact of
535 genetics on immunity, underscore the need for more studies of diverse ancestral backgrounds,
536 and provide a greater understanding of the relationship between the baseline immune state and
537 the magnitude and quality of acute immunity.
538
539 In further support of unique immune signatures across distinct populations, analysis of plasma
540 cytokines in three independent human cohorts of influenza infection showed two influenza
541 disease severity profiles, largely driven by age, wherein the severe pediatric response exhibits
542 global hyper-inflammatory cytokine levels and the severe adult response exhibits generally
543 decreased cytokine levels, with increases of select cytokines (IL6, TNFa, GCSF, IL8). This
544 indicates different therapeutic approaches may be needed and underscores the need for future
545 work to search for therapeutic targets across diverse populations in order to identify the most
546 robust and broadly applicable options. Such studies would advance the scientific and medical
547 communities closer to achieving personalized medicine by providing valuable insight into the
548 number of distinct severity profiles, the appropriate method to treat severe populations/profiles,
549 and the minimum number of therapeutic approaches required to effectively treat patients of all
550 demographics.
551
552 Results from individual cytokine regression models, both at baseline and during infection, show
553 each category of independent variables contributes to immune variation, these effects are not

554 mutually exclusive, the magnitudes of modulation vary by the cytokine of interest, and there are
555 interactive effects between and within categories. For example, herpesviruses have unique and
556 interactive effects on cytokine levels, suggesting independent and collective mechanisms of
557 immune modulation and emphasizing the need for future studies to include additional
558 herpesviruses in their experimental design. Interestingly, the effects of herpesviruses at
559 baseline are much smaller compared to the effects during acute infection with influenza. This
560 may be due, at least in part, to the limited transcriptional profile of herpesviruses during latency,
561 whereas increased transcriptional activity, perhaps in response to immune cell differentiation or
562 signaling, would allow for greater interference due to increased viral protein production. Indeed,
563 established methods for reactivating herpesviruses include stimulation with PRR ligands and
564 differentiation of myeloid cells.[57,58]
565
566 Of the 23 cytokine models with acute infection (study day, viral load), genetic (SNP genotypes,
567 AIMs), chronic infectious exposure (herpesvirus serostatus) and biological (age, sex) factors, 8
568 had a significant genetic component and improved model performance, each of which were
569 identified as a correlate of influenza disease severity. Functionally relevant SNPs, identified in
570 eQTL and protein cytokine level regression analyses, included SNPs with different allele
571 frequencies across populations and SNPs with similar allele frequencies across populations but
572 variations within populations, highlighting genetic contributions to both inter- and intra-
573 population immune variation, respectively. Importantly, variation in genotype prevalence across
574 or within populations may lead to different immune responses and subsequent immune
575 outcomes if the altered protein is a correlate of severity. For example, SNP rs3024496 was
576 identified as an eQTL for IL10, its distribution of genotypes varied between Memphis and
577 Nicaragua, and prior work in infected subjects from the Memphis cohort found IL10 to be a
578 correlate of hospitalization and symptom severity during influenza infection.[33] Moreover, IL10
579 had the second highest adjusted R^2 value in the infected regression model, and factors which
580 contributed to its variation, at the protein level, included SNP rs3024496, age, sex, viral load,
581 and herpesviruses. Luciferase assays further support the modeling results and show the
582 functional impact of rs3204496, such that the C allele is associated with higher promoter activity
583 when unstimulated or infected with influenza A virus, with the latter condition indicating a higher
584 magnitude of differential gene expression. This demonstrates the powerful implication of
585 genetics on the overall immunophenotype across ancestrally distinct populations, and supports
586 recent work which highlights the value in and need for increased studies of diverse, particularly
587 multiethnic, populations to improve our understanding of human immunology.[25,59,60]
588
589 It's postulated that immune variation can be modeled as an "immunological landscape", with
590 individuals lying at different coordinates according to their immune profile, this provides a
591 framework for the interaction of genetic, environmental, and biological factors in shaping the
592 human immune response.[53] In the initial description, it was suggested that the starting position
593 of an individual within the landscape is largely due to genetics, whereas age changes this
594 position in a consistent and predictable manner across subjects, and local environment may
595 drive convergence between two individuals if shared (such as cohabitation) or diverge two
596 individuals if they are from distinct environmental backgrounds.[53] The results presented here
597 suggest an even more complex and multidimensional immune space with dynamic landscapes

598 depending on the immune pressure and interactions between variables which change the
599 position of an individual, in addition to single variable influences. Our results show there are
600 clusters of distinct immunophenotypes across this landscape, both at baseline and during
601 infection. Importantly, the observation that gene expression differences between two diverse
602 populations depended on the stimulatory condition suggests that the immunological landscape
603 is context dependent, thus unique immunophenotype profiles and clusters may exist during
604 infection that were not observed at baseline. Additionally, unique correlates of severity across
605 diverse populations and shared correlates in cohorts with common features (e.g. age range)
606 suggest distinct clusters of immune profiles are associated with differential illness outcomes and
607 more than one severity-immunotype cluster may exist for a given disease. These observations,
608 and others presented here, likely would have been obscured if the study solely focused on a
609 single population, especially if biased in a given demographic. This underscores the need for
610 future studies to include diverse subjects when examining immune variation, whether at
611 baseline or during infection. While it is challenging to conduct studies that are diverse in all
612 desired categories, this study shows we can leverage multiple cohorts enriched in specific target
613 populations and conduct comparative analyses to gain insight into overlaps or unique signatures
614 of immunity as it relates to immune competence and disease severity.

615 Limitations of the Study

616 Although simultaneous investigation of multiple cohorts provided novel insight, there are several
617 limitations in this study. The independent designs of each cohort and differences in study
618 endpoints affected sampling and data collection. One of the baseline cohorts collected serum,
619 as opposed to plasma; however, we believe the impact was minimal, as it only applies to basal
620 analyses, and a study which compared plasma versus serum cytokine levels, using Luminex
621 multiplex assays, shows results are well correlated.[61] Differences in sampling types, timing, and
622 quantities limited baseline and acute infection PBMC availability for several cohorts. Thus, we
623 were unable to include acute cell phenotypes and connect baseline immune profiles with acute
624 immunophenotypes. Additionally, multiple approaches were utilized for symptom reporting and
625 clinical data collection, which required different statistical approaches for correlates of severity
626 analyses and limited our ability to include other factors, such as body mass index. Lastly, there
627 may be unknown differences in handling and/or storage across the multiple study sites, which
628 we attempted to account for by including "Cohort" as a factor in our models, when appropriate.

Main Figure Legends

Figure 1. *Study populations and model summary.* Table of cohort characteristics (top) and overview of regression model analyses (bottom, labeled A-D in order of appearance).

Figure 2. *Variation of explanatory factors across diverse populations.* Age (A) and sex (B) distribution for each cohort (n = 1,705). (C) Results from the principal component analysis of AIMs (n = 285). (D) Donut plots show the proportion of individuals in each cohort with a given genotype for a SNP of interest (n = 352), plot labels are encoded as follows: Protein-Gene-SNP. (E) Table of herpesvirus serostatus (%) by cohort (n = 1,532). (F) Venn-diagram of herpesvirus infection by cohort. Numbers in the bottom right represent the number of individuals negative for all indicated herpesviruses. For easier visualization, the herpesvirus with the highest overall seroprevalence (VZV) was excluded.

Figure 3. *Basal immune profiles in ancestrally and environmentally distinct populations.* (A) A t-SNE plot based on plasma/serum cytokines and colored by cohort (n = 631). (B) Violin plot demonstrates individual cytokine variation for EGF (n = 623). Black points in the center of the violin plot represent the mean cytokine level and line segments represent +/- the standard deviation. (C) Schematic of herpesvirus effects on basal cytokine levels based on linear regression modeling results (model displayed above diagram, n ≈ 500). Cytokines listed have an adjusted R^2 value > 0.08 and FDR < 0.05. (D) Boxplots depict differences in basal cell phenotypes in PBMC samples from healthy subjects in the Memphis (n = 28) and Nicaragua (n = 19) cohorts. Presented p-values are from Mann-Whitney U tests, post adjustment for multiple comparisons (when applicable).

Figure 4. *Distinct basal immune profiles mold response to immune challenge.* (A) Gene expression by cohort under different stim conditions for IFI16 ($n_{AllStim}$ = 903), CD86 ($n_{AllStim}$ = 906), and IFNG ($n_{AllStim}$ = 774). Displayed p-values were adjusted for multiple comparisons and are from Mann Whitney U tests. (B) Principal component analysis of baseline (unstim) gene expression colored by Cohort (n = 102). (C) Summary of linear regression Models #2 (n ≈ 188) and #3 ($n_{AllStim}$ ≈ 350) that were utilized to determine factors which contribute to baseline and post stim gene expression variation, respectively (see Summary of Models section for details). Displayed values are the number of genes that a factor significantly affected out of a total 90 genes.

Figure 5. *Acute cytokine profiles delineate two age-associated signatures of influenza disease severity.* (A) Heatmap of highlighted results from plasma (left) and nasal wash (right) cytokine correlates of severity analysis, with significant adjusted p-values displayed, for the Memphis (n = 165), Baltimore (n = 222), and Taiwan (n = 226) cohorts. (B) Boxplots show example associations between plasma or nasal wash cytokine levels and symptom severity during influenza infection for each cohort. Presented p-values are from Mann Whitney U tests, and results were adjusted for multiple comparisons by controlling the FDR (< 0.1).

Figure 6. *Herpesviruses have unique and age specific effects on influenza disease severity.* (A-B) Significant herpesvirus associations with influenza disease severity. Severe Memphis was

673 categorized based on lower respiratory symptom severity scores. (C-D) Schematic of
674 herpesvirus effects on acute cytokine levels in plasma (C, n ≈ 200) and nasal wash (D, n ≈ 300)
675 based on linear regression modeling results (models displayed above diagrams). Cytokines
676 listed have an adjusted R^2 value > 0.08 and FDR < 0.05.
677
678 **Figure 7.** *Transcription and protein level of IL10 is modulated by rs3024496.* (A-B) For
679 cytokines with an eQTL, a linear regression model was performed including: sex, age,
680 herpesvirus (HV) serostatus, interactions between HV serostatus and age, interactions between
681 HVs, study day, viral load, AIMs PC1 and PC2, and eQTL SNP genotype (n ≈ 200). The top two
682 results, based on adjusted R^2 value, are shown and include the coefficient estimate of the
683 change in the dependent variable per unit change in the independent variable (Coeff Estimate),
684 the corresponding standard error (Std Error), t score (Statistic), and p-value (highlighted when
685 significant). (C) Schematic of luciferase assays. Boxplots compare Firefly:Renilla ratios from
686 unstimulated (D) or influenza virus infected (E) luciferase assays for A and G alleles of IRF3
687 rs2304204 (left), and T and C alleles of IL10 rs3024496 (right). Presented p-values are from
688 Mann-Whitney U tests.

689 Supplemental Item Legends

690 **Supplemental Figure 1.** *Explanatory variables by Race/Ethnicity and age.* (A-C) Principal
691 component analysis of AIMs colored by Race and/or Ethnicity. (D) Total number of
692 herpesviruses by age, separated by cohort.
693
694 **Supplemental Figure 2.** *Cohort tSNE clusters are not associated with demographics or chronic*
695 *virus infection.* Basal cytokine tSNE plot colored by age (A), sex (B), herpesvirus serostatus (C-
696 F), and Race (G, for Memphis only, all other cohorts in gray). (H) Basal cytokine tSNE with
697 Memphis cohort data removed.
698
699 **Supplemental Figure 3.** *Heatmap summary of baseline cytokine regression models.* (A)
700 Summary of coefficient estimates for each factor included in the baseline plasma/serum linear
701 regression model. Cytokine models are listed in decreasing order of the adjusted R^2 value
702 (shown on the far right) and coefficients are colored if FDR < 0.05. Red is associated with an
703 increase in the cytokine level per unit change in the factor, and green is associated with a
704 decrease in the cytokine level. (B) Basal plasma/serum cytokine level distributions.
705
706 **Supplemental Figure 4.** *Highlights of gene expression associations with independent*
707 *variables.* (A-B) Example of associations based on results from Model #2 (A, determinants of
708 baseline gene expression) and Model #3 (B, determinants of stimulated gene expression). The
709 y-axis is gene expression and the x-axis is the independent variable (age in years, herpesvirus
710 serostatus, ancestry informative marker principal component, baseline gene expression
711 principal component). (C) Examples of eQTLs. *MCP1* (n = 833) and *FLT3LG* (n = 868) gene
712 expression by genotype for rs13900 and rs2946871, respectively.
713
714 **Supplemental Figure 5.** *Plasma and nasal wash cytokine correlates of severity.* Heatmap of
715 plasma (A) and nasal wash (B) cytokine correlates of severity analysis, with significant adjusted

716 p-values displayed. Colors are based on correlation coefficients or a Cohen's D effect size,
717 depending on whether a symptom is continuous or binary, respectively. Presented p-values are
718 from correlations (Spearman and Pearson) or Mann Whitney U tests, and results were adjusted
719 for multiple comparisons by controlling the FDR (< 0.1).
720
721 **Supplemental Figure 6.** *Heatmap summary of acute plasma cytokine regression models.*
722 Samples were from influenza infected subjects. (A) Summary of coefficient estimates for each
723 factor included in the infected plasma linear regression model. Cytokine models are listed in
724 decreasing order of the adjusted R^2 value (shown on the far right) and coefficients are colored if
725 FDR < 0.05. Red is associated with an increase in the cytokine level per unit change in the
726 factor, and green is associated with a decrease in the cytokine level. (B) Acute plasma cytokine
727 level distributions.
728
729 **Supplemental Figure 7.** *Heatmap summary of acute nasal wash cytokine regression models.*
730 Samples were from influenza infected subjects. (A) Summary of coefficient estimates for each
731 factor included in the infected nasal wash linear regression model. Cytokine models are listed in
732 decreasing order of the adjusted R^2 value (shown on the far right) and coefficients are colored if
733 FDR < 0.05. Red is associated with an increase in the cytokine level per unit change in the
734 factor, and green is associated with a decrease in the cytokine level. (B) Acute nasal wash
735 cytokine level distributions.
736
737 **Supplemental Table 1**. *Immune-related SNP genotype frequency in Memphis and Nicaragua*.
738
739 **Supplemental Table 2**. *Results from basal cell phenotype comparisons in PBMCs*
740 *from healthy Nicaraguans and Memphians*.
741
742 **Supplemental Table 3**. *Gene expression Mann Whitney U and linear regression model results*.
743
744 **Supplemental Table 4**. *Regression results for eQTL analysis*.
745
746 **Supplemental Table 5**. *Cytokine correlates of severity analysis results*.
747
748 **Supplemental Table 6**. *Biological and herpesvirus severity associations*.
749
750 **Supplemental Table 7**. *Results from integrative analysis of biological, genetic, and herpesvirus*
751 *effects on cytokine levels during influenza infection*.

752 Acknowledgements

753 We thank the participants and clinical study support staff in each of the cohorts; Emily Walker
754 and the Hartwell Center for processing Infinium Global Screening Array-24 v3.0 chips; Jessica
755 Gaevert for technical assistance with RNA extractions; and ALSAC (American Lebanese Syrian
756 Associated Charities). Select figures (Fig 1, Fig 3C, Fig 6E-F) were created using an academic
757 license of BioRender software. This work was funded by: the Hercules Foundation – Belgium
758 and the University of Antwerp (Methusalem funding; BOF Concerted Research Action), and the

Research Foundation Flanders, project grant G.0409.12N and 1861219N (B.O.); the National Institute of Allergy and Infectious Diseases, US National Institutes of Health, under HHS contract HHSN272201400007C for Johns Hopkins Center of Excellence in Influenza Research and Surveillance (JH CEIRS) (A.P. and R.R.), HHS contract HHSN272201400006C for the St. Jude Center of Excellence for Influenza Research and Surveillance (SJ CEIRS) (R.W., S.S.C., P.G.T.), HHS contract 75N93021C00016 for the St. Jude Centers of Excellence for Influenza Research and Response (SJ CEIRR) (R.W., S.S.C., P.G.T.), HHS contract 75N93019C00052 for the Center for Influenza Vaccine Research for High Risk Populations (CIVR-HRP) of the Collaborative Influenza Vaccine Innovation Centers (R.W., S.S.C., P.G.T.), RO1 5R01AI120997 (A.G.), and RO1 AI107625 (P.G.T.).

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

Conceptualization, A.S. and P.G.T.; Methodology, A.S., E.K.A., L.T., and P.G.T.; Investigation, A.S., E.K.A., C.O., S.W., and T.J.; Formal Analysis, A.S., L.S., and S.P.; Visualization, A.S.; Writing – Original Draft, A.S.; Writing – Review & Editing, A.S., E.K.A., and P.G.T.; Resources, G.K., A.B., R.Z., K.S.S. J.C.D., S.H., G.E., P.V.D., V.V.T., B.O., R.W., S.S.C., A.P., R.R., A.G., P.G.T.; Supervision, P.G.T.; Funding Acquisition, B.O., R.W., S.S.C., A.P., R.R., A.G., P.G.T.

Declaration of Interests

P.G.T. has consulted or received honorarium and/or travel support from Illumina, JNJ, Pfizer, and 10X. P.G.T. serves on the Scientific Advisory Board of ImmunoScape and CytoAgents. All remaining authors declare no competing interests.

Data Availability

Supporting data can be obtained from the authors upon reasonable request.

Code Availability

Analyses were conducted using R, a publicly available software, as detailed in the Methods. All code is available upon request.

Methods

Study Subjects

Memphis Cohort

Plasma, nasal wash, and PBMCs were collected from healthy and influenza infected (n = 383) subjects in Memphis, Tennessee.[33] Recruited subjects were 75% Black/African American and 25% White/Caucasian. To be included in the study, a flu positive patient must meet the clinical

definition of influenza at the time of enrollment and symptoms must have begun within the previous 96 hours or less. Enrollment included index patients and household contacts (both flu positive and negative). Exclusion criteria included refusal to consent to nasal lavage samples, a history of receiving immunoglobulin or other blood products within the three months prior to enrollment, or any condition that in the opinion of the clinical site investigator place the subject at any risk or injury or render the subject unable to meet the study requirements. For index patients, nasal lavages and blood were collected at the day of enrollment (day 0) and days 3, 7, 10, and 28. Household contacts were also enrolled and provided nasal swabs on days 0. 3. 7, and 14, and 28, with blood on days 0 and 28. If at any point a household contact tested influenza positive, they were re-enrolled as a converted contact case and followed the same course of sample collection as the initial index patient. Symptom severity scores were collected throughout the duration of infection and were based on participant rankings using a visual analog scale and data consisted of five categories- total, upper respiratory, lower respiratory, gastrointestinal, and systemic. The peak total and lower respiratory symptom scores were utilized for severity analyses. Hospitalization status was also utilized and occurred in 25.9% of influenza positive participants. Written informed consent was obtained for participants or their guardians and written assent was acquired from age-appropriate subjects. This study was approved by the Institutional Review Boards of St. Jude Children's Research Hospital and the University of Tennessee Health Science Center/Le Bonheur Children's Hospital.

Colombia Cohorts

Healthy plasma samples were collected at a single time point from subjects in the local city of Calabazo (n = 100) and two indigenous populations, Seywiaka (n = 50) and Umandita (n = 50), which are 1.5 and 9.5 hours walk away from Calabazo, respectively.[35] Recruited subjects were 100% Hispanic/Latino. All individuals willing to participate by completing the study questionnaires and permitting the withdrawal of blood and nasal swabs were included in the study. Exclusion criteria included children who are less than two years old when baseline enrollment is performed and anyone with a known immunosuppressive condition or ongoing receipt of immunosuppressive therapy. Signed consent or assent was obtained as appropriate and this study was approved by the Indigenous Health Council, the Tropical Health Foundation ethics committee, and the St. Jude Children's Research Hospital Institutional Review Board.

Belgium Cohort

Healthy serum samples were collected at a single time point from 40 neonatal intensive care unit or pediatric unit nurses.[36] Recruited subjects were 100% White/Caucasian. Inclusion criteria included a minimum of two years working in the respective hospital unit, a maximum of one year between blood sampling and last VZV exposure for pediatric unit nurses, and a minimum of one year between blood sampling and last VZV exposure for neonatal intensive care unit nurses. Exclusion criteria included immunosuppressed state (due to disease or medication), any malignancy in the last five years, receipt of blood products within three months of blood sampling, coagulation disorders, and any condition the investigator believed might interfere with the study. Written informed consent was obtained from all study participants and this study was approved by the ethics board of the Antwerp University Hospital, Antwerp, Belgium.

835 **Baltimore and Taiwan Cohorts**

836 Plasma and/or nasal wash samples were collected from healthy or influenza-infected subjects
837 recruited from the East Baltimore and Bayview campuses in Baltimore, Maryland,
838 USA (n = 225) and the Chang Gung Memorial Hospitals in the Taipei, Taiwan metropolitan area
839 (n = 266).[37,38,62,63] Recruited subjects from Taiwan were 100% Asian, and those from Baltimore
840 were 67% Black/African American, 23% White/Caucasian, 1% American Indian or Alaskan
841 Native, 9% Other. Possible index cases were recruited if, within the past 7 days of enrollment,
842 they reported or documented fever and new or increased cough, headache, or sore throat.
843 Cases were confirmed influenza positive if they received a positive influenza virus test result
844 from the hospital visit (GeneXpert). Exclusion criteria included children younger than the age of
845 five, currently incarcerated, unable to write or speak English, unable to provide informed
846 consent, or unable to provide a telephone number for follow-up. Clinical coordinators obtained
847 written informed consent, demographic, and clinical data. Samples were collected at the initial
848 visit and at a 3-5 week follow-up visit. Additionally, symptom information included the presence
849 or absence of wheezing and/or shortness of breath (WSOB), pneumonia (based on radiological
850 or microscopic findings), hospitalization (for influenza associated illness), and administration of
851 supplemental oxygen. In the Baltimore cohort, these symptoms occurred in 75%, 26%, 31%,
852 and 17% of subjects, respectively. In the Taiwan cohort, the prevalence was 67%, 39%, 24%,
853 and 19%, respectively. Due to the high prevalence of WSOB, it was excluded from downstream
854 analyses for both cohorts. These studies were reviewed and approved by the Institutional
855 Review Boards of Johns Hopkins School of Medicine and the Chang Gung Memorial Hospital
856 (IRB00091667).

857 **Nicaragua Cohort**

858 Plasma samples were collected from healthy and influenza-infected subjects (n = 591) recruited
859 at the Health Center Sócrates Flores Vivas (HCSFV) in Managua, Nicaragua.[34] Recruited
860 subjects were 100% Hispanic/Latino. Inclusion criteria include that an index case must live in
861 the HCSFV catchment area, be positive for influenza by QuickVue Influenza A+B rapid test,
862 experience onset of symptoms within the previous 48 hours, live with at least one additional
863 person who agrees to participate in the study, and have no household contacts who have had
864 influenza-like illness in the previous two weeks. There are no exclusion criteria for participation
865 in this study. Household members were also recruited and monitored for influenza infection
866 through 5 home visits, 2-3 days apart, over a 10-14 day period. Blood samples were collected at
867 the initial visit and a 30 day follow-up visit, and participants submitted questionnaires regarding
868 signs and symptoms of acute respiratory illness. Participants provided written informed consent
869 and parental permission was obtained from parents or legal guardians of children, in addition to
870 verbal assent from children aged six years and older. This study was approved by the
871 Institutional Review Boards at the University of Michigan, Centro Nacional de Diagnóstico y
872 Referencia (Ministry of Health, Nicaragua), and University of California, Berkeley.

873 Herpesvirus Serostatus

874 Cytomegalovirus (CMV), Epstein-Barr virus (EBV), herpes simplex virus 1 (HSV1), herpes
875 simplex virus 2 (HSV2), and human herpesvirus 6 (HHV6) serostatus was determined

using plasma and the following ELISA kits, respectively: GWB-892399, GWB- DA817E, GWB-ECB3C0, GWB-C2349A, and Abnova KA1457.

Cytokine Measurement

Plasma and nasal wash cytokine, chemokine, and growth factor levels were measured using the Luminex MAP system with the Milliplex HCYTOMAG-60K assay, and samples were processed according to the kit manual. Prior to statistical analyses, analytes were removed from the dataset if less than 15-40% of their values were above or below the lower and upper limits of detection, respectively. This value was selected based on total sample size and the prevalence of independent and dependent variables of interest.

Flow Cytometry

PBMCs were stained for the following surface markers: Panel 1) CCR7-FITC (Biolegend Cat. 353216), CD25-APC (Biolegend Cat. 302610), CD3-PerCP eFluor710 (Thermo Cat. 46-0037-42), CD38-BV785 (Biolegend Cat. 303530), CD4-BV650 (BD Biosciences Cat. 563875), CD45RA-PE Dazzle/594 (Biolegend Cat. 304146), CD45RO-BV605 (Biolegend Cat. 304238), CD69-AF700 (BD Biosciences Cat. 310922), CD8-APC Fire750 (Biolegend Cat. 344746), gd TCR-PE/Cy7 (Biolegend Cat. 331222), PD1-BV711 (Biolegend Cat. 329928), TIGIT-PE (Biolegend Cat. 372704), and TIM3-BV421 (BD Biosciences Cat. 565562). Panel 2) CD11b-BV785 (Biolegend Cat. 301346), CD11c-PE/Cy7 (Biolegend Cat. 301608), CD123-BV421 (Biolegend Cat. 306018), CD14-APC eFluor780 (Thermo Cat. 47-0149-42), CD16-FITC (Biolegend Cat. 360715), CD19-PerCP eFluor710 (Thermo Cat. 46-0198-42), CD1c-BV711 (Biolegend Cat. 331536), CD3-PerCP eFluor710 (Thermo Cat. 46-0037-42), CD56-BV605 (Biolegend Cat. 318334), CD68-APC (Biolegend Cat. 333810), CD80-PE (Biolegend Cat. 305208), and HLA DR-BV650 (Biolegend Cat. 307650). Panel 3) CD3-AF700 (Biolegend Cat. 300323), CD4-APC Fire750 (Biolegend Cat. 300559), FOXP3-PE (Biolegend Cat. 320107). Briefly, after thawing from cryopreservation and plating in a 96-well round bottom plate, cells were spun down and resuspended in 50µL of human Fc block (BD Biosciences Cat. 564220) in DPBS and incubated for 10 minutes at room temperature. Afterwards, 50µL of a Live/Dead Aqua (Tonbo Cat. 13-0870-T100) and pre-titrated surface antibody cocktail in DPBS were added to each well and cells were incubated for 30 minutes on ice and in the dark. Additionally, the BD Cytofix/Cytoperm kit (BD Biosciences Cat. 554714) and FOXP3 Transcription Factor Staining Buffer set (eBioscience Cat. 00-5523-00) were used for intracellular staining of CD68-APC (Panel 2) and FOXP3-PE (Panel 3), respectively and in accordance with the kit manual. Cells were washed, resuspended in FACS (fluorescence-activated cell sorting) buffer and analyzed on a BD Fortessa cell analyzer and analyzed using FlowJo software.

Microneutralization Assays

Microneutralization titers were performed according to guidelines outlined by the World Health Organization (WHO) (Global Influenza Surveillance Network, 2011). Reported results are the reciprocal dilution which inhibited virus growth as determined by an ELISA readout or hemagglutination assay.

Viral Titers

Influenza viral load was measured with quantitative reverse-transcription real-time polymerase chain reaction (qRT-PCR) assays and RNA from nasal swabs. Nucleic acid was extracted from samples using the MagMAXtm- 96 AI/ND Viral RNA Isolation kit on the KingFisher magnetic processor platform. Afterwards, a 25mL qRT-PCR reaction was set up utilizing the Qiagen One-Step RT-PCR kit, influenza A matrix or influenza B non-structural protein primers and the protocol outlined by the WHO(Global Influenza Surveillance Network, 2011).

Human Genetic Studies

Prioritization of SNPs

The UCSC Human Genome Browser build GRCh37/19 was used to curate a list of 308 common SNPs in genes encoding 41 cytokines measured by the Milliplex assay HCYTOMAG-60K and for 3 transcription factors important in initiating antiviral immunity, namely IRF3, IRF7, and all 3 subunits of NFkB. Briefly, putative functional SNPs were prioritized and included the following classes of SNPs: missense, splice, non-coding, and those which are located in the 3' or 5' untranslated region (UTR) of the promoter region. Utilizing 1000 Genomes Project data, the curated list of SNPs was reduced to 209 by filtering for SNPs with minor allele frequencies (MAF) which significantly vary across ancestrally distinct populations. For a given SNP, this was calculated using the MAFs of the African (AFR), American (AMR), Eastern Asian (EAS), European (EUR), and South Asian (SAS) populations. A SNP was considered to significantly vary if the standard deviation of the MAFs across these populations was greater than 5. Lastly, the prioritized list of SNPs was reduced to 94 following a tagSNP analysis to identify the most informative SNPs in a region based on linkage disequilibrium. The tagSNP analysis was conducted by utilizing the SNPclip tool from the National Cancer Institute (https://ldlink.nci.nih.gov/?tab=snpclip) and data from African (YRI), African American (ASW), Mexican (MXL), and European (CEU) populations, as these most reflect the populations in this study.

DNA Extraction

For the Memphis cohort, 90% of samples were extracted in a clinical HLA core at St. Jude, the remaining 10% were extracted from peripheral blood mononuclear cells using the Zymo Quick-DNA Microprep Plus Kit (Cat. D4074). All DNA samples for the Nicaragua cohort were extracted from 200mL of neutrophil pellets using the Qiagen QIAamp DNA Mini Kit (Cat. 51306).

All samples were quantified using a Nanodrop, the average 260/280 ratio was 1.86 (range 1.16 - 2.04). For Fluidigm SNP Type assays, all samples met the minimum DNA input requirement of 10ng/mL, and all ancestry informative marker samples had an input of 250ng.

SNP Genotyping

Genotypes for the curated list of 94 SNPs were determined using the Fluidigm SNP Type

954 Assay on the BioMark HD System and 96.96 Dynamic Array integrated fluidic circuits.
955 SNP Type assays were designed using Fluidigm's D3 Assay Design Tool. The
956 assay was performed in accordance with the Fluidigm protocol, using PCR grade water
957 as a negative control, and including the specific target amplification step with 14 cycles of
958 amplification.

959 **Ancestry Informative Markers**

960 Genetic based ancestry was determined using the Infinium Global Screening Array-24 v3.0
961 BeadChip which measures 654,027 markers including multiethnic genome-wide content,
962 curated clinical research variants, and quality control markers. High quality DNA was submitted
963 to the St. Jude Hartwell Center for preparation and processing.

964 ## Gene Expression Assays

965 **PBMC Stimulation**

966 PBMCs from 188 subjects (94 Memphians and 94 Nicaraguans) were plated at a max of
967 850,000 cells per well and resuspended in 1µg/mL of lipopolysaccharide, Pam3CSK4,
968 transfected poly(I:C), transfected poly(dA:dT), or plain media as an unstimulated control.
969 Plates were incubated at 37°C for 24 hours. Following stimulation, plates were spun at
970 500g for 5 minutes. Supernatants were transferred to clean 96-well plates and stored at
971 -80°C. Pellets were resuspended in 300µL of lysis buffer and stored at -80°C in a labeled
972 epitube until ready for extraction.

973 **Gene Expression**

974 RNA was extracted using the Zymo Quick-RNA extraction kit (Cat. R1050), and stored in a 96-
975 well plate at -80°C until ready for PCR. Prior to running gene expression assays, stored RNA
976 was thawed and normalized to the highest amount of RNA possible, with respect to stimulation
977 condition. The input RNA was 14ng for unstimulated control, 18ng for Pam3CSK4, 40ng for
978 lipopolysaccharide, 50ng for transfected poly(I:C), and 50ng for transfected poly(dA:dT). The
979 expression of 96 immune genes was measured using the Fluidigm Fast Gene Expression
980 Analysis Using EvaGreen on the BioMark HD System and 96.96 Dynamic Array integrated
981 fluidic circuits. Primers were designed using Fluidigm's D3 Assay Design Tool. The assay was
982 performed in accordance with the Fluidigm protocol, including the specific target amplification
983 step with 16 cycles of amplification, and 2 water samples per chip as negative controls. Quality
984 control of gene expression results included removing samples/genes with no measurable gene
985 expression or samples with melting curve peaks that were more than one degree away from the
986 consensus peak melting temperature. Following QC steps, 4 genes were removed altogether
987 (*IFNL2*, *IL9, SFTPD*, *TMEM173*) and 92 remained for further analysis (2 housekeeping - *ACTB,*
988 *GAPDH*; 90 immune - *AIM2*, *AREG*, *BIK*, *CCL11*, *CCL2*, *CCL22*, *CCL3*, *CCL4*, *CCL5*, *CCL7*,
989 *CD40LG*, *CD69*, *CD80*, *CD86*, *CGAS*, *CSF3*, *CX3CL1*, *CX3CR1*, *CXCL1*, *CXCL10*, *CXCL2*,
990 *DDX58*, *EGF*, *EOMES*, *FGF2*, *FLT3LG*, *GATA3*, *GZMB*, *IFI16*, *IFIT1*, *IFITM3*, *IFNA2*, *IFNB1*,
991 *IFNG*, *IFNL1*, *IL10*, *IL12A*, *IL12B*, *IL13*, *IL15*, *IL17A*, *IL18*, *IL1A*, *IL1B*, *IL1RN*, *IL2*, *IL4*, *IL5*, *IL6*,
992 *IL8*, *IRF1*, *IRF3*, *IRF7*, *IRF9*, *ISG15*, *JUN*, *KLRC1*, *KLRK1*, *LTA*, *MAVS*, *MX1*, *MX2*, *NFATC1*,
993 *NFKB1*, *NFKB2*, *NFKBIA*, *NLRP3*, *OAS1*, *PDGFA*, *PDGFB*, *PTGS2*, *REL*, *RELA*, *RELB*,

994 *RORC*, *SELL*, *SOCS1*, *SOCS3*, *STAT1*, *STAT3*, *STAT4*, *TBK1*, *TBX21*, *TGFA*, *TGFB1*, *TNF*,
995 *TNFAIP3*, *TREX1*, *VEGFA*, *VEGFC*). Lastly, prior to conducting comparative analyses, results
996 were normalized again using housekeeping gene expression (*ACTB*).

997 ## Luciferase Assays

998 Determination of IL10-rs3024496 and IRF3-rs2304204 as functional SNPs was achieved using
999 luciferase assay constructs ordered from GenScript. For IL10-rs3024496, a 1,1114bp region
1000 (chr1:206,945,768-206,946,881, GRCh37/19), encompassing the IL10 promoter, was inserted
1001 upstream of the *luc2* gene in the pGL4.10[*luc2*] plasmid (Promega Cat. E6651). Following this
1002 step, two final constructs were created, one with the C allele and one with the T allele for
1003 rs3024496, by cloning a 2,364bp region (chr. 1:206,939,683-206,942,046, GRCh37/19) that
1004 encompassed the IL10 3'UTR and the SNP of interest. For IRF3-rs2304204, a 1,394bp region
1005 (chr19:50,168,220-50,169,613, GRCh37/19), encompassing the IRF3 promoter, was inserted
1006 upstream of the *luc2* gene in the pGL4.10[*luc2*] plasmid (Promega Cat. E6651). As the SNP
1007 was located within the promoter region, two versions of this construct were created and cloned,
1008 one with the G allele and one with the A allele for rs2304204.
1009
1010 A549s (ATCC Cat. CCL-185) were transfected with each plasmid using the TransIT LT
1011 Transfection Reagent (Mirius Bio Cat. MIR 2300). Briefly, $1.5x10^5$ cells/well were seeded into
1012 12-well plates, and incubated overnight at 37°C and 5% CO_2. The following day, 1µg of
1013 pGL4.10 test plasmid and 0.01µg of pRL-TK control plasmid (Promega Cat. E2241) were added
1014 per well, according to the TransIT LT protocol. All transfections were conducted in triplicate. At
1015 24 hours post transfection, cells were collected and luciferase expression was measured using
1016 the Dual-Luciferase Assay kit (Promega Cat. E1910) on the GloMax platform in accordance with
1017 the kit protocol. For influenza stimulated results, at 24 hours post transfection, cells were
1018 washed with PBS, 1mL of $6.0x10^5$ pfu/mL Influenza A virus (A/Singappore/INFIMH-16-
1019 019/2016, multiplicity of infection = 4) was added to each well, and plates were incubated at
1020 37°C and 5% CO_2 for 2 hours. Untreated control plates received infection media instead of
1021 inoculum. Following the 2 hour incubation, inoculum was removed, cells were washed with PBS,
1022 1mL of infection media was added to each well, and plates were incubated for an additional 24
1023 hours at 37°C and 5% CO_2, after which luciferase expression was measured.

1024 ## Statistical Analysis

1025 Detailed descriptions of specific applied statistical approaches can be found in figure legends
1026 and/or the main text. All analyses were performed with R using base functions (lm, p.adjust) and
1027 generally the following packages: Rtsne, GGally, pheatmap, ggpubr, broom, formattable, and
1028 tidyverse.[64–70] Dimension reduction techniques included principal component analysis and t-
1029 Distributed Stochastic Neighbor Embedding. For the cytokine tSNE analysis, any missing values
1030 were replaced with the median of that cytokine level, with respect to the cohort. Pearson and
1031 Spearman correlations were used to determine cytokine correlates of severity for continuous
1032 symptom outcomes. The unstimulated gene expression PCA required complete observations
1033 (i.e. no missing values), and thus was generated from 102 subjects and 67 genes. Multivariate
1034 analyses were conducted utilizing logistic or linear regression modeling, depending on whether
1035 the outcome variable was binary or continuous, respectively. For all models, continuous

variables were standardized using the scale() function in R. All remaining analyses were Mann-Whitney U or Kruskal Wallis tests, depending on whether there were two or more comparison groups, respectively. All herpesvirus association analyses only included CMV, EBV, and HSV, due to the overall high seroprevalence of HHV6 (%) and VZV (%). When appropriate and indicated, p-values were adjusted for multiple comparisons using the Benjamini-Hochberg method for controlling the false discovery rate.[71]

1042 ## References

1043 1. Tsang, J.S., Schwartzberg, P.L., Kotliarov, Y., Biancotto, A., Xie, Z., Germain, R.N., Wang,
1044 E., Olnes, M.J., Narayanan, M., Golding, H., et al. (2014). Global Analyses of Human
1045 Immune Variation Reveal Baseline Predictors of Postvaccination Responses. Cell *157*, 499–
1046 513. 10.1016/j.cell.2014.03.031.
1047 2. Hwang, S., Kwon, A.-Y., Jeong, J.-Y., Kim, S., Kang, H., Park, J., Kim, J.-H., Han, O.J., Lim,
1048 S.M., and An, H.J. (2020). Immune gene signatures for predicting durable clinical benefit of
1049 anti-PD-1 immunotherapy in patients with non-small cell lung cancer. Sci. Rep. *10*, 643.
1050 10.1038/s41598-019-57218-9.
1051 3. Vujkovic-Cvijin, I., Dunham, R.M., Iwai, S., Maher, M.C., Albright, R.G., Broadhurst, M.J.,
1052 Hernandez, R.D., Lederman, M.M., Huang, Y., Somsouk, M., et al. (2013). Dysbiosis of the
1053 Gut Microbiota Is Associated with HIV Disease Progression and Tryptophan Catabolism.
1054 Sci. Transl. Med. *5*, 193ra91-193ra91. 10.1126/scitranslmed.3006438.
1055 4. Orrù, V., Steri, M., Sole, G., Sidore, C., Virdis, F., Dei, M., Lai, S., Zoledziewska, M.,
1056 Busonero, F., Mulas, A., et al. (2013). Genetic Variants Regulating Immune Cell Levels in
1057 Health and Disease. Cell *155*, 242–256. 10.1016/j.cell.2013.08.041.
1058 5. Drake, K.A., Torgerson, D.G., Gignoux, C.R., Galanter, J.M., Roth, L.A., Huntsman, S., Eng,
1059 C., Oh, S.S., Yee, S.W., Lin, L., et al. (2014). A genome-wide association study of
1060 bronchodilator response in Latinos implicates rare variants. J. Allergy Clin. Immunol. *133*,
1061 370-378.e15. 10.1016/j.jaci.2013.06.043.
1062 6. Ye, C.J., Feng, T., Kwon, H.-K., Raj, T., Wilson, M.T., Asinovski, N., McCabe, C., Lee, M.H.,
1063 Frohlich, I., Paik, H., et al. (2014). Intersection of population variation and autoimmunity
1064 genetics in human T cell activation. Science *345*. 10.1126/science.1254665.
1065 7. Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S.,
1066 Dekker, C.L., Swan, G.E., Butte, A.J., et al. (2015). Variation in the Human Immune System
1067 Is Largely Driven by Non-Heritable Influences. Cell *160*, 37–47. 10.1016/j.cell.2014.12.020.
1068 8. Carr, E.J., Dooley, J., Garcia-Perez, J.E., Lagou, V., Lee, J.C., Wouters, C., Meyts, I., Goris,
1069 A., Boeckxstaens, G., Linterman, M.A., et al. (2016). The cellular composition of the human
1070 immune system is shaped by age and cohabitation. Nat. Immunol. *17*, 461–468.
1071 10.1038/ni.3371.
1072 9. Lu, Y., Biancotto, A., Cheung, F., Remmers, E., Shah, N., McCoy, J.P., and Tsang, J.S.
1073 (2016). Systematic Analysis of Cell-to-Cell Expression Variation of T Lymphocytes in a
1074 Human Cohort Identifies Aging and Genetic Associations. Immunity *45*, 1162–1175.
1075 10.1016/j.immuni.2016.10.025.
1076 10. Li, Y., Oosting, M., Smeekens, S.P., Jaeger, M., Aguirre-Gamboa, R., Le, K.T.T., Deelen,
1077 P., Ricaño-Ponce, I., Schoffelen, T., Jansen, A.F.M., et al. (2016). A Functional Genomics
1078 Approach to Understand Variation in Cytokine Production in Humans. Cell *167*, 1099-
1079 1110.e14. 10.1016/j.cell.2016.10.017.
1080 11. Molony, R.D., Nguyen, J.T., Kong, Y., Montgomery, R.R., Shaw, A.C., and Iwasaki, A.
1081 (2017). Aging impairs both primary and secondary RIG-I signaling for interferon induction in
1082 human monocytes. Sci. Signal. *10*, eaan2392. 10.1126/scisignal.aan2392.
1083 12. Allen, E.K., Randolph, A.G., Bhangale, T., Dogra, P., Ohlson, M., Oshansky, C.M., Zamora,
1084 A.E., Shannon, J.P., Finkelstein, D., Dressen, A., et al. (2017). SNP-mediated disruption of
1085 CTCF binding at the IFITM3 promoter is associated with risk of severe influenza in humans.
1086 Nat. Med. *23*, 975–983. 10.1038/nm.4370.
1087 13. Henrickson, S.E., Manne, S., Dolfi, D.V., Mansfield, K.D., Parkhouse, K., Mistry, R.D.,
1088 Alpern, E.R., Hensley, S.E., Sullivan, K.E., Coffin, S.E., et al. (2018). Genomic Circuitry
1089 Underlying Immunological Response to Pediatric Acute Respiratory Infection. Cell Rep. *22*,
1090 411–426. 10.1016/j.celrep.2017.12.043.
1091 14. Gignoux, C.R., Torgerson, D.G., Pino-Yanes, M., Uricchio, L.H., Galanter, J., Roth, L.A.,

1092 Eng, C., Hu, D., Nguyen, E.A., Huntsman, S., et al. (2019). An admixture mapping meta-
1093 analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos. J.
1094 Allergy Clin. Immunol. *143*, 957–969. 10.1016/j.jaci.2016.08.057.
1095 15. Lim, H.K., Huang, S.X.L., Chen, J., Kerner, G., Gilliaux, O., Bastard, P., Dobbs, K.,
1096 Hernandez, N., Goudin, N., Hasek, M.L., et al. (2019). Severe influenza pneumonitis in
1097 children with inherited TLR3 deficiency. J. Exp. Med. *216*, 2038–2056.
1098 10.1084/jem.20181621.
1099 16. Kotliarov, Y., Sparks, R., Martins, A.J., Mulè, M.P., Lu, Y., Goswami, M., Kardava, L.,
1100 Banchereau, R., Pascual, V., Biancotto, A., et al. (2020). Broad immune activation underlies
1101 shared set point signatures for vaccine responsiveness in healthy individuals and disease
1102 activity in patients with lupus. Nat. Med. *26*, 618–629. 10.1038/s41591-020-0769-8.
1103 17. Herati, R.S., Silva, L.V., Vella, L.A., Muselman, A., Alanio, C., Bengsch, B., Kurupati, R.K.,
1104 Kannan, S., Manne, S., Kossenkov, A.V., et al. (2021). Vaccine-induced ICOS+CD38+
1105 circulating Tfh are sensitive biosensors of age-related changes in inflammatory pathways.
1106 Cell Rep. Med. *2*. 10.1016/j.xcrm.2021.100262.
1107 18. Brenes-Chacon, H., Garcia-Mauriño, C., Moore-Clingenpeel, M., Mertz, S., Ye, F., Cohen,
1108 D.M., Ramilo, O., and Mejias, A. (2021). Age-dependent Interactions Among Clinical
1109 Characteristics, Viral Loads and Disease Severity in Young Children With Respiratory
1110 Syncytial Virus Infection. Pediatr. Infect. Dis. J. *40*, 116–122.
1111 10.1097/INF.0000000000002914.
1112 19. Erwin, E.A., Jaramillo, L.M., Smith, B., Kruszewski, P.G., Kahwash, B., Grayson, M.H.,
1113 Mejias, A., and Ramilo, O. (2021). Sex Differences in Blood Transcriptional Profiles and
1114 Clinical Phenotypes in Pediatric Patients with Eosinophilic Esophagitis. J. Allergy Clin.
1115 Immunol. Pract. *9*, 3350-3358.e8. 10.1016/j.jaip.2021.06.043.
1116 20. O'Neill, M.B., Quach, H., Pothlichet, J., Aquino, Y., Bisiaux, A., Zidane, N., Deschamps, M.,
1117 Libri, V., Hasan, M., Zhang, S.-Y., et al. (2021). Single-Cell and Bulk RNA-Sequencing
1118 Reveal Differences in Monocyte Susceptibility to Influenza A Virus Infection Between
1119 Africans and Europeans. Front. Immunol. *12*.
1120 21. Randolph, H.E., Fiege, J.K., Thielen, B.K., Mickelson, C.K., Shiratori, M., Barroso-Batista,
1121 J., Langlois, R.A., and Barreiro, L.B. (2021). Genetic ancestry effects on the response to
1122 viral infection are pervasive but cell type specific. Science. 10.1126/science.abg0928.
1123 22. Alanio, C., Verma, A., Mathew, D., Gouma, S., Liang, G., Dunn, T., Oldridge, D.A., Weaver,
1124 J.E., Kuri-Cervantes, L., Pampena, M.B., et al. (2022). Cytomegalovirus latent infection is
1125 associated with an increased risk of COVID-19-related hospitalization. J. Infect. Dis.,
1126 jiac020. 10.1093/infdis/jiac020.
1127 23. Schönfelder, K., Breuckmann, K., Elsner, C., Dittmer, U., Fistera, D., Herbstreit, F., Risse,
1128 J., Schmidt, K., Sutharsan, S., Taube, C., et al. (2021). The influence of IFITM3
1129 polymorphisms on susceptibility to SARS-CoV-2 infection and severity of COVID-19.
1130 Cytokine *142*, 155492. 10.1016/j.cyto.2021.155492.
1131 24. Cuesta-Llavona, E., Albaiceta, G.M., García-Clemente, M., Duarte-Herrera, I.D., Amado-
1132 Rodríguez, L., Hermida-Valverde, T., Enríquez-Rodriguez, A.I., Hernández-González, C.,
1133 Melón, S., Alvarez-Argüelles, M.E., et al. (2021). Association between the interferon-
1134 induced transmembrane protein 3 gene (IFITM3) rs34481144 / rs12252 haplotypes and
1135 COVID-19. Curr. Res. Virol. Sci. *2*, 100016. 10.1016/j.crviro.2021.100016.
1136 25. Graham, S.E., Clarke, S.L., Wu, K.-H.H., Kanoni, S., Zajac, G.J.M., Ramdas, S., Surakka, I.,
1137 Ntalla, I., Vedantam, S., Winkler, T.W., et al. (2021). The power of genetic diversity in
1138 genome-wide association studies of lipids. Nature *600*, 675–679. 10.1038/s41586-021-
1139 04064-3.
1140 26. Ishigaki, K., Sakaue, S., Terao, C., Luo, Y., Sonehara, K., Yamaguchi, K., Amariuta, T., Too,
1141 C.L., Laufer, V.A., Scott, I.C., et al. (2022). Multi-ancestry genome-wide association
1142 analyses identify novel genetic mechanisms in rheumatoid arthritis. Nat. Genet. *54*, 1640–

1651. 10.1038/s41588-022-01213-w.
27. Mersha, T.B., and Abebe, T. (2015). Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. Hum. Genomics *9*, 1. 10.1186/s40246-014-0023-x.
28. Borrell, L.N., Elhawary, J.R., Fuentes-Afflick, E., Witonsky, J., Bhakta, N., Wu, A.H.B., Bibbins-Domingo, K., Rodríguez-Santana, J.R., Lenoir, M.A., Gavin, J.R., et al. (2021). Race and Genetic Ancestry in Medicine — A Time for Reckoning with Racism. N. Engl. J. Med. *384*, 474–480. 10.1056/NEJMms2029562.
29. Virgin, H.W., Wherry, E.J., and Ahmed, R. (2009). Redefining Chronic Viral Infection. Cell *138*, 30–50. 10.1016/j.cell.2009.06.036.
30. Kahan, S.M., Wherry, E.J., and Zajac, A.J. (2015). T cell exhaustion during persistent viral infections. Virology *479–480*, 180–193. 10.1016/j.virol.2014.12.033.
31. Souquette, A., and Thomas, P.G. (2018). Past Life and Future Effects—How Heterologous Infections Alter Immunity to Influenza Viruses. Front. Immunol. *9*, 1071. 10.3389/fimmu.2018.01071.
32. Sehrawat, S., Kumar, D., and Rouse, B.T. (2018). Herpesviruses: Harmonious Pathogens but Relevant Cofactors in Other Diseases? Front. Cell. Infect. Microbiol. *8*.
33. Oshansky, C.M., Gartland, A.J., Wong, S.-S., Jeevan, T., Wang, D., Roddam, P.L., Caniza, M.A., Hertz, T., Devincenzo, J.P., Webby, R.J., et al. (2014). Mucosal immune responses predict clinical outcomes during influenza infection independently of age and viral load. Am. J. Respir. Crit. Care Med. *189*, 449–462. 10.1164/rccm.201309-1616OC.
34. Ng, S., Lopez, R., Kuan, G., Gresh, L., Balmaseda, A., Harris, E., and Gordon, A. (2016). The Timeline of Influenza Virus Shedding in Children and Adults in a Household Transmission Study of Influenza in Managua, Nicaragua. Pediatr. Infect. Dis. J. *35*, 583–586. 10.1097/INF.0000000000001083.
35. Altan, E., Dib, J.C., Gulloso, A.R., Juandigua, D.E., Deng, X., Bruhn, R., Hildebrand, K., Freiden, P., Yamamoto, J., Schultz-Cherry, S., et al. (2019). Effect of Geographic Isolation on the Nasal Virome of Indigenous Children. J. Virol. *93*. 10.1128/JVI.00681-19.
36. Elias, G., Souquette, A., Heynderickx, S., Meester, I.D., Jansens, H., Beutels, P., Damme, P.V., Smits, E., Thomas, P.G., Tendeloo, V.V., et al. (2018). Altered CD4+ T cell immunity in nurses occupationally exposed to viral pathogens. Clin. Exp. Immunol. *194*, 192–204. 10.1111/cei.13193.
37. Liu, H., Gong, Y.-N., Shaw-Saliba, K., Mehoke, T., Evans, J., Liu, Z.-Y., Lewis, M., Sauer, L., Thielen, P., Rothman, R., et al. (2021). Differential disease severity and whole-genome sequence analysis for human influenza A/H1N1pdm virus in 2015–2016 influenza season. Virus Evol. *7*, veab044. 10.1093/ve/veab044.
38. Blumenkrantz, D.R., Mehoke, T., Shaw-Saliba, K., Powell, H., Wohlgemuth, N., Liu, H., Macias, E., Evans, J., Lewis, M., Medina, R., et al. (2021). Identification of H3N2 NA and PB1-F2 genetic variants and their association with disease symptoms during the 2014–15 influenza season. Virus Evol. *7*, veab047. 10.1093/ve/veab047.
39. Mudd, P.A., Crawford, J.C., Turner, J.S., Souquette, A., Reynolds, D., Bender, D., Bosanquet, J.P., Anand, N.J., Striker, D.A., Martin, R.S., et al. (2020). Distinct inflammatory profiles distinguish COVID-19 from influenza with limited contributions from cytokine storm. Sci. Adv., eabe3024. 10.1126/sciadv.abe3024.
40. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet *395*, 497–506. 10.1016/S0140-6736(20)30183-5.
41. Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science *369*, 718–724. 10.1126/science.abc6027.

1194 42. Bonnet, B., Cosme, J., Dupuis, C., Coupez, E., Adda, M., Calvet, L., Fabre, L., Saint-
1195 Sardos, P., Bereiziat, M., Vidal, M., et al. (2021). Severe COVID-19 is characterized by the
1196 co-occurrence of moderate cytokine inflammation and severe monocyte dysregulation.
1197 eBioMedicine *73*. 10.1016/j.ebiom.2021.103622.
1198 43. Mathew, D., Giles, J.R., Baxter, A.E., Oldridge, D.A., Greenplate, A.R., Wu, J.E., Alanio, C.,
1199 Kuri-Cervantes, L., Pampena, M.B., D'Andrea, K., et al. (2020). Deep immune profiling of
1200 COVID-19 patients reveals distinct immunotypes with therapeutic implications. Science *369*,
1201 eabc8511. 10.1126/science.abc8511.
1202 44. Furman, D., Jojic, V., Sharma, S., Shen-Orr, S.S., Angel, C.J.L., Onengut-Gumuscu, S.,
1203 Kidd, B.A., Maecker, H.T., Concannon, P., Dekker, C.L., et al. (2015). Cytomegalovirus
1204 infection enhances the immune response to influenza. Sci. Transl. Med. *7*, 281ra43-
1205 281ra43. 10.1126/scitranslmed.aaa2293.
1206 45. Mettelman, R.C., and Thomas, P.G. (2020). Human Susceptibility to Influenza Infection and
1207 Severe Disease. Cold Spring Harb. Perspect. Med., a038711.
1208 10.1101/cshperspect.a038711.
1209 46. Wikby, A., Johansson, B., Olsson, J., Löfgren, S., Nilsson, B.-O., and Ferguson, F. (2002).
1210 Expansions of peripheral blood CD8 T-lymphocyte subpopulations and an association with
1211 cytomegalovirus seropositivity in the elderly: the Swedish NONA immune study. Exp.
1212 Gerontol. *37*, 445–453. 10.1016/S0531-5565(01)00212-1.
1213 47. Savva, G.M., Pachnio, A., Kaul, B., Morgan, K., Huppert, F.A., Brayne, C., and Moss, P.A.H.
1214 (2013). Cytomegalovirus infection is associated with increased mortality in the older
1215 population. Aging Cell *12*, 381–387. 10.1111/acel.12059.
1216 48. Wang, G.C., Kao, W.H.L., Murakami, P., Xue, Q.-L., Chiou, R.B., Detrick, B., McDyer, J.F.,
1217 Semba, R.D., Casolaro, V., Walston, J.D., et al. (2010). Cytomegalovirus Infection and the
1218 Risk of Mortality and Frailty in Older Women: A Prospective Observational Cohort Study.
1219 Am. J. Epidemiol. *171*, 1144–1152. 10.1093/aje/kwq062.
1220 49. Netea, M.G., Joosten, L.A.B., Latz, E., Mills, K.H.G., Natoli, G., Stunnenberg, H.G., O'Neill,
1221 L.A.J., and Xavier, R.J. (2016). Trained immunity: A program of innate immune memory in
1222 health and disease. Science *352*. 10.1126/science.aaf1098.
1223 50. Tu, W., and Rao, S. (2016). Mechanisms Underlying T Cell Immunosenescence: Aging and
1224 Cytomegalovirus Infection. Front. Microbiol. *7*.
1225 51. Souquette, A., Frere, J., Smithey, M., Sauce, D., and Thomas, P.G. (2017). A constant
1226 companion: immune recognition and response to cytomegalovirus with aging and
1227 implications for immune fitness. GeroScience *39*, 293–303. 10.1007/s11357-017-9982-x.
1228 52. Jergović, M., Contreras, N.A., and Nikolich-Žugich, J. (2019). Impact of CMV upon immune
1229 aging: facts and fiction. Med. Microbiol. Immunol. (Berl.) *208*, 263–269. 10.1007/s00430-
1230 019-00605-w.
1231 53. Liston, A., Carr, E.J., and Linterman, M.A. (2016). Shaping Variation in the Human Immune
1232 System. Trends Immunol. *37*, 637–646. 10.1016/j.it.2016.08.002.
1233 54. Evans, D.M., Frazer, I.H., and Martin, N.G. (1999). Genetic and environmental causes of
1234 variation in basal levels of blood cells. Twin Res. Hum. Genet. *2*, 250–257.
1235 10.1375/twin.2.4.250.
1236 55. Roederer, M., Quaye, L., Mangino, M., Beddall, M.H., Mahnke, Y., Chattopadhyay, P., Tosi,
1237 I., Napolitano, L., Terranova Barberio, M., Menni, C., et al. (2015). The Genetic Architecture
1238 of the Human Immune System: A Bioresource for Autoimmunity and Disease Pathogenesis.
1239 Cell *161*, 387–403. 10.1016/j.cell.2015.02.046.
1240 56. Lakshmikanth, T., Muhammad, S.A., Olin, A., Chen, Y., Mikes, J., Fagerberg, L.,
1241 Gummesson, A., Bergström, G., Uhlen, M., and Brodin, P. (2020). Human Immune System
1242 Variation during 1 Year. Cell Rep. *32*, 107923. 10.1016/j.celrep.2020.107923.
1243 57. Cook, C.H., Trgovcich, J., Zimmerman, P.D., Zhang, Y., and Sedmak, D.D. (2006).
1244 Lipopolysaccharide, Tumor Necrosis Factor Alpha, or Interleukin-1β Triggers Reactivation of

Latent Cytomegalovirus in Immunocompetent Mice. J. Virol. *80*, 9151–9158. 10.1128/JVI.00216-06.

58. Sinclair, J. (2008). Human cytomegalovirus: Latency and reactivation in the myeloid lineage. J. Clin. Virol. *41*, 180–185. 10.1016/j.jcv.2007.11.014.
59. Pino-Yanes, M., Gignoux, C.R., Galanter, J.M., Levin, A.M., Campbell, C.D., Eng, C., Huntsman, S., Nishimura, K.K., Gourraud, P.-A., Mohajeri, K., et al. (2015). Genome-wide association study and admixture mapping reveal new loci associated with total IgE levels in Latinos. J. Allergy Clin. Immunol. *135*, 1502–1510. 10.1016/j.jaci.2014.10.033.
60. Peng, K., Safonova, Y., Shugay, M., Popejoy, A.B., Rodriguez, O.L., Breden, F., Brodin, P., Burkhardt, A.M., Bustamante, C., Cao-Lormeau, V.-M., et al. (2021). Diversity in immunogenomics: the value and the challenge. Nat. Methods *18*, 588–591. 10.1038/s41592-021-01169-5.
61. Rosenberg-Hasson, Y., Hansmann, L., Liedtke, M., Herschmann, I., and Maecker, H.T. (2014). Effects of serum and plasma matrices on multiplex immunoassays. Immunol. Res. *58*, 224–233. 10.1007/s12026-014-8491-6.
62. Fink, A.L., Liu, H., Shaw-Saliba, K., Mehoke, T., Evans, J., Liu, Z.-Y., Lewis, M., Sauer, L., Thielen, P., Chen, K.-F., et al. (2021). Regional differences in vaccine uptake and serological responses to vaccine and circulating strains of H1N1 viruses among patients with confirmed influenza. J. Clin. Virol. Plus *1*, 100034. 10.1016/j.jcvp.2021.100034.
63. Zapf, A.J., Hardick, J., McBryde, B., Sauer, L.M., Fenstermacher, K.Z.J., Ricketts, E.P., Lin, Y.-C., Chen, K.-F., Hsieh, Y.-H., Dugas, A., et al. (2022). Impact of coinfection status and comorbidity on disease severity in adult emergency department patients with influenza B. Influenza Other Respir. Viruses *16*, 236–246. 10.1111/irv.12907.
64. R Core Team (2021). R: A Language and Environment for Statistical Computing.
65. Krijthe, J.H. (2015). Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut.
66. Kolde, R. (2019). pheatmap: Pretty Heatmaps.
67. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. 10.21105/joss.01686.
68. David Robinson, Alex Hayes, and Simon Couch (2021). broom: Convert Statistical Objects into Tidy Tibbles.
69. Kun Ren and Kenton Russell (2021). formattable: Create “Formattable” Data Structures.
70. Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., Toomet, O., Crowley, J., Hofmann, H., et al. (2021). GGally: Extension to “ggplot2.”
71. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. *57*, 289–300.

Study n = 1,705	Unit/Key	Memphis n = 383	Nicaragua n = 591	Baltimore n = 225	Taiwan n = 266	Colombia n = 200	Belgium n = 40
Age	Median (Range)	17 (0-71)	11 (0-84)	43 (18-89)	41 (20-92)	13 (2-77)	42 (22-53)
Sex	% Female	61.4	55.7	56.0	57.9	56.5	85
Subjects Included	Healthy; Influenza Infected						
Sample Types	Nasal Wash; Plasma; PBMC						
Timepoints	Single or Multiple	Multiple	Single	Single	Single	Single	Single
Symptoms	Continuous (Scale) or Binary (presence/ absence)	Continuous	Binary	Binary	Binary	-	-
Analysis	Cytokine Severity Correlates; Baseline Immunity; Acute Immunity						

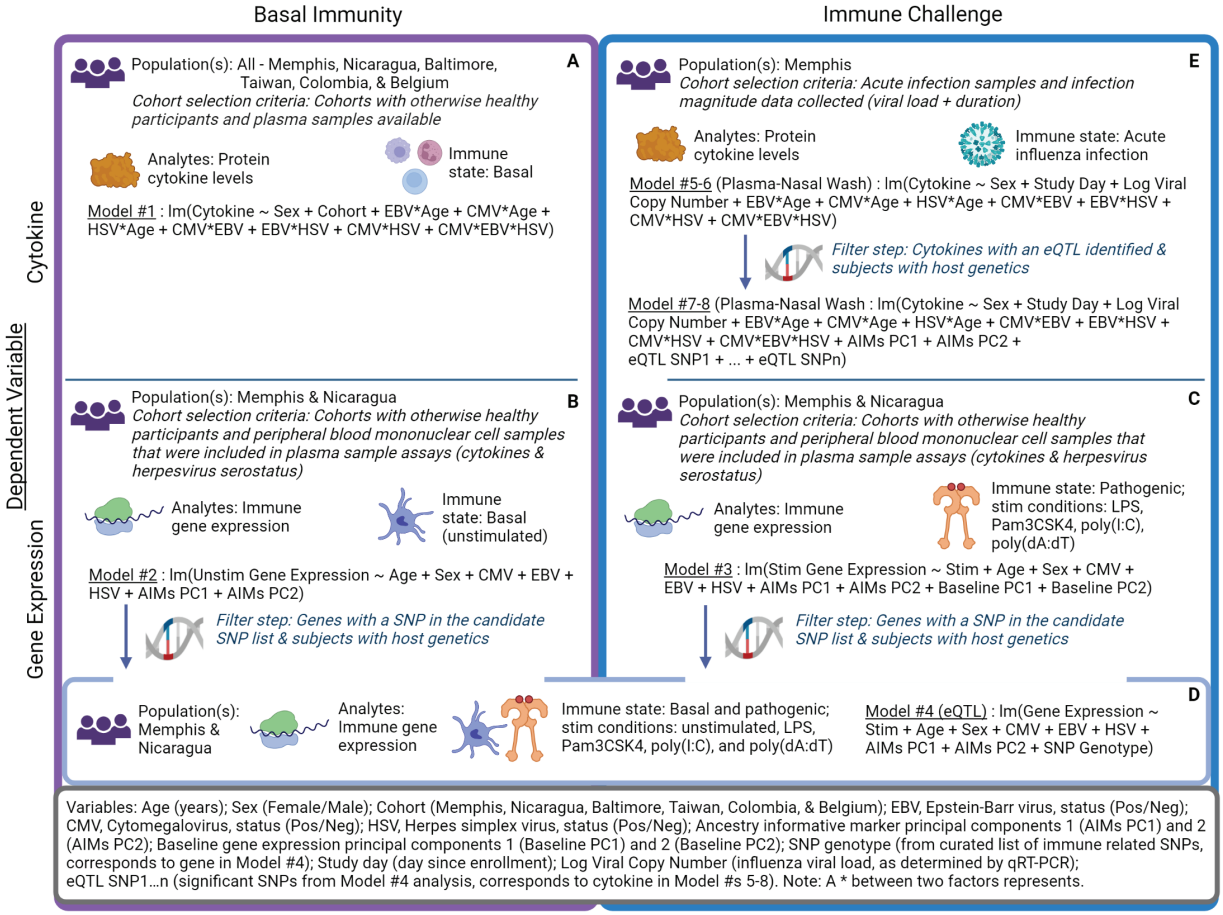
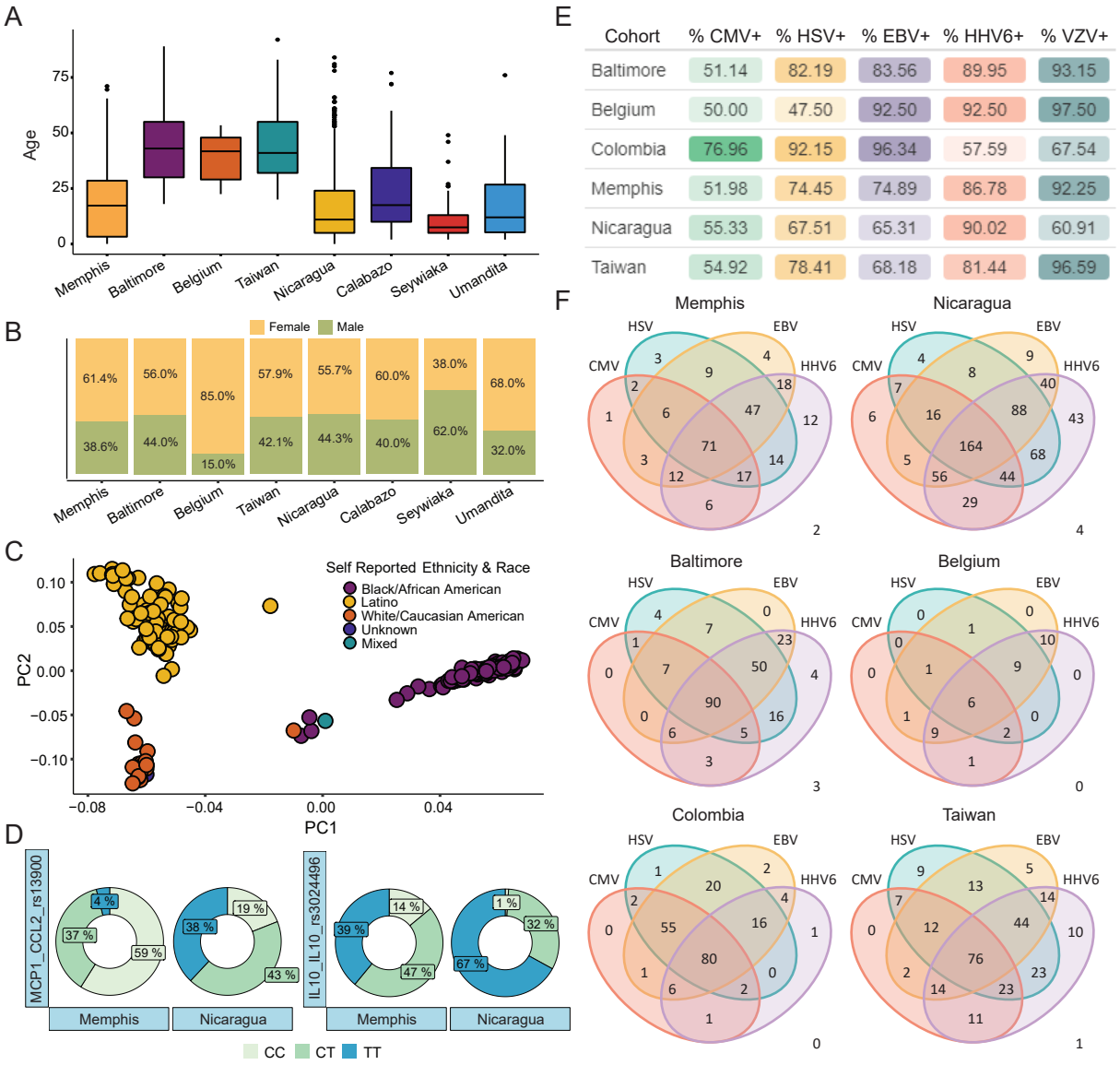


Figure 1. Study populations and model summary. Table of cohort characteristics (top) and overview of regression model analyses (bottom, labeled A-D in order of appearance).



Cohort	% CMV+	% HSV+	% EBV+	% HHV6+	% VZV+
Baltimore	51.14	82.19	83.56	89.95	93.15
Belgium	50.00	47.50	92.50	92.50	97.50
Colombia	76.96	92.15	96.34	57.59	67.54
Memphis	51.98	74.45	74.89	86.78	92.25
Nicaragua	55.33	67.51	65.31	90.02	60.91
Taiwan	54.92	78.41	68.18	81.44	96.59

Figure 2. Variation of explanatory factors across diverse populations. Age (A) and sex (B) distribution for each cohort (n = 1,705). (C) Results from the principal component analysis of AIMs (n = 285). (D) Donut plots show the proportion of individuals in each cohort with a given genotype for a SNP of interest (n = 352), plot labels are encoded as follows: Protein-Gene-SNP. (E) Table of herpesvirus serostatus (%) by cohort (n = 1,532). (F) Venn-diagram of herpesvirus infection by cohort. Numbers in the bottom right represent the number of individuals negative for all indicated herpesviruses. For e sier visualization, the herpesvirus with the highest overall seroprevalence (VZV) was excluded.

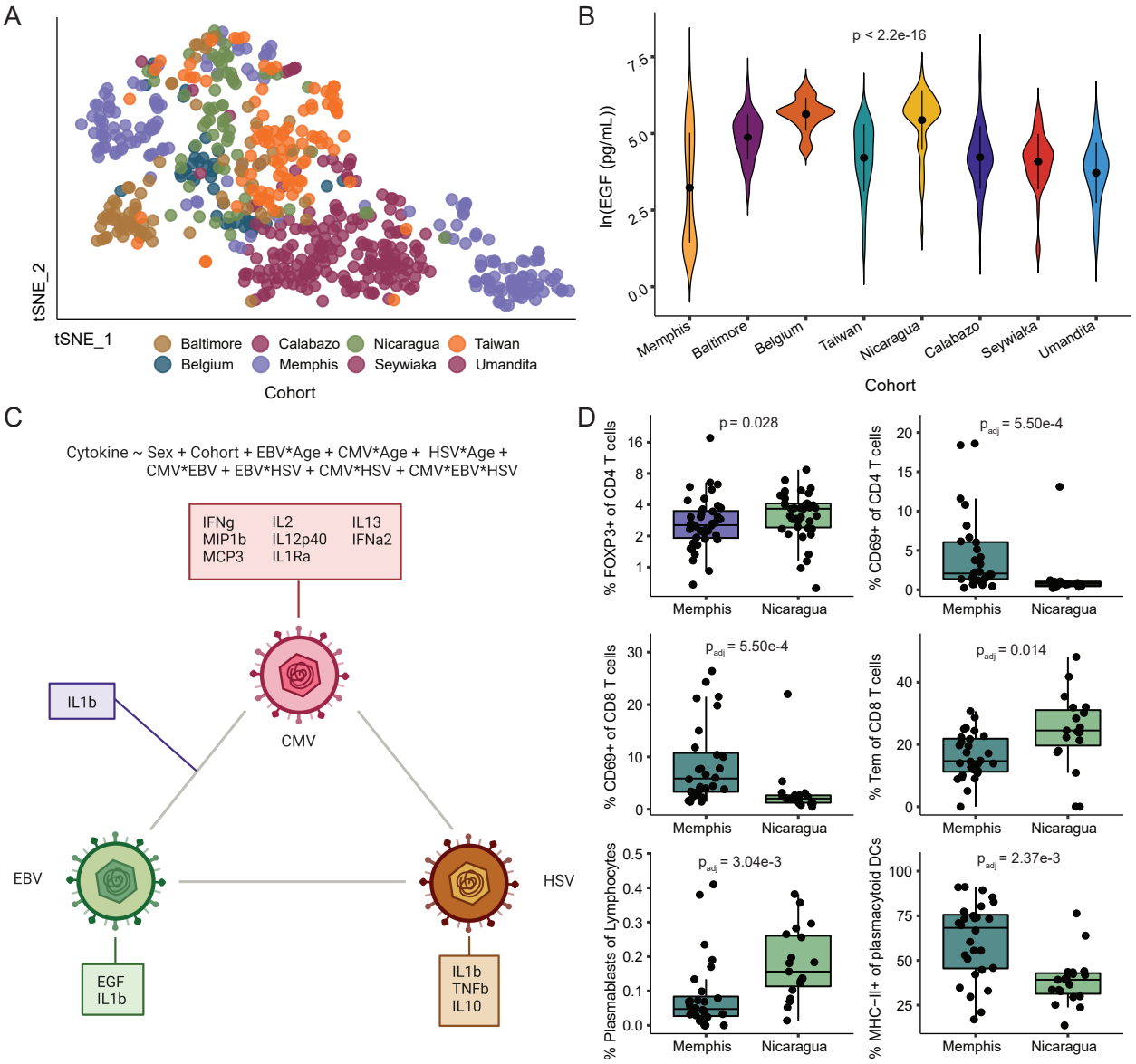
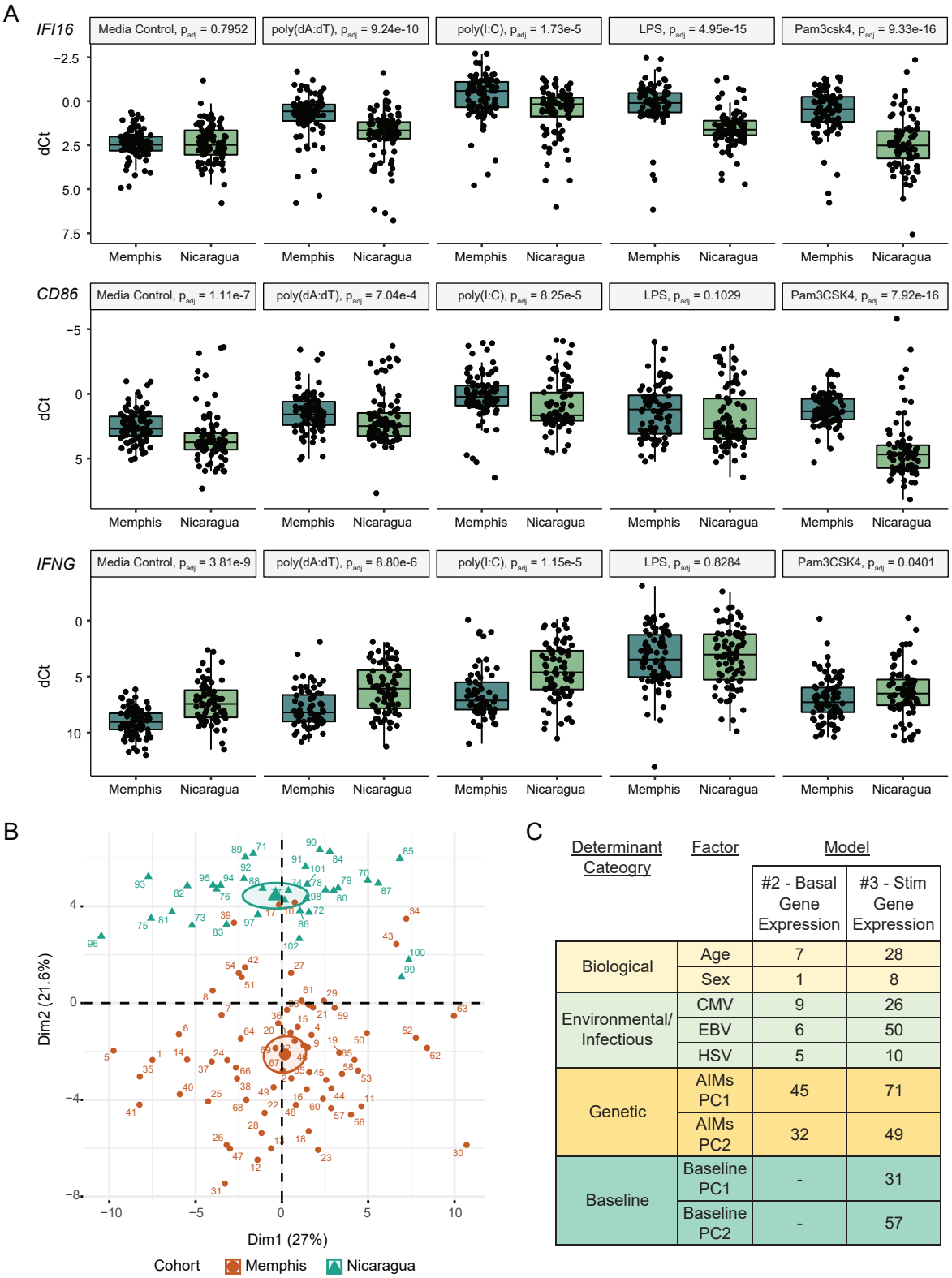


Figure 3. Basal immune profiles in ancestrally and environmentally distinct populations. (A) A t-SNE plot based on plasma/serum cytokines and colored by cohort (n = 631). (B) Violin plot demonstrates individual cytokine variation for EGF (n = 623). Black points in the center of the violin plot represent the mean cytokine level and line segments represent +/- the standard deviation. (C) Schematic of herpesvirus effects on basal cytokine levels based on linear regression modeling results (model displayed above diagram, n ≈ 500). Cytokines listed have an adjusted R^2 value > 0.08 and FDR < 0.05. (D) Boxplots depict differences in basal cell phenotypes in PBMC samples from healthy subjects in the Memphis (n = 28) and Nicaragua (n = 19) cohorts. Presented p-values are from Mann-Whitney U tests, post adjustment for multiple comparisons (when applicable).



C

Determinant Cateogry	Factor	Model: #2 - Basal Gene Expression	Model: #3 - Stim Gene Expression
Biological	Age	7	28
	Sex	1	8
Environmental/ Infectious	CMV	9	26
	EBV	6	50
	HSV	5	10
Genetic	AIMs PC1	45	71
	AIMs PC2	32	49
Baseline	Baseline PC1	-	31
	Baseline PC2	-	57

Figure 4. Distinct basal immune profiles mold response to immune challenge. (A) Gene expression by cohort under different stim conditions for *IFI16* ($n_{AllStim}$ = 903), *CD86* ($n_{AllStim}$ = 906), and *IFNG* ($n_{AllStim}$ = 774). Displayed p-values were adjusted for multiple comparisons and are from Mann Whitney U tests. (B) Principal component analysis of baseline (unstim) gene expression colored by Cohort (n = 102). (C) Summary of linear regression Models #2 (n ≈ 188) and #3 ($n_{AllStim}$ ≈ 350) that were utilized to determine factors which contribute to baseline and post stim gene expression variation, respectively (see Summary of Models section for details). Displayed values are the number of genes that a factor significantly affected out of a total 90 genes.

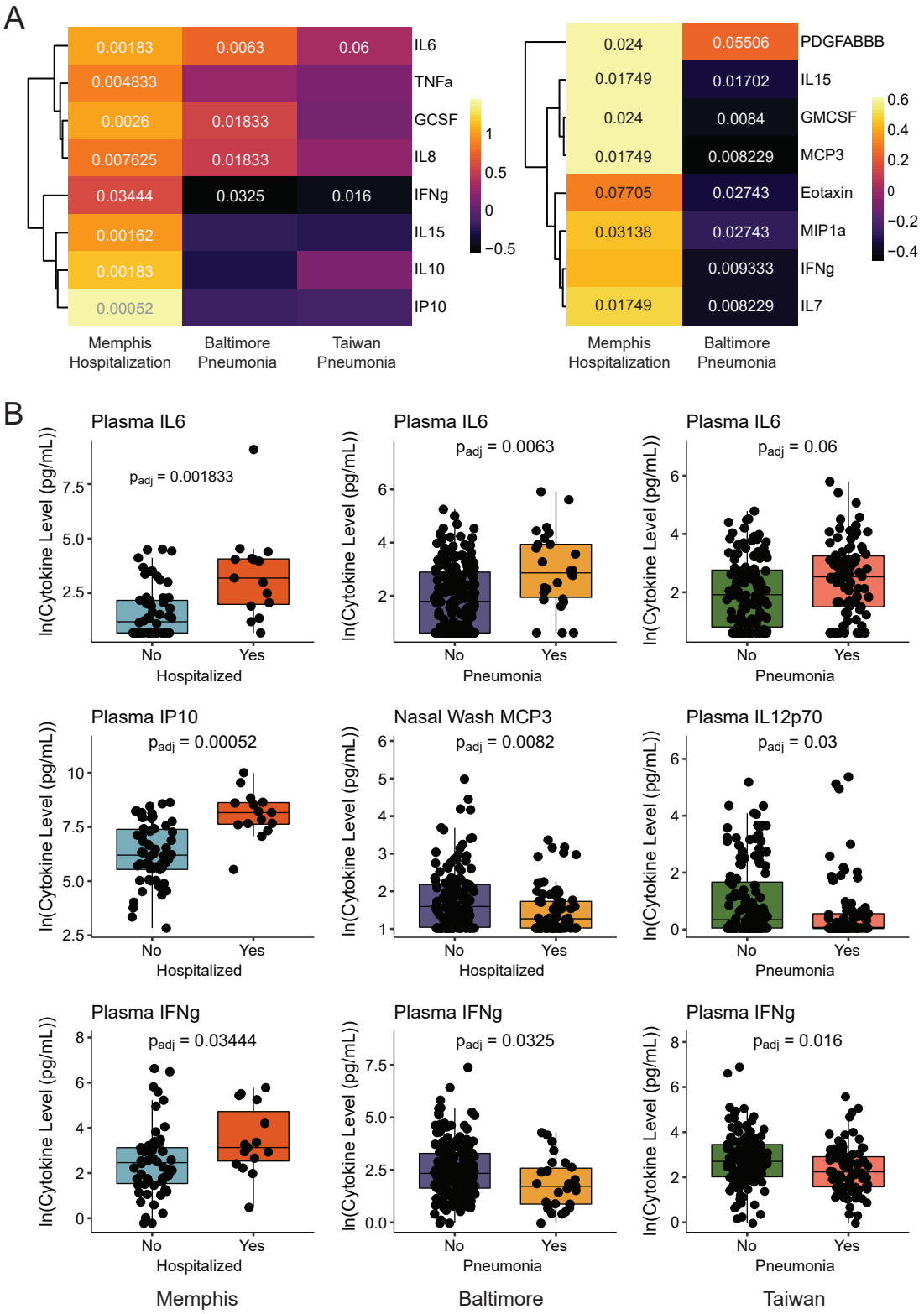


Figure 5. Acute cytokine profiles delineate two age-associated signatures of influenza disease severity. (A) Heatmap of highlighted results from plasma (left) and nasal wash (right) cytokine correlates of severity analysis, with significant adjusted p-values displayed, for the Memphis (n = 165), Baltimore (n = 222), and Taiwan (n = 226) cohorts. (B) Boxplots show example associations between plasma or nasal wash cytokine levels and symptom severity during influenza infection for each cohort. Presented p-values are from Mann Whitney U tests, and results were adjusted for multiple comparisons by controlling the FDR (< 0.1).

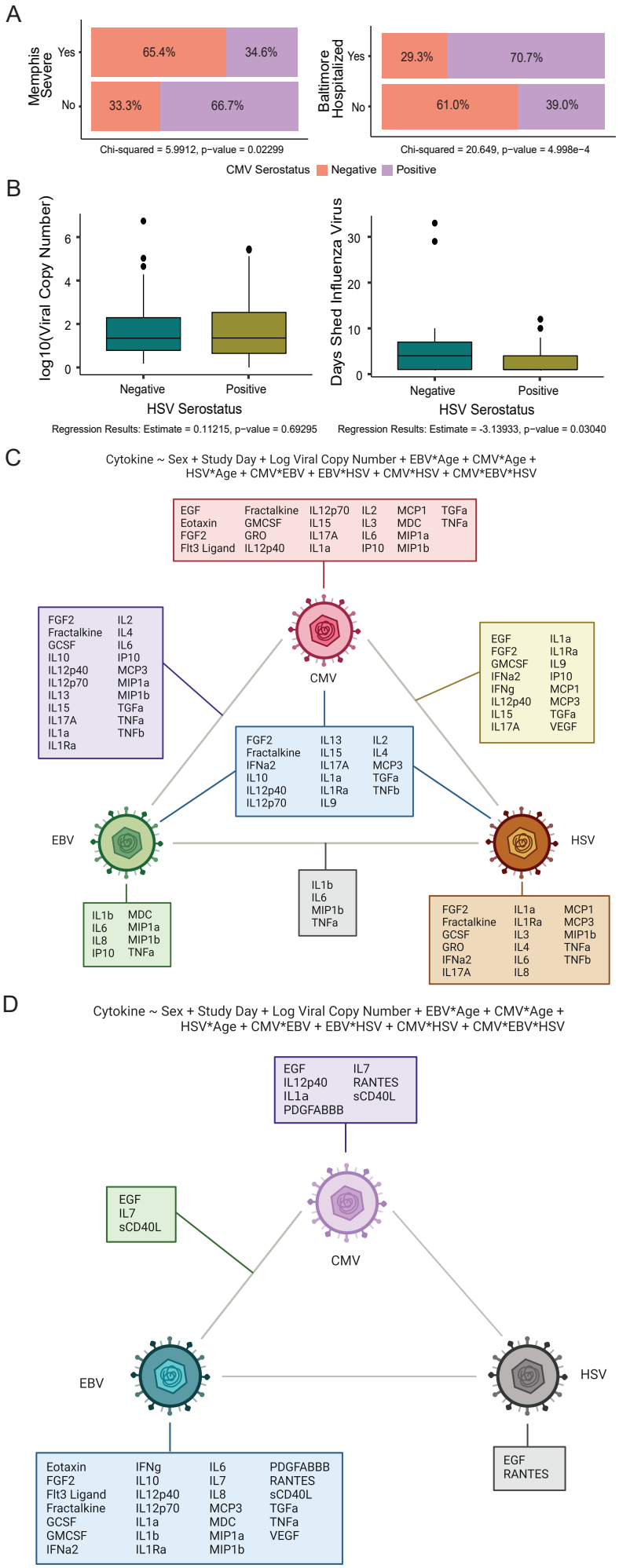


Figure 6. Herpesviruses have unique and age specific effects on influenza disease severity. (A-B) Significant herpesvirus associations with influenza disease severity. Severe Memphis was categorized based on lower respiratory symptom severity scores. (C-D) Schematic of herpesvirus effects on acute cytokine levels in plasma (C, n ≈ 200) and nasal wash (D, n ≈ 300) based on linear regression modeling results (models displayed above diagrams). Cytokines listed have an adjusted R^2 value > 0.08 and FDR < 0.05.

A

lm(Eotaxin ~ Study Day + Influenza Viral Load (Log.VCN) + Sex + EBV*Age + CMV*Age + HSV*Age + CMV*EBV + EBV*HSV + CMV*HSV + CMV*EBV*HSV + PC1 + PC2 + rs5027581)

Factor	Coeff Estimate	Std Error	Statistic	P-value
(Intercept)	↑ 4.56	0.4010	11.400	4.46e-23
EBVPositive	↓ -0.767	0.4080	-1.880	6.19e-02
scale(Age)	↑ 0.464	0.3550	1.310	1.92e-01
CMVPositive	↓ -0.976	0.3980	-2.450	1.52e-02
HSVPositive	↓ -0.654	0.4370	-1.500	1.36e-01
SexMale	↑ 0.525	0.1280	4.100	6.23e-05
scale(Study.Day)	↑ 0.0676	0.0610	1.110	2.69e-01
scale(Log.VCN)	↑ 0.18	0.0618	2.920	3.93e-03
scale(AIMs PC1)	↓ -0.403	0.2910	-1.380	1.68e-01
scale(AIMs PC2)	↓ -0.298	0.2860	-1.040	2.98e-01
rs5027581AT	↓ -0.251	0.1220	-2.060	4.06e-02
rs5027581TT	↑ 0.483	0.1750	2.760	6.32e-03
EBVPositive:scale(Age)	↑ 0.157	0.3460	0.454	6.50e-01
scale(Age):CMVPositive	↓ -0.549	0.1600	-3.430	7.41e-04
scale(Age):HSVPositive	↓ -0.027	0.2080	-0.130	8.97e-01
EBVPositive:CMVPositive	↑ 1.07	0.4340	2.450	1.51e-02
EBVPositive:HSVPositive	↑ 0.319	0.4720	0.676	5.00e-01
CMVPositive:HSVPositive	↑ 1.98	0.7000	2.830	5.18e-03
EBVPositive:CMVPositive:HSVPositive	↓ -2.02	0.7530	-2.680	7.97e-03

Adjusted R^2 = 0.5568

B

lm(IL10 ~ Study Day + Influenza Viral Load (Log.VCN) + Sex + EBV*Age + CMV*Age + HSV*Age + CMV*EBV + EBV*HSV + CMV*HSV + CMV*EBV*HSV + PC1 + PC2 + rs3024496 + rs3024498)

Factor	Coeff Estimate	Std Error	Statistic	P-value
(Intercept)	↑ 0.971	0.4150	2.340	2.02e-02
EBVPositive	↓ -0.13	0.3770	-0.346	7.30e-01
scale(Age)	↓ -0.751	0.3230	-2.320	2.13e-02
CMVPositive	↓ -0.432	0.3390	-1.280	2.04e-01
HSVPositive	↑ 0.171	0.3900	0.438	6.62e-01
SexMale	↑ 0.343	0.1240	2.770	6.21e-03
scale(Study.Day)	↓ -0.0774	0.0576	-1.340	1.80e-01
scale(Log.VCN)	↑ 0.396	0.0578	6.860	9.42e-11
scale(AIMs PC1)	↓ -0.232	0.2760	-0.841	4.02e-01
scale(AIMs PC2)	↑ 0.233	0.2760	0.846	3.99e-01
rs3024496CT	↓ -0.331	0.1690	-1.960	5.17e-02
rs3024496TT	↓ -0.683	0.1690	-4.050	7.49e-05
rs3024498AG	↑ 0.0476	0.1850	0.258	7.97e-01
rs3024498GG	↓ -0.45	0.3330	-1.350	1.79e-01
EBVPositive:scale(Age)	↑ 0.048	0.3140	0.153	8.79e-01
scale(Age):CMVPositive	↑ 0.0979	0.1500	0.652	5.15e-01
scale(Age):HSVPositive	↑ 0.504	0.1990	2.530	1.21e-02
EBVPositive:CMVPositive	↑ 1.33	0.3730	3.560	4.65e-04
EBVPositive:HSVPositive	↑ 0.557	0.4150	1.340	1.81e-01
CMVPositive:HSVPositive	↑ 0.436	0.6500	0.670	5.03e-01
EBVPositive:CMVPositive:HSVPositive	↓ -1.28	0.6820	-1.870	6.29e-02

Adjusted R^2 = 0.3915

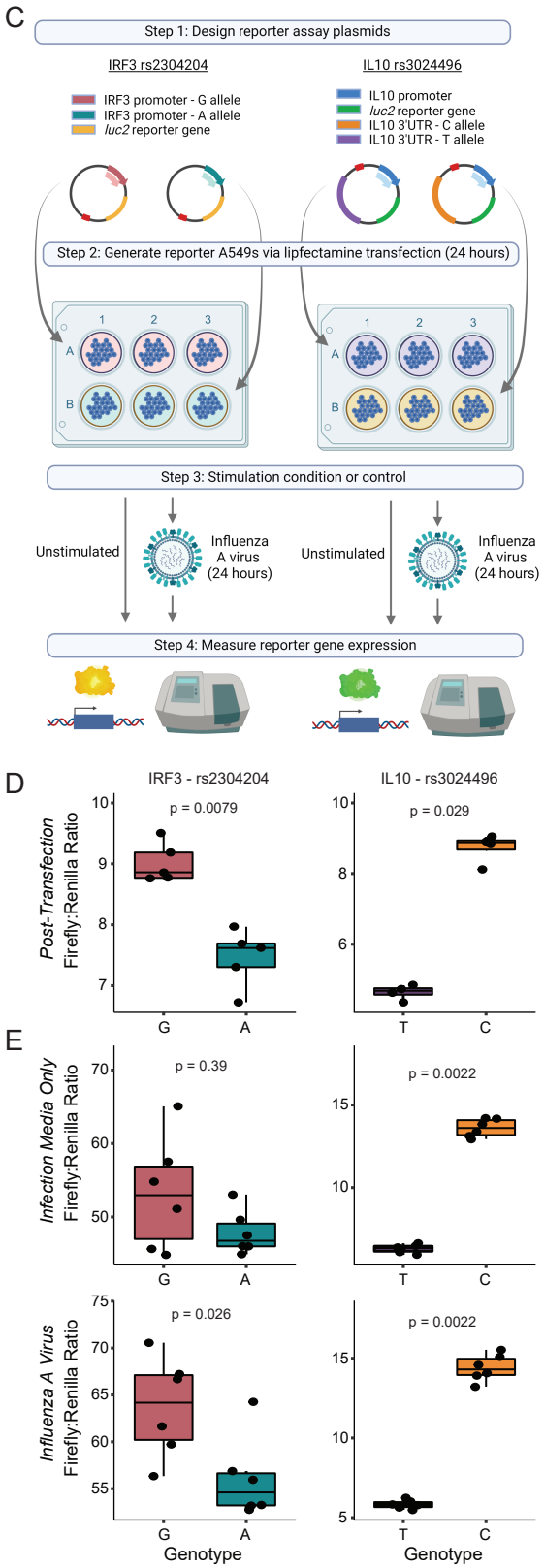
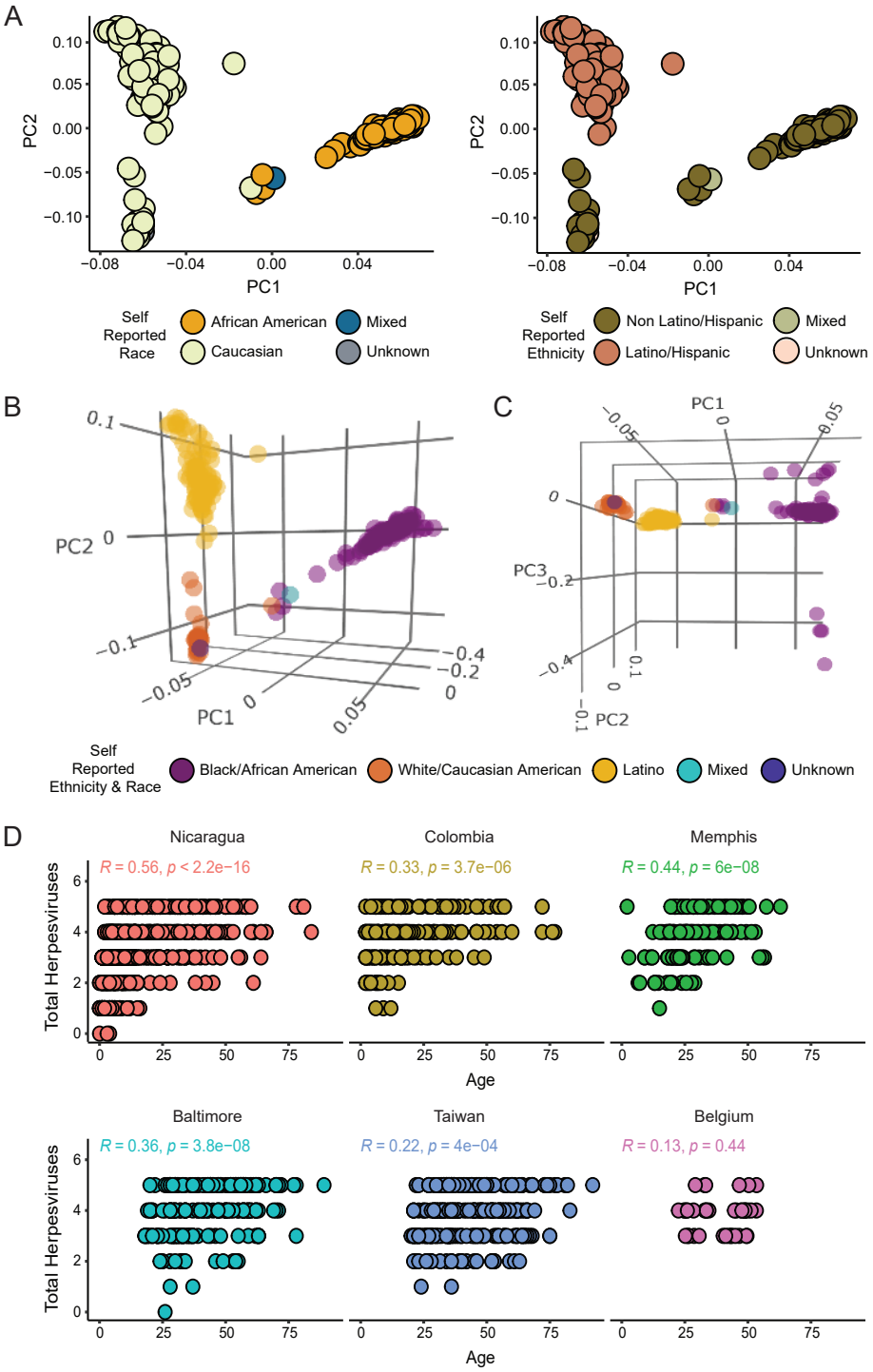
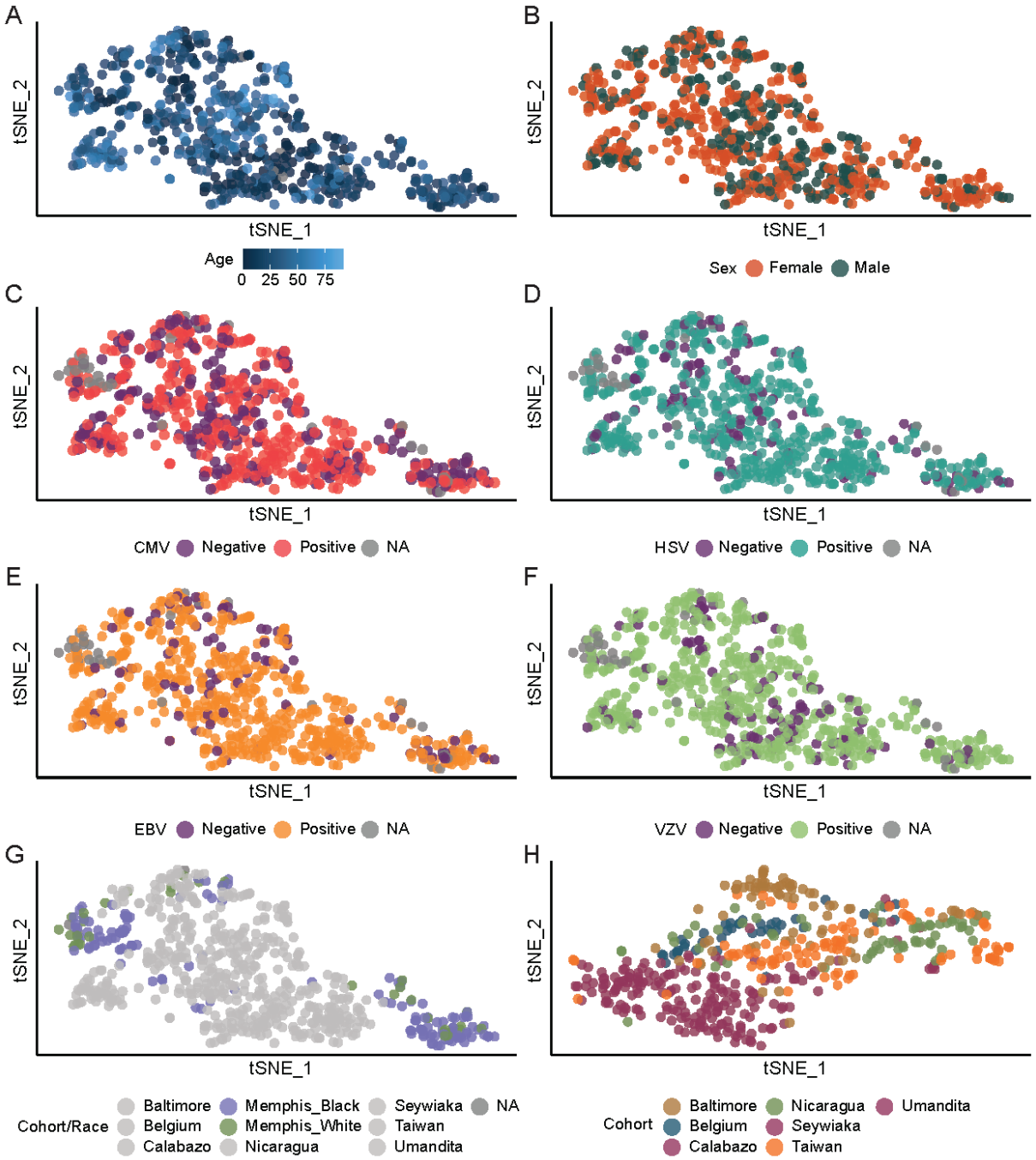


Figure 7. Transcription and protein level of IL10 is modulated by rs3024496. (A-B) For cytokines with an eQTL, a linear regression model was performed including: sex, age, herpesvirus (HV) serostatus, interactions between HV serostatus and age, interactions between HVs, study day, viral load, AIMs PC1 and PC2, and eQTL SNP genotype (n ≈ 200). The top two results, based on adjusted R^2 value, are shown and include the coefficient estimate of the change in the dependent variable per unit change in the independent variable (Coeff Estimate), the corresponding standard error (Std Error), t score (Statistic), and p-value (highlighted when significant). (C) Schematic of luciferase assays. Boxplots compare Firefly:Renilla ratios from unstimulated (D) or influenza virus infected (E) luciferase assays for A and G alleles of IRF3 rs2304204 (left), and T and C alleles of IL10 rs3024496 (right). Presented p-values are from Mann-Whitney U tests.



Supplemental Figure 1. Explanatory variables by Race/Ethnicity and age. (A-C) Principal component analysis of AIMs colored by Race and/or Ethnicity. (D) Total number of herpesviruses by age, separated by cohort.



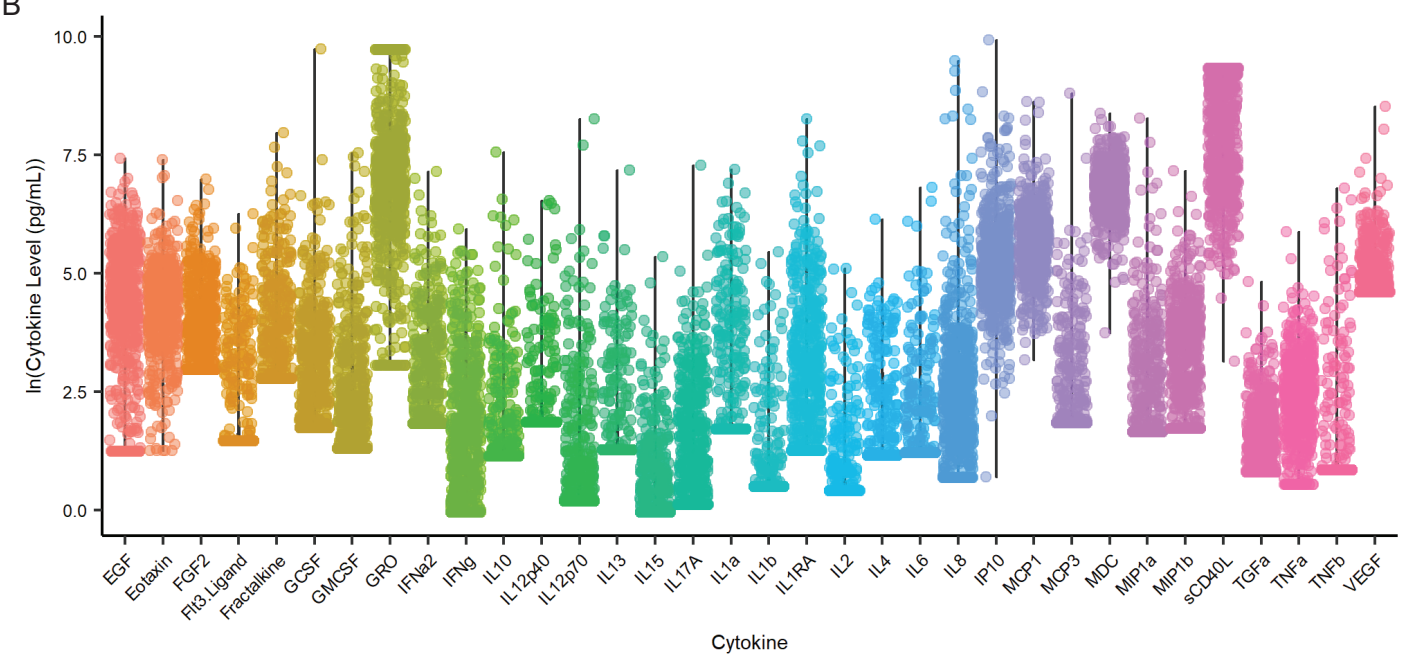
Supplemental Figure 2. Cohort tSNE clusters are not associated with demographics or chronic virus infection. Basal cytokine tSNE plot colored by age (A), sex (B), herpesvirus serostatus (C-F), and Race (G, for Memphis only, all other cohorts in gray). (H) Basal cytokine tSNE with Memphis cohort data removed.

A

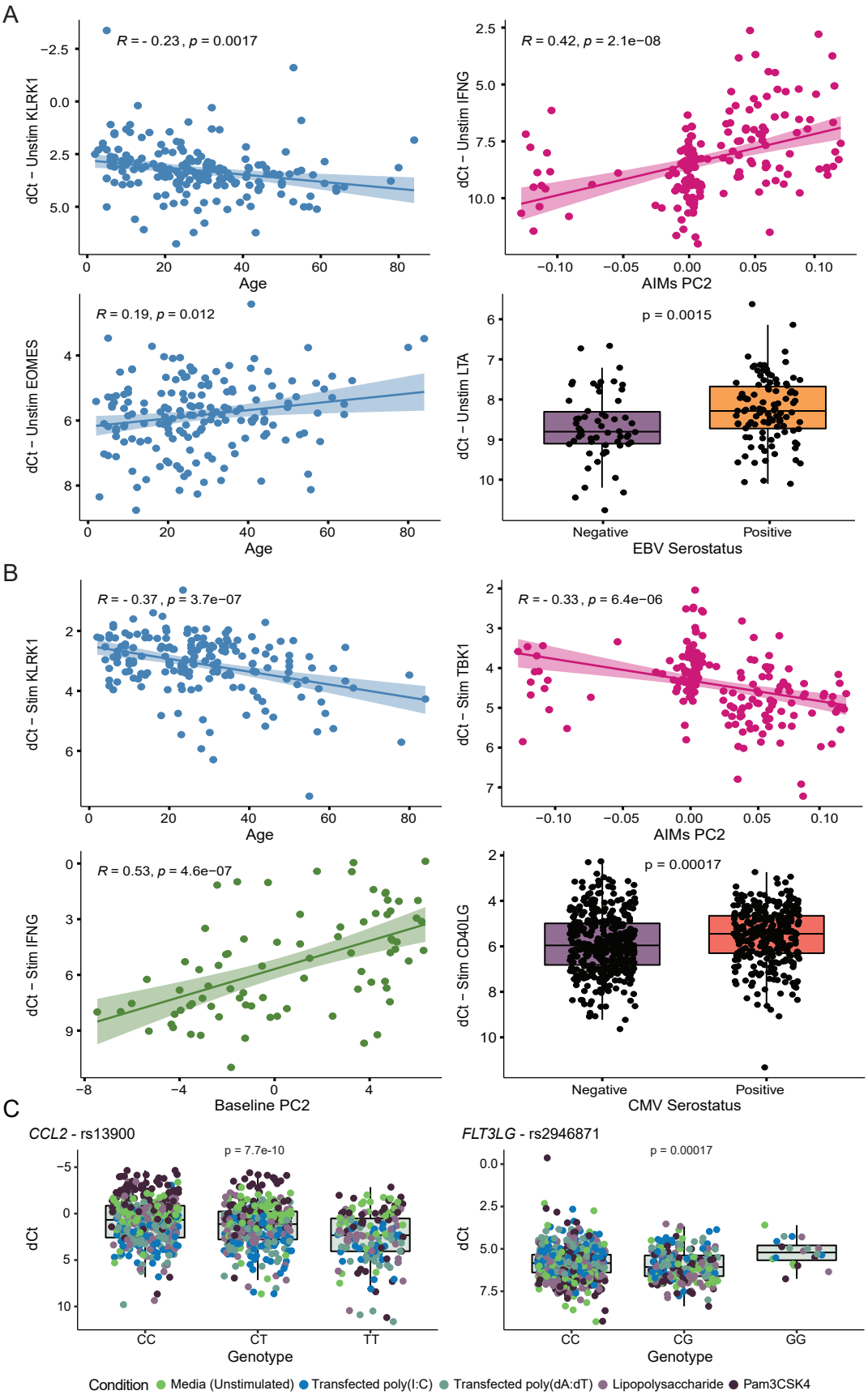
Cytokine	Age	Sex.Male	CMVPos	Age: CMVPos	HSVPos	Age: HSVPos	EBVPos	Age: EBVPos	EBVPos: CMVPos	CMVPos: HSVPos	EBVPos: HSVPos	EBVPos: CMVPos: HSVPos	Cohort Belgium	Cohort Calabazo	Cohort Memphis	Cohort Nicaragua	Cohort Seywiaka	Cohort Taiwan	Cohort Umandita	Adj.R2
IL4	-0.1088	-0.1187	0.0732	-0.0120	0.0176	0.0470	0.2812	0.0541	-0.4330	0.1111	-0.2029	0.1721	-1.8587	-1.8428	-1.7906	-1.8317	-1.9067	-1.2541	-1.9483	0.4709
GRO	-0.6083	-0.1828	0.5188	0.0176	0.6678	0.1884	0.9633	0.3739	-0.6153	-0.1712	-0.8587	0.4887	-1.2974	-1.4589	-3.3088	-1.4349	-2.4034	-1.2187	-2.6628	0.4569
sCD40L	-0.1498	0.0320	0.2733	-0.0063	0.4675	-0.0766	0.3363	0.1580	-0.0530	-0.2256	-0.4437	0.0684	0.0267	-1.8331	-1.3420	-0.1152	-2.1052	-0.5331	-1.8609	0.3721
EGF	-0.3680	-0.0638	0.1286	-0.1198	-0.0218	-0.0842	0.1586	0.4614	0.1728	-0.2553	-0.1412	0.1270	0.5745	-0.8113	-1.8451	0.3987	-1.0405	-0.6201	-1.3360	0.3222
TNFa	-0.1924	0.1087	-0.3801	-0.1434	-0.2171	0.2375	-0.0024	0.0227	0.3725	0.4636	-0.0266	-0.3256	-0.3021	-0.6190	-0.8691	0.3070	0.0880	0.4516	0.0620	0.3145
MIP1b	0.0118	-0.0054	0.1301	-0.2592	-0.1558	0.0151	0.2618	0.1504	-0.0750	0.0339	-0.1831	-0.0257	-0.7073	-1.3946	-0.3720	0.2491	-1.0017	0.2679	-1.1129	0.2908
GCSF	0.2368	-0.1938	-0.3033	-0.1047	-0.3581	0.0370	-0.4705	-0.1489	0.3321	0.3941	0.4278	-0.4806	-1.3700	-1.2568	-0.6437	0.3056	-1.2812	0.1689	-1.0055	0.2707
FGF2	0.0477	-0.1376	-0.1519	-0.1824	-0.4421	0.0613	-0.3831	0.0135	0.1077	0.0934	0.3625	-0.1602	-0.4969	-0.5252	-0.5932	0.7341	-0.1671	0.2722	-0.4174	0.2456
IL15	0.1659	0.0308	0.1109	-0.1188	-0.2467	0.0471	-0.2469	-0.0874	-0.3082	-0.1951	0.1223	0.4140	-0.5866	-0.6610	-0.4542	-0.0102	-0.6615	0.3304	-0.6074	0.2446
IFNg	0.5748	-0.0621	-0.0782	-0.3273	-0.5823	-0.2260	-0.5147	-0.1701	-0.0694	-0.2742	0.4063	0.3563	-0.0580	-0.9925	0.4446	0.3864	-1.0031	0.7206	-1.0720	0.2373
Eotaxin	0.4972	0.0587	0.0545	-0.1183	0.0555	-0.0861	0.2356	-0.2032	0.1275	-0.2862	-0.1655	0.2181	0.0882	0.0130	-0.7203	0.1396	0.2171	0.0646	0.3574	0.2262
MCP1	0.0113	0.0475	-0.4893	-0.1279	-0.2861	0.0624	-0.2374	0.1047	0.6307	0.5500	0.1808	-0.6520	0.1039	-0.3207	-0.6930	0.1921	-0.0696	0.0964	0.0760	0.2251
IL10	-0.0720	0.1462	-0.1937	-0.0798	-0.0833	0.4251	-0.3078	-0.2187	0.1091	0.4206	0.0041	-0.2021	0.1391	0.0583	0.0943	0.5325	0.3709	0.9496	0.2144	0.2187
MIP1a	-0.1897	0.0489	-0.2085	-0.1010	0.2399	0.0405	0.3956	0.2503	0.2535	0.0191	-0.4526	0.0301	0.0186	-0.8508	0.4048	0.2604	-0.8784	-0.6070	-0.8068	0.2172
GMCSF	-0.0420	-0.1299	-0.3589	-0.2353	-0.2667	0.1421	-0.1504	0.0603	0.6074	0.2353	0.2579	-0.4960	-0.5482	-0.7287	0.5406	0.6355	-0.5638	0.2539	-0.6541	0.1934
TGFa	0.1096	-0.2056	-0.2944	-0.0777	-0.2855	-0.0444	-0.1424	-0.0366	0.2444	0.3439	0.2182	-0.3986	0.4111	-0.2401	-0.1366	0.8170	0.2590	-0.0253	-0.1227	0.1914
MDC	-0.1701	0.0475	-0.0030	-0.0357	0.0114	0.0750	0.0614	0.0291	0.0310	0.0788	0.0274	-0.0529	-0.5556	-0.6780	-0.3266	-0.3886	-0.8918	-0.4388	-0.8713	0.1854
IL1RA	-0.1225	-0.2492	-0.3479	-0.3561	-0.1754	0.0903	-0.2355	0.1782	0.1426	0.6326	-0.0311	-0.3671	-0.8538	0.7046	-0.5987	1.1648	0.0248	0.7523	-0.4907	0.1841
IFNa2	-0.1405	-0.0134	-0.6085	-0.2437	-0.7392	0.2623	-0.6239	0.0823	0.6951	0.7066	0.6490	-0.7791	-0.5000	-0.5509	0.0846	0.9336	-0.5658	0.2821	-0.2326	0.1797
Fractalkine	-0.1954	-0.0797	0.3100	-0.1096	0.0769	0.2075	-0.0076	-0.0092	-0.1755	-0.3041	-0.0156	0.0794	-0.3072	-0.6019	-0.1779	0.7816	-0.7219	-0.1262	-0.4398	0.1509
Flt3.Ligand	0.0986	0.0026	0.4639	-0.1211	-0.4175	-0.0003	-0.5485	0.0870	-0.2417	-0.4601	0.5555	0.2208	0.0102	-0.5042	-0.2653	0.7390	-0.2370	-0.2853	-0.3680	0.1497
IL17A	0.4172	-0.0808	0.2671	-0.1709	-0.3182	-0.0586	-0.4994	-0.1961	-0.3258	-0.8089	0.2834	0.7680	0.2855	-0.6726	-0.2004	0.6091	-0.7006	0.2974	-0.6203	0.1465
VEGF	0.0869	-0.0136	0.4686	-0.1109	-0.0462	0.0133	0.0006	0.0578	-0.4046	-0.4549	0.0484	0.3181	-0.0270	-0.0227	0.1478	0.6481	0.0207	0.0315	-0.0022	0.1461
IL1a	0.2248	0.1829	-1.1481	-0.0985	-0.7966	-0.0128	-0.8889	-0.1142	0.9937	1.1015	0.6754	-0.9216	0.2377	-0.0466	0.1490	0.9348	-0.1420	0.9872	0.1367	0.1233
MCP3	-0.0118	-0.0920	-0.5290	-0.2323	-0.6159	0.1024	-0.4191	0.0784	0.7628	0.6483	0.4008	-0.7783	-0.3575	-0.1492	0.3900	0.7740	0.1501	0.1652	-0.1290	0.1125
IL12p70	0.5703	0.0794	0.2996	-0.1585	-0.4564	-0.1648	-0.5545	-0.2743	-0.3378	-0.6801	0.2929	0.7445	-0.1870	-0.4830	0.3366	0.3190	-0.5343	-0.0240	-0.4817	0.1036
IL8	-0.1511	-0.0145	-0.0129	-0.1972	0.6285	0.0617	0.5794	0.3017	0.1426	-0.1806	-1.0008	0.1956	0.7903	-0.6442	0.6688	0.5908	0.1193	0.0702	-0.2125	0.1002
IP10	-0.1878	0.0471	-0.0404	-0.1303	0.0631	0.2009	-0.1414	0.1662	0.0457	0.1263	-0.1158	0.1388	-0.1655	0.4218	0.3975	0.2474	0.8169	0.6865	0.4257	0.0840
IL6	0.1951	0.0601	0.2211	-0.0953	-0.2674	0.0235	-0.0431	-0.1372	-0.2875	-0.0399	0.1320	0.2065	-0.2199	-0.2561	0.3124	-0.0198	-0.2808	0.2884	-0.2456	0.0777
IL1b	0.3180	-0.0189	-0.7883	-0.1694	-0.7561	-0.0158	-0.7097	-0.1642	0.8976	0.5514	0.7211	-0.6407	-0.0731	-0.0862	0.4355	0.3602	-0.0538	0.0319	-0.0863	0.0771
IL12p40	0.1574	0.0094	-0.7549	-0.2386	-0.4214	0.1237	-0.5394	-0.0696	0.8085	0.5738	0.3174	-0.5956	-0.3054	-0.2273	0.0002	0.3845	-0.0974	-0.1160	-0.0919	0.0693
TNFb	0.0286	-0.1087	-0.5919	-0.1881	-0.9711	0.0448	-0.7963	0.0793	0.5967	0.8155	0.8129	-0.7705	-0.4168	-0.2177	-0.1584	0.5117	-0.0477	0.1885	-0.3609	0.0621
IL13	0.1430	-0.0892	-0.3477	-0.1975	-0.6860	-0.0172	-0.6252	0.0084	0.4115	0.3556	0.6543	-0.4551	-0.3710	-0.2655	0.0542	0.2374	-0.1504	0.0387	-0.3375	0.0460
IL2	0.2218	-0.0281	-0.5328	-0.1709	-0.4215	0.0323	-0.6362	-0.0919	0.5919	0.3098	0.4380	-0.4106	-0.1207	-0.1569	-0.0214	0.4073	-0.1495	0.0150	-0.0307	0.0449

Coefficient Scale: -3.5 -2.5 -1.5 -0.5 0 0.5 1.5 2.5 3.5

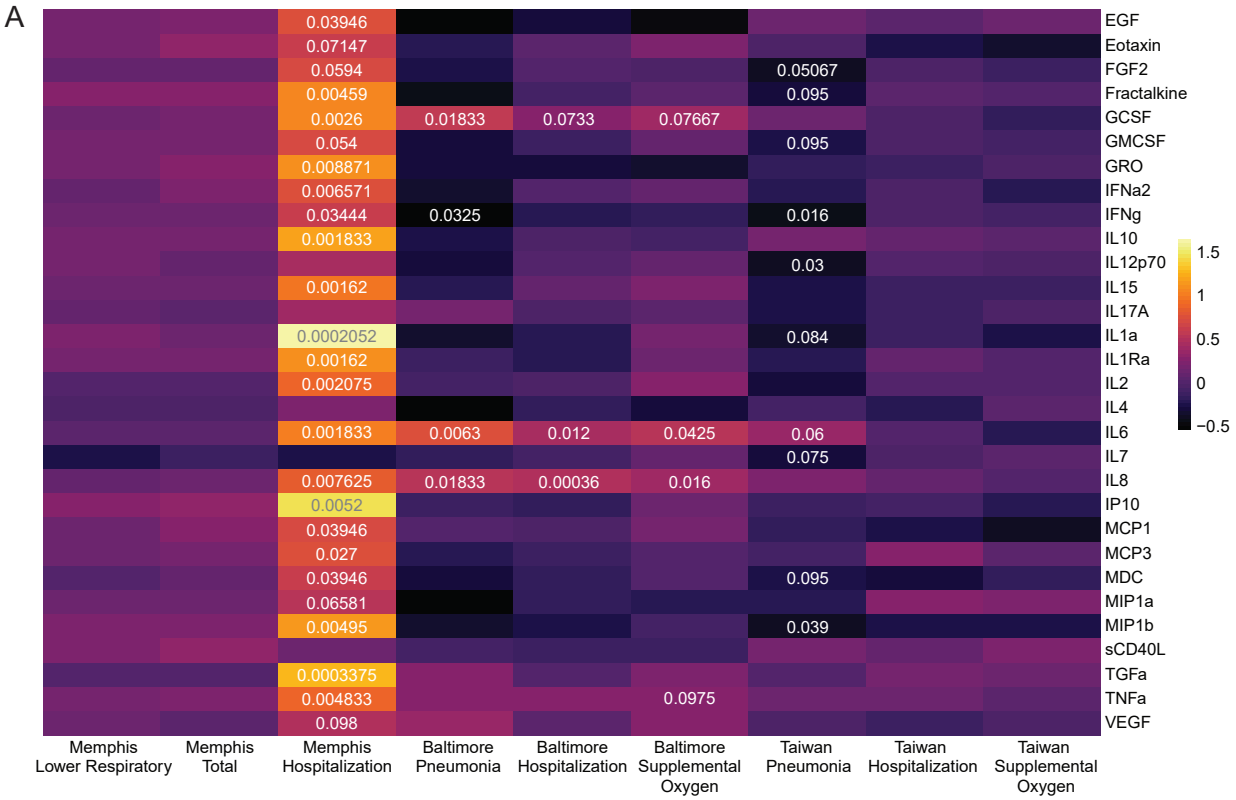
B

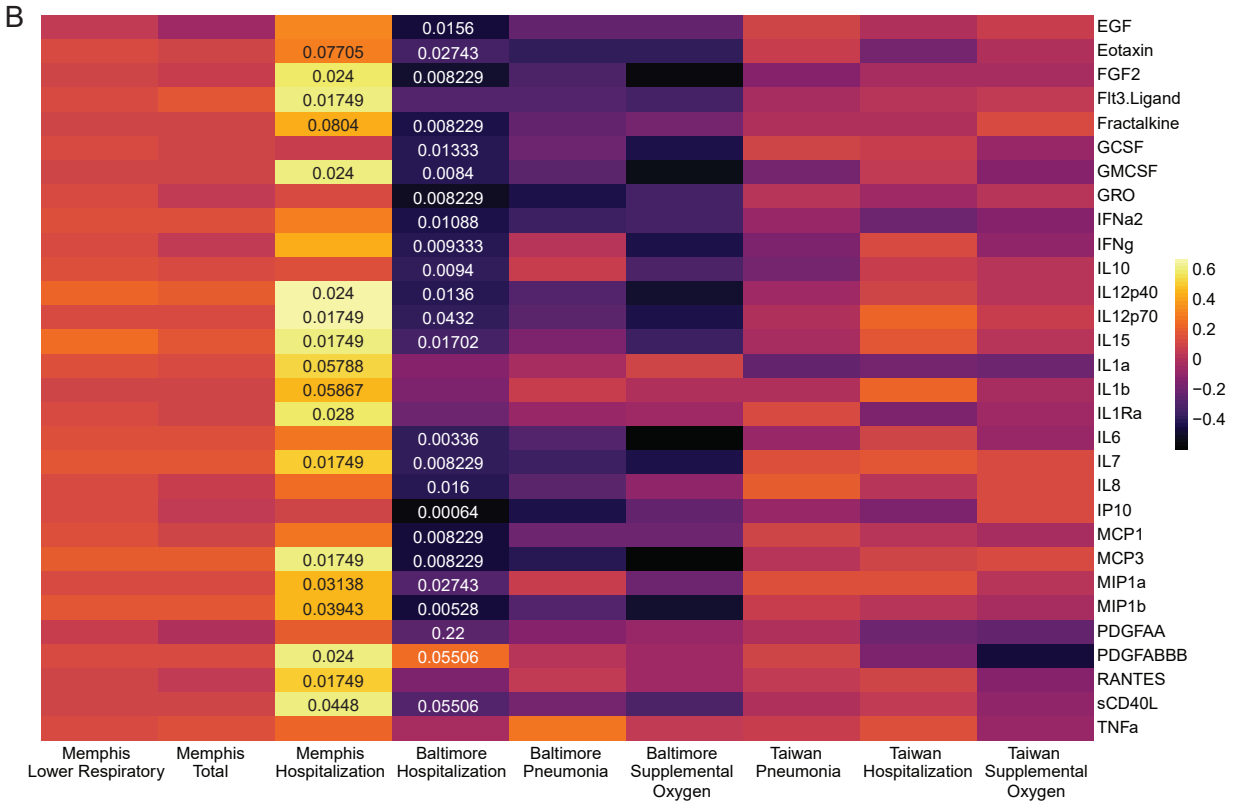


Supplemental Figure 3. Heatmap summary of baseline cytokine regression models. (A) Summary of coefficient estimates for each factor included in the baseline plasma/serum linear regression model. Cytokine models are listed in decreasing order of the adjusted R^2 value (shown on the far right) and coefficients are colored if FDR < 0.05. Red is associated with an increase in the cytokine level per unit change in the factor, and green is associated with a decrease in the cytokine level. (B) Basal plasma/serum cytokine level distributions.



Supplemental Figure 4. Highlights of gene expression associations with independent variables. (A-B) Example of associations based on results from Model #2 (A, determinants of baseline gene expression) and Model #3 (B, determinants of stimulated gene expression). The y-axis is gene expression and the x-axis is the independent variable (age in years, herpesvirus serostatus, ancestry informative marker principal component, baseline gene expression principal component). (C) Examples of eQTLs. *MCP1* (n = 833) and *FLT3LG* (n = 868) gene expression by genotype for rs13900 and rs2946871, respectively.

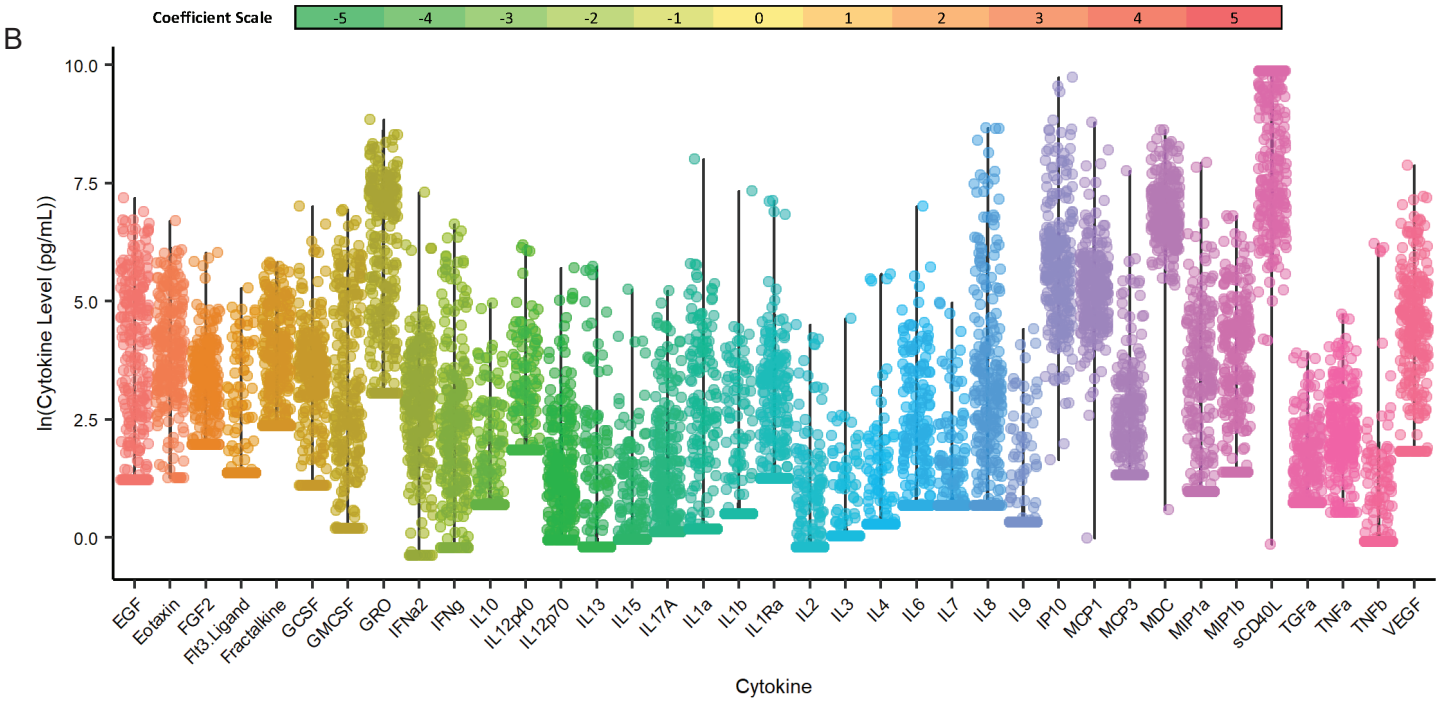




Supplemental Figure 5. Plasma and nasal wash cytokine correlates of severity. Heatmap of plasma (A) and nasal wash (B) cytokine correlates of severity analysis, with significant adjusted p-values displayed. Colors are based on correlation coefficients or a Cohen's D effect size, depending on whether a symptom is continuous or binary, respectively. Presented p-values are from correlations (Spearman and Pearson) or Mann Whitney U tests, and results were adjusted for multiple comparisons by controlling the FDR (< 0.1).

A

Cytokine	Age	Study.Day	Sex.Male	Log.VCN	CMVPos	Age: CMVPos	HSVPos	Age: HSVPos	EBVPos	Age: EBVPos	EBVPos: CMVPos	CMVPos: HSVPos	EBVPos: HSVPos	EBVPos: CMVPos: HSVPos	Adj.R2
IL10	-0.5874	-0.0587	0.2509	0.4600	-0.6830	-0.1276	0.0425	0.2695	-0.1583	0.2630	1.5257	0.9956	0.4518	-1.7824	0.3584
IP10	-0.6166	0.0080	0.2041	0.6615	-1.2799	-0.5005	-0.1757	0.0431	-0.1985	1.2459	1.3050	2.2717	-0.1422	-1.9641	0.3489
IL15	0.0194	-0.0056	0.3004	0.2022	-0.6360	-0.3822	-0.0968	0.4781	-0.3892	-0.2145	1.8998	2.9631	0.6586	-4.1683	0.3010
TNFb	-0.0570	-0.0041	0.3022	0.1136	0.0288	0.0495	0.6813	0.8302	-0.3057	-0.7850	1.9288	0.7604	0.1641	-2.7216	0.2942
MIP1b	-2.1096	0.1970	0.1437	0.1408	-1.5901	-0.9589	2.9331	1.3542	1.6877	1.8147	1.5511	0.7308	-3.1939	-0.6401	0.2521
IL4	-0.2507	-0.0301	0.1829	0.0568	0.0586	0.0144	0.5657	0.5337	-0.0145	-0.3234	1.4769	0.8452	0.1422	-2.5686	0.2490
IL1a	-1.1253	0.0378	0.2607	0.5205	-1.8751	-0.6607	0.0804	1.0447	0.1628	0.8981	3.3974	2.7584	-0.0032	-3.5819	0.2464
MCP1	-0.1291	0.0940	-0.1171	0.3070	-0.8931	-0.5235	-0.9928	0.0853	-0.4618	0.5924	0.3722	2.2769	0.2700	-1.2507	0.2420
TNFa	-1.2841	0.1046	0.2800	0.1284	-0.9306	-0.2655	1.9287	0.8857	0.7262	0.8280	1.3744	0.3214	-1.5506	-0.7523	0.2249
IFNa2	-1.0016	0.0474	0.1560	0.3892	-0.0615	-0.1000	1.0870	1.1556	0.6803	0.3403	0.8192	3.0638	-1.2218	-3.5546	0.2214
Fractalkine	-0.7381	0.0531	0.0261	0.2013	-1.1847	-0.6522	1.0444	0.6399	0.3609	0.8132	2.0190	1.6061	-0.9654	-2.2066	0.2048
IL13	0.1338	0.0105	0.3999	0.1737	-0.0269	-0.2466	0.2157	0.3634	-0.4284	-0.6090	1.9360	1.4272	0.6039	-2.9886	0.2021
FGF2	-0.4958	0.0618	0.3218	0.1446	-0.9162	-0.2508	0.8524	0.6301	0.2189	0.3021	1.6238	1.7633	-0.6028	-2.2256	0.1947
IL12p70	-0.1672	0.0352	0.4088	0.1452	-1.0714	-0.5844	0.6820	0.1821	0.3160	0.6094	2.1132	1.7136	-0.0988	-2.7607	0.1890
EGF	-0.5943	0.2254	0.2382	0.3282	-1.7527	-0.8944	-1.1247	0.9160	-0.5112	0.8791	1.3319	4.1539	0.0541	-3.0222	0.1834
MCP3	-0.9384	0.1528	0.0983	0.2180	-0.6707	-0.1748	0.6836	1.0980	-0.1973	0.0844	1.9107	2.6431	0.0211	-3.7237	0.1740
IL2	0.0183	-0.0137	0.4153	0.1262	-0.8029	-0.7567	0.1492	0.2193	0.1643	0.3543	1.4518	1.7735	-0.0757	-2.0331	0.1732
GMCSF	-0.7331	0.0554	0.4098	0.3321	-1.5228	-1.4543	-0.9168	0.6720	0.0500	1.3538	1.3310	3.2242	-0.3184	-2.0213	0.1672
sCD40L	-0.4422	0.2783	-0.0512	0.3645	-0.5653	-0.3441	-0.9859	0.7662	-0.6361	0.5531	0.2032	2.3928	0.3830	-1.9433	0.1585
IL1Ra	-0.3795	0.1372	-0.0451	0.2043	-0.9516	0.0224	0.2344	0.8292	0.0720	-0.0254	1.9640	3.0805	-0.0448	-4.3557	0.1549
IL6	-0.7663	0.1117	0.4695	0.1763	-1.0349	-0.5904	1.7590	-0.3755	1.2898	1.4923	1.6158	0.5870	-1.9787	-0.9350	0.1542
Flt3.Ligand	0.6184	0.0532	0.5497	0.2149	-0.3946	-0.4729	-0.2288	-0.4844	-0.1184	0.1700	0.8500	1.1729	0.2747	-1.2686	0.1488
IL8	-1.7305	0.5313	-0.0053	0.2662	-0.7178	-0.2163	2.4720	0.2646	1.1112	1.8013	1.0996	2.4508	-1.7499	-3.4827	0.1461
IL12p40	-0.1598	-0.0652	0.2525	0.0584	-0.8477	-0.6326	-0.3722	0.2447	-0.3926	0.3431	1.6514	2.0133	0.7380	-2.5471	0.1460
GRO	-1.2059	0.1903	0.1118	0.3476	-1.1951	-0.8754	-0.1583	1.1881	-0.0013	1.1770	1.1350	2.4497	-0.2608	-1.8318	0.1415
TGFa	0.3623	0.1109	0.2830	0.0699	-0.7113	-0.5358	-0.0178	0.0092	-0.1350	0.0021	1.0660	1.6230	-0.1107	-1.6913	0.1273
MDC	-0.0337	0.0884	0.0143	0.1374	-0.4573	-0.3960	-0.3726	-0.3434	-0.2581	0.6558	0.2235	1.0355	-0.0414	-0.3764	0.1255
Eotaxin	-0.1278	0.1375	0.1731	0.2429	-0.3089	-0.6248	-0.2619	0.2205	-0.0332	0.7042	-0.0434	1.2704	-0.4706	-0.6325	0.1233
IL17A	0.9563	0.0394	0.4188	0.1510	-1.7088	-0.6783	-1.0232	-0.8812	0.0553	0.4966	1.6534	3.3257	0.1436	-3.0768	0.1203
GCSF	-1.1270	-0.0436	-0.2817	0.2141	-0.8567	-0.4152	1.7337	0.5869	0.3970	0.9875	1.5321	0.4635	-1.4084	-0.7805	0.1186
IL1b	-0.9188	0.1101	-0.2564	-0.0428	-0.3924	-0.2932	1.2011	-0.0083	2.1916	1.0974	-0.4535	-0.0092	-2.4446	0.6532	0.1149
MIP1a	-1.1026	0.1199	-0.1066	0.1179	-1.6884	-1.2280	1.0940	-0.1170	0.5728	1.9323	1.9453	1.6499	-1.5113	-1.3446	0.1148
VEGF	0.1944	0.0473	0.6644	0.2818	-1.2003	-0.3599	-0.5785	-0.3831	0.3459	0.5275	1.2994	2.9979	-0.2648	-2.6507	0.1040
IL3	-0.1420	0.0019	0.0837	0.0227	-0.2677	-0.3249	0.2363	0.4353	0.2337	0.0320	0.5315	0.7438	-0.3464	-0.8183	0.0981
IFNg	0.3701	-0.0429	0.5934	0.2926	-1.3498	-0.7027	-0.7589	-0.3456	0.0905	0.4439	1.8268	3.0936	0.1470	-3.1874	0.0851
IL9	0.1107	-0.0550	0.3552	-0.0201	-0.3268	0.0646	-0.3496	0.1152	-0.3608	-0.4376	0.7928	1.9024	0.6888	-2.1145	0.0685
IL7	-0.0377	0.0675	0.1323	0.0283	0.0167	0.2053	0.2239	0.1528	-0.0270	-0.2860	0.5495	1.1632	0.0828	-1.5705	0.0208

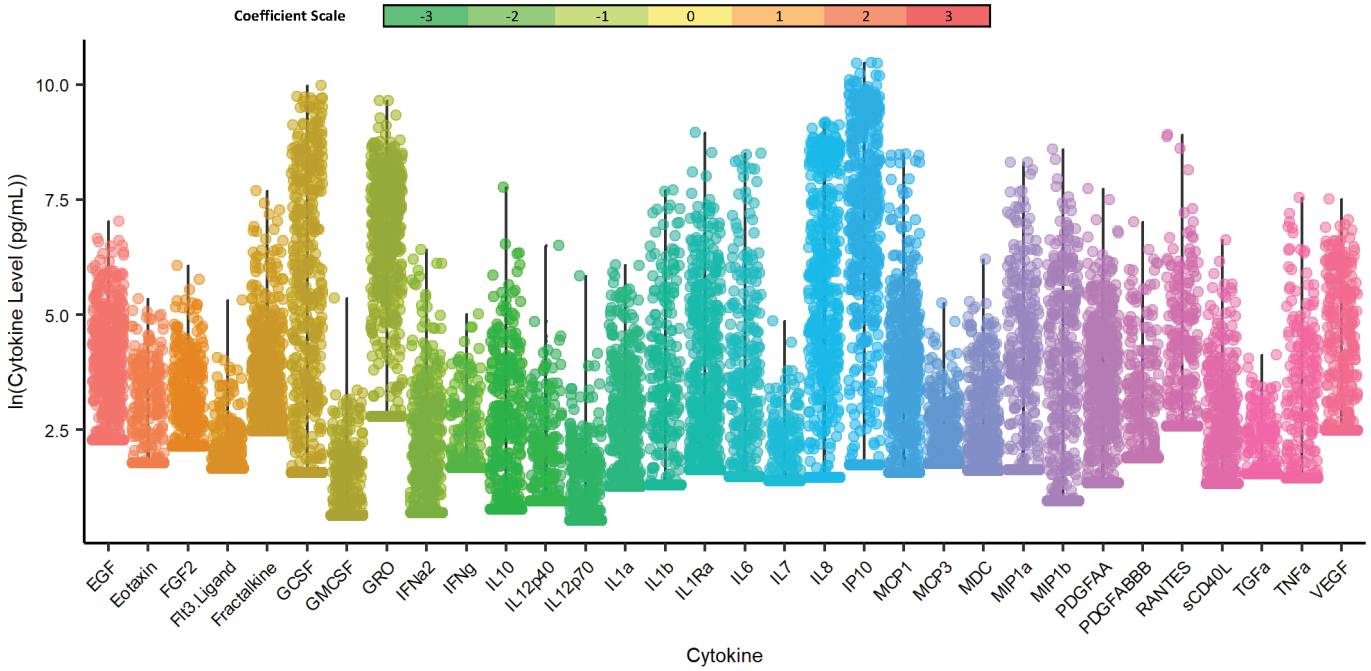


Supplemental Figure 6. Heatmap summary of acute plasma cytokine regression models. Samples were from influenza infected subjects. (A) Summary of coefficient estimates for each factor included in the infected plasma linear regression model. Cytokine models are listed in decreasing order of the adjusted R^2 value (shown on the far right) and coefficients are colored if FDR < 0.05. Red is associated with an increase in the cytokine level per unit change in the factor, and green is associated with a decrease in the cytokine level. (B) Acute plasma cytokine level distributions.

A

Cytokine	Age	Study.Day	Sex.Male	Log.VCN	CMVPos	Age: CMVPos	HSVPos	Age: HSVPos	EBVPos	Age: EBVPos	EBVPos: CMVPos	CMVPos: HSVPos	EBVPos: HSVPos	EBVPos: CMVPos: HSVPos	Adj.R2
RANTES	-1.4776	-0.0022	-0.1340	0.1326	-0.6570	0.0316	-0.1059	0.4904	-0.4976	1.0175	0.6927	0.4233	0.0585	-0.4770	0.332
IFNa2	-0.6786	-0.1032	-0.1711	0.3624	0.4912	-0.1265	-0.4204	-0.1818	-0.6756	0.7716	-0.4320	-0.6776	0.3987	1.0873	0.321
MIP1b	-2.0749	-0.0859	0.0685	0.5850	-0.4289	-0.0608	0.4615	0.4124	-0.0169	1.5636	0.4344	0.1856	-0.6486	-0.1606	0.314
FGF2	-0.9453	0.0054	-0.0363	0.1898	-0.4048	-0.0185	0.0024	0.2167	-0.2611	0.7276	0.5318	0.4335	-0.0917	-0.5024	0.306
PDGFABBB	-1.0112	-0.0144	-0.1526	0.1787	-0.4939	-0.0454	-0.1640	0.2566	-0.4063	0.8132	0.4856	0.5301	0.0294	-0.3881	0.301
IL10	-1.2074	-0.0908	-0.0346	0.4968	-0.1795	-0.0075	0.1769	0.1550	-0.2585	0.9289	0.1132	-0.2709	-0.1841	0.5233	0.300
IL6	-1.3389	-0.0603	-0.0887	0.5361	-0.4547	-0.0856	0.3458	0.2499	-0.3158	1.0644	0.5040	-0.2156	-0.3461	0.3720	0.295
IFNg	-0.5473	-0.0120	-0.0721	0.1977	-0.1668	-0.0704	-0.0058	0.1069	-0.2155	0.4809	0.1567	0.1250	-0.0383	0.0356	0.289
VEGF	-1.2603	-0.0204	-0.1035	0.2228	-0.4571	-0.1230	-0.1465	0.2048	-0.2344	1.0962	0.4564	0.8210	-0.2707	-0.5466	0.289
Fractalkine	-0.9536	-0.0627	-0.0777	0.3599	-0.0663	-0.2316	0.1727	-0.0968	-0.3078	1.1164	-0.0745	0.0414	-0.4442	0.6122	0.286
IL1Ra	-1.5814	-0.0808	-0.1758	0.2606	-0.1519	-0.1531	-0.0443	0.1136	-0.5321	1.4833	0.1563	0.6069	-0.1814	-0.3167	0.285
sCD40L	-1.0431	-0.0310	0.0013	0.2368	-0.7654	-0.0721	-0.3201	0.2271	-0.4050	0.8180	0.8582	0.8641	0.0351	-0.7212	0.280
TNFa	-0.9269	-0.0169	-0.0985	0.3963	-0.2459	-0.0121	0.1210	0.2813	-0.1593	0.6229	0.2086	-0.2959	-0.2360	0.3603	0.257
MIP1a	-1.5491	-0.0418	-0.0631	0.4249	-0.3489	-0.0444	0.3185	0.3686	-0.0028	1.1586	0.3076	0.1130	-0.5590	-0.0429	0.256
GMCSF	-0.7425	-0.0036	-0.0716	0.1493	-0.2507	-0.0562	-0.0154	0.1673	-0.0763	0.6065	0.1228	0.3017	-0.2294	-0.0552	0.255
MDC	-0.8485	-0.0478	-0.0280	0.0818	-0.4146	-0.0355	-0.3037	0.1879	-0.3376	0.6294	0.3510	0.5851	0.1050	-0.3136	0.237
IL12p70	-0.7346	-0.0168	-0.0352	0.1454	-0.4191	-0.0267	-0.0433	0.1728	-0.2587	0.5753	0.5139	0.3598	-0.0595	-0.3769	0.236
Eotaxin	-0.6210	-0.0123	-0.0609	0.0896	-0.3334	-0.0274	-0.1205	0.1168	-0.1387	0.5131	0.2653	0.4822	-0.0637	-0.3557	0.235
GCSF	-2.8086	-0.3114	-0.0862	0.0777	-0.9625	-0.1586	-0.1860	0.5492	-0.7696	1.8876	1.4126	0.4714	0.2993	-0.5759	0.232
IL1b	-1.5241	-0.0248	-0.1652	0.1845	-0.4864	-0.0436	0.1096	0.4152	-0.1683	1.1221	0.2826	0.1023	-0.3644	0.1624	0.221
Flt3.Ligand	-0.3865	-0.0110	-0.0404	0.1006	-0.2368	-0.0312	-0.2022	0.1058	-0.2252	0.3026	0.1796	0.2419	0.0896	-0.1123	0.214
IL1a	-1.1099	-0.0584	-0.1163	0.2400	-0.5514	-0.3175	0.0336	0.2933	-0.1003	1.0428	0.6520	0.6426	-0.5574	-0.2620	0.211
IP10	-0.1511	-0.5635	-0.0908	0.6460	0.1458	-0.0607	0.3088	-0.6153	-1.0982	-0.0377	0.0748	-1.0626	0.0199	1.8618	0.203
IL12p40	-0.6352	-0.0118	-0.0883	0.1569	-0.5532	-0.0622	-0.2406	0.1319	-0.2871	0.5500	0.4877	0.6876	0.0324	-0.4913	0.199
MCP3	-0.4562	-0.0073	-0.0187	0.1593	-0.3465	-0.0278	-0.0533	0.1007	-0.1650	0.3659	0.3396	0.2897	-0.0884	-0.1794	0.199
MCP1	-0.7382	-0.1627	-0.2032	0.4581	-0.1866	-0.1374	0.6990	-0.0148	-0.2400	0.7581	0.2491	-0.4924	-0.9743	0.8910	0.178
TGFa	-0.3740	-0.0011	-0.0473	0.0157	-0.2120	-0.0136	-0.0499	0.0983	-0.0870	0.2870	0.1674	0.1740	-0.0402	-0.0986	0.163
IL8	-2.1434	-0.2907	-0.1995	0.1478	-0.1616	-0.0700	0.7039	0.1712	-0.2434	1.4735	0.0179	-0.7922	-0.5714	1.2125	0.160
GRO	-1.2874	-0.2565	-0.1324	0.0403	-0.1200	-0.1532	0.3730	0.1332	-0.6436	0.6075	0.2947	-0.5608	-0.0695	0.7886	0.159
IL7	-0.2078	-0.0147	0.0189	0.1046	-0.3871	-0.0271	-0.2584	0.0368	-0.3235	0.1907	0.4509	0.3308	0.2146	-0.3213	0.154
EGF	0.3233	-0.1739	-0.1145	-0.1100	-1.2474	-0.0413	-0.2612	-0.6705	-0.5924	0.1431	1.0814	1.1319	0.0997	-0.7124	0.118
PDGFAA	-0.5016	-0.1477	0.0180	-0.0420	-0.1029	-0.1962	0.0005	-0.0014	-0.4792	0.3703	0.0085	0.1481	-0.2555	0.4710	0.105

B



Supplemental Figure 7. Heatmap summary of acute nasal wash cytokine regression models. Samples were from influenza infected subjects. (A) Summary of coefficient estimates for each factor included in the infected nasal wash linear regression model. Cytokine models are listed in decreasing order of the adjusted R^2 value (shown on the far right) and coefficients are colored if FDR < 0.05. Red is associated with an increase in the cytokine level per unit change in the factor, and green is associated with a decrease in the cytokine level. (B) Acute nasal wash cytokine level distributions.