Continuous immunotypes describe human immune variation and predict diverse responses

Kevin J. Kaczorowski[a,b,c], Karthik Shekhar[d], Dieudonné Nkulikiyimfura[e], Cornelia L. Dekker[f], Holden Maecker[g], Mark M. Davis[g,h,i,1], Arup K. Chakraborty[a,b,c,j,k,1], and Petter Brodin[e,l,1]

[a]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, MA 02139; [b]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; [c]Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139; [d]Klarman Cell Observatory, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA 02142; [e]Science for Life laboratory, Department of Medicine, Karolinska Institutet, Solna 17177, Sweden; [f]Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94304; [g]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94304; [h]The Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94304; [i]Institute for Immunity, Transplantation, and Infection, Stanford University School of Medicine, Stanford, CA 94304; [j]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139; [k]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139; and [l]Department of Neonatology, Karolinska University Hospital, Stockholm 17176, Sweden

The immune system consists of many specialized cell populations that communicate with each other to achieve systemic immune responses. Our analyses of various measured immune cell population frequencies in healthy humans and their responses to diverse stimuli show that human immune variation is continuous in nature, rather than characterized by discrete groups of similar individuals. We show that the same three key combinations of immune cell population frequencies can define an individual's immunotype and predict a diverse set of functional responses to cytokine stimulation. We find that, even though interindividual variations in specific cell population frequencies can be large, unrelated individuals of younger age have more homogeneous immunotypes than older individuals. Across age groups, cytomegalovirus seropositive individuals displayed immunotypes characteristic of older individuals. The conceptual framework for defining immunotypes suggested by our results could guide the development of better therapies that appropriately modulate collective immunotypes, rather than individual immune components.

human immune variation | immune cell composition | systems immunology | aging

In the peripheral blood of humans over 100 different immune cell populations can be distinguished based on the expression of cell surface and intracellular markers. These cell populations span various activation and differentiation states within broader classes of principal immune cell lineages in the adaptive and innate immune systems (1). With recent advances in single-cell measurement technologies such as high-dimensional flow cytometry and mass cytometry (2), these different immune cell populations can be simultaneously quantified in an individual sample. Leveraging these developments, recent population studies have profiled the global immune cell composition in healthy adult humans (3–5) and found these to be highly stable within an individual over the course of weeks to months, but highly variable between individuals (3, 6). Although variations in a few immune cell populations are heritable, the majority of immune cell population frequencies and functional responses are largely determined by nonheritable influences in healthy individuals (4). For example, cohabitation has been shown to be a strong influence, emphasizing the importance of shared environmental factors within a household (6). Most of the specific environmental factors that shape human immune systems have yet to be determined, but monozygotic (MZ) twin pairs discordant for human cytomegalovirus (HCMV) seropositivity were more different in the majority of immune measurements compared with seronegative twins (4).

Many studies have focused on individual immune cell populations, their functional properties, regulatory roles, and clinical significance. However, a prototypical immune response involves the coordinated action of multiple cell populations (7, 8), providing both inhibitory and stimulatory feedback, leading to a balanced response that fulfills its protective function while preventing immunopathology. The rules for immune cell interdependence are understood in only a few cases, but there may be many that are unknown. Many current therapies target individual cell populations and sometimes give rise to unexpected side effects, involving completely different cell populations from the ones directly targeted. For example, a subgroup of patients treated with the B-cell–depleting antibody Rituximab, which targets the CD20 antigen expressed on B cells, developed delayed neutropenia despite the fact that neutrophils do not express CD20 (9). An understanding of why such complications occur in selected patients requires a more complete knowledge of how specific immune cell populations depend on each other, regulate each other, and affect systemic responses.

The immune system is a complex adaptive system, without any apparent master regulator. Therefore, all systemic behaviors are emergent properties dependent upon the collective actions of all system components. As such, immune responses will be unpredictable until these emergent properties are understood. Furthermore, specific immune cell populations are dependent upon each other for stimulation and inhibition, and different cell population frequencies vary widely across healthy individuals. Therefore, we hypothesized that differences in immune responses

Significance

The human immune system consists of many different white blood cells that coordinate their actions to fight infections. The balance between these cell populations is determined by direct interactions and soluble factors such as cytokines, which serve as messengers between cells. Understanding how the interactions between cell populations influence the function of the immune system as a whole will allow us to better distinguish patients most at risk for specific infections or immune-mediated diseases and inform vaccination strategies. Here, we determine key collective interactions between white blood cells present in blood samples taken from healthy individuals. This perspective allows us to predict functional responses and describe previously unappreciated differences between age groups and in individuals carrying cytomegalovirus.

between individuals are influenced by differences in some collective measures of their frequencies of various immune cell populations; i.e., certain combinations of immune cell populations are determinants of response. In this paper, we test this hypothesis by analyzing responses to stimuli in cohorts of healthy humans.

We first analyzed immune cell composition data from a total of 1,575 healthy individuals across multiple cohorts, both published (6, 10) and previously unpublished (149 individuals) in addition to our previously published data (4), combined in Dataset S1. We also took advantage of a set of diverse functional responses measured in 329 individuals from a number of Stanford cohorts (refs. 4 and 11 and previously unpublished data, combined in Datasets S2 and S3). Our goal was to define a small number of collective variables, or specific combinations of cell populations, variations in which could predict functional responses. Thus, our results depend on the specific cell populations and functional responses measured in this study. As such, future studies across additional independent cohorts from different geographical locations measuring the same or different cell populations and functional responses are needed to further validate the concept that a few properly chosen collective variables can predict a person's immune response to stimulation. Nonetheless, our current analysis is a proof of concept that immunology can benefit from a more collective view of the immune system as a whole.

We identified three key combinations of immune cell frequencies (not individual cell frequencies) that robustly predict diverse functional responses in healthy individuals. Viewed from the perspective of how individual humans are different in terms of key combinations of cell populations, we find that, rather than separate clusters of individuals sharing immunological traits, there is a continuum of human immune variation. This variation is influenced, but not dictated, by age, sex and HCMV seropositivity. Even though young individuals vary greatly in their frequencies of individual cell populations, they are highly similar to each other in terms of these three key combinations of immune cell frequencies. With age, unrelated individuals become more heterogeneous when viewed from this perspective, likely reflecting the diversity of environmental exposures with time. Our results show that HCMV seropositivity is associated with a shift in immune phenotype along an age axis, leading to a phenotype characteristic of older populations, irrespective of the actual age of the individual. This conceptual framework of defining immunotypes using the collective states of many immune system components provides a sensitive way of describing immune variation and could guide the development of better immune modulatory therapies in the future.

Results

Human Cohorts and Measurements. We analyzed human immune cell frequency data from multiple cohorts, described in further detail in *Materials and Methods*. These are referred to as the Stanford cohorts of 398 individuals, 226 MZ and dizygotic (DZ) twins from a published twin study (4), 23 individuals from a published longitudinal study (11), and 149 previously unpublished subjects (combined in Dataset S1); the Roederer et al. (10) cohort, consisting of 669 female twins sampled in the United Kingdom; and the Carr et al. (6) cohort, consisting of 670 nontwin subjects, both men and women, sampled in Belgium.

All three cohorts measured the frequencies of various immune cell populations in samples of peripheral blood mononuclear cells (PBMCs). In addition, for the Stanford cohort, a diverse set of functional responses was measured, including 168 signaling responses to cytokine stimulation in vitro (Dataset S2), as well as response to vaccination to five influenza strains (Dataset S3), listed in *Materials and Methods*. The cell signaling and antibody responses are referred to as functional responses.

Variation of Immune Cell Composition Across Healthy Individuals Is Continuous Rather than Discrete. There has been much interest recently in characterizing immune variation across healthy individuals over a variety of ages (3–6, 10, 12–15). It is also well established that healthy individuals respond differently to the same immunological stimulation. A puzzling example of this is that although most patients receiving certain influenza vaccines mount robust antibody responses, some otherwise healthy patients fail to produce a protective response. The reasons for such disparities are only partially understood (16). Along the same lines, novel immune-modulatory drugs used for cancer therapy yield variable and unpredictable responses in different patients (17, 18).

To understand such similarities and differences in immune responsiveness, a conceptual framework for assessing and studying interindividual differences in immune cell composition will be helpful. Using data from the three cohorts, we asked whether a healthy population formed well-separated clusters of individuals with similar immune cell composition traits. We analyzed each dataset separately as differences in the measurement panels prevented us from combining them (Table S1). Principal component analysis (PCA) (19) was used to project high-dimensional immune cell compositions onto the two principal components (PCs) that account for most of the underlying variance in each dataset (Fig. 1A). In each of the three datasets, individuals were spread across a continuous "cloud" and did not segregate into separate clusters.

We quantified this by performing *k*-means clustering on the top PCs for each dataset (Fig. S1). The extent of clustering, when quantified (by average silhouette coefficient and between-cluster variance, Fig. S1 *B and C*), is no higher for the cohorts in this study compared with randomized "null" datasets (*Materials and Methods*, *Clustering Methods*). Although Fig. 1A shows only the top two PCs, the analyses in Fig. S1 *B and C* show lack of clustering using up to the top 10 PCs (which explain around 70% of total variance; Fig. S1A). Finally, *t*-distributed stochastic neighbor embedding (20–22), a nonlinear dimensionality reduction technique that overcomes the limitation of linearity in PCA, also reveals a similarly continuous distribution of individuals with respect to their immune cell composition (Fig. S1D). Thus, we concluded that there is no evidence
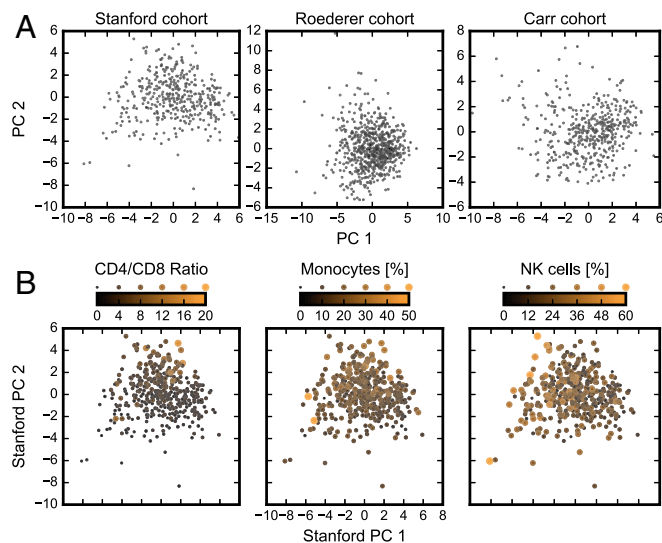
Fig. 1. Immune cell composition varies continuously across the healthy human population. (*A*) Projection of immune cell compositions of the Stanford, Roederer, and Carr cohorts onto the top two of their respective PCs (Fig. S1). (*B*) Overlay of individuals from the Stanford cohorts with extreme values of individual immune cell populations (or two, in the case of CD4/CD8 ratio) on the top two PCs of immune cell composition data. Solid circles scale linearly in both area and color (Fig. S2).

of clustering of individuals based on their immune cell composition in any of these datasets. We acknowledge the possibility that discrete clusters were not found simply because the particular cell populations measured in these studies represent a limited panel. However, this possibility is unlikely given the number and diversity of cell populations analyzed, and instead our results suggest that human immune cell composition variation is continuous rather than discrete among healthy individuals.

Healthy Individuals with Extreme Values of Specific Cell Types Are Not Outliers in the Context of Their Collective Immune Cell Composition. Within these continuously distributed immune cell compositions, we sought to understand whether individuals with extreme outlier values of specific immune cell population frequencies are also outliers when taking the full set of measurements into account. As three key examples involving both innate and adaptive cells, we analyzed CD4/CD8 ratios (i.e., the ratio between CD4$^+$ and CD8$^+$ T-cell population frequencies) and monocyte and NK cell frequencies. The CD4/CD8 ratios are commonly studied and have been related to diverse immunological conditions (23–25), most notably as a prime indicator of AIDS before the identification of the virus (26). Normal ratios are believed to be ~2 and our results show that CD4/CD8 ratios among healthy individuals varied broadly, from 0.34 to 16 (Fig. 1B). To our surprise, the outlier individuals with very high or very low CD4/CD8 ratios were distributed throughout the cloud of points representing healthy individuals when viewed in the context of their collective cell composition (Fig. 1B). In other words, the CD4/CD8 ratio, although exhibiting a large range, did not significantly contribute to the underlying correlated variation captured by PC1 and PC2. A similar pattern was found for other well-defined cell populations such as monocytes, which varied broadly from 0% to 49% of non–T- and non–B-cell PBMCs, and NK cells, which varied from 0% to 59% of non–T- and non–B-cell PBMCs (Fig. 1B). Despite such extreme individual values of specific cell population frequencies, these individuals were not outliers with respect to their global immune cell composition (Fig. 1B). Fig. S2 corroborates this finding by overlaying the bottom and top 5% of these three individual immune cell frequencies onto the distribution of the top four PCs individually. We conclude from these findings that individual immune cell measurements alone cannot identify individuals with outlier phenotypes and only by taking the collective state of their immune cell composition can such individuals be defined. This has important implications for the definition of immunological health and disease risk.

Diverse Functional Responses Can Be Predicted Based on an Individual's Immune Cell Composition. Given that the immune cell compositions are continuously distributed among healthy individuals and that the possible combinations of immune cell frequencies in any individual are large, we explored whether the extent to which an individual's immune system responds to stimuli can be adequately predicted by a smaller set of specific combinations of individual immune cell frequencies; each such combination is a collective variable. In other words, we sought to choose a small set of collective variables in a way such that individuals with similar coordinates in the space defined by these collective variables respond similarly to a particular stimulation. Toward this end, we used partial least-squares (PLS) regression (27) to correlate the immune cell compositions of individuals in the Stanford cohort with a large set of measured functional responses, thereby defining a particular choice of collective variables [termed the PLS "latent variables" (LVs)]; note that response measurements were not available for the Roederer and Carr cohorts so this analysis applies only to the Stanford cohort. The functional responses analyzed involved three JAK-STAT signaling pathways. These pathways are prototypic membrane to nucleus pathways stimulated by a range of cytokines and growth factors and

regulate growth, survival, differentiation, and pathogen resistance in the immune system (28). We analyzed the phosphorylation of STAT1(pY701), STAT3(pY705), and STAT5(pY694) in response to one of seven different cytokines (IL-2, Il-6, IL-7, IL-10, IL-21, IFN-α, or IFN-γ) in PBMCs in vitro (4, 11). The three responses were analyzed by intracellular antibody staining and flow cytometry, allowing for the responses of eight different immune cell populations [monocytes, B cells, total CD4$^+$, naive CD4$^+$ (CD45RA$^+$), memory CD4$^+$ (CD45RA$^-$), total CD8$^+$, naive CD8$^+$ (CD45RA$^+$), and memory CD8$^+$ (CD45RA$^-$) T cells] to be analyzed separately. Also, antibody responses 28 d after a seasonal flu vaccination were included as a different form of functional responses (4). Interindividual differences in such responses have previously been shown to distinguish differentially regulated immune systems (29–31).

For each of the 168 responses above, PLS (*Materials and Methods* and Dataset S4) was used to find the LVs or the set of linear combinations of immune cell frequencies that have the highest covariance with the functional response. Our measurements span a high-dimensional space defined by axes, wherein each axis represents the frequency of a particular cell population. The LV signatures for the k-th functional response make up the columns of the LV matrix, $\mathbf{R}^{(k)}$, where the j-th column represents the direction in this high-dimensional space most correlated with the portion of the functional response unexplained by the previous $j-1$ LVs. This front loading (ordering of LVs in the order of explanatory power) means that potentially only a small number of such directions would have predictive value. Each direction corresponding to a LV can be viewed as an axis representing a specific combination of individual cell frequencies. A particular person's coordinates in the low-dimensional space defined by a small number of combinations of cell populations (or axes) defines how he or she will respond to the stimulus—i.e., location of an individual's cell composition in this lower-dimensional space defines what we term the individual's immunotype.

We compared the performance of PLS with principal component regression (PCR), another widely used technique in which the few important axes are chosen to be the PCs, ordered to explain the most to least variance within the multidimensional immune cell compositions across patients. Model performance was measured using cross-validation to estimate an out-of-sample negative Spearman correlation between measured and predicted response (*Materials and Methods*). Both PLS and PCR models perform significantly better than a null model in which response measurements are permuted across individuals to destroy any relationship between immune cell composition and response (Fig. 2 A and B). This indicates that the correlations between immune cell composition and these functional responses are unlikely to occur by chance. Although this is true for all but six signaling responses (*Materials and Methods*, *PLS Regression*), only one of the five flu-vaccine responses (antibodies to influenza B/Brisbane/60/2008 strain) could be predicted by PLS/PCR better than by a null model. The 10 responses with poor PLS fit were removed from further analysis because the results are not statistically significant. Compared with PCR, the PLS models achieve lower cross-validation error with fewer LVs, a result of the "front-loaded" explanatory power in the first LVs. For a typical functional response, the first LV of PLS accounts for 50% of the possible explanatory power, whereas the first PC of PCR accounts for only 20%, requiring approximately nine PCs to reach the same model performance (Fig. 2A).

Remarkably, the top three most predictive LV signatures were highly concordant across different functional responses (Fig. 2 C–E). To quantify this, we calculated the similarities of predictive signatures as absolute dot products between their respective LV signatures. According to this metric, more than 64% of response pairs showed similar LV1–3 signatures (P values < 0.05 compared with a null distribution of randomly chosen vector pairs on a 34-dimensional hypersphere, Fig. S3). This result suggests that
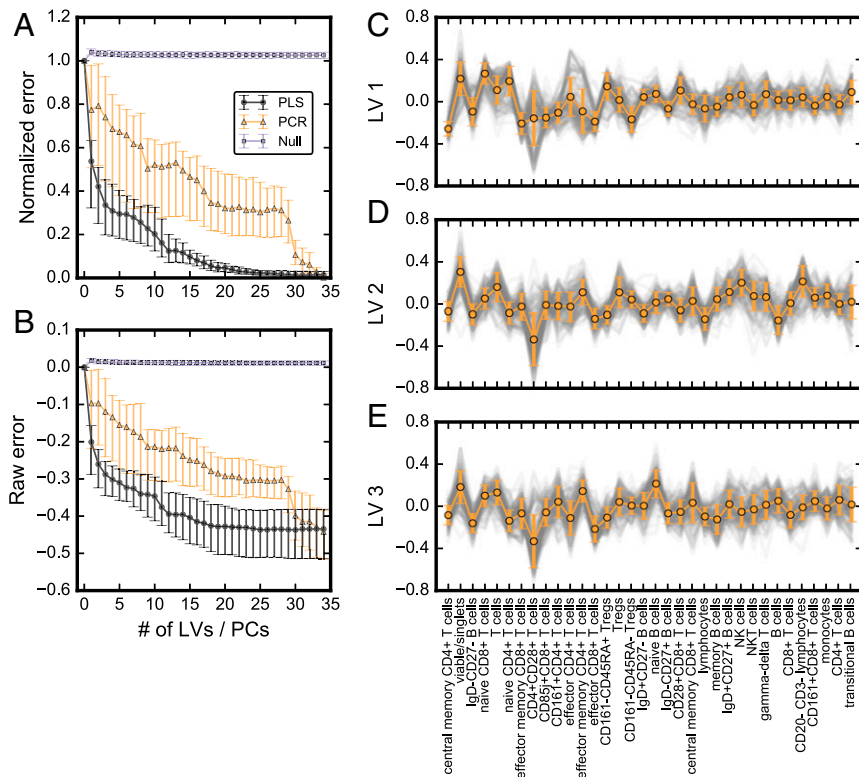
Fig. 2. A few key combinations of immune cell frequencies predict functional responses: regression model analysis. (*A* and *B*) Learning curves for PLS and PCR models, which plot model error as a function of number of variables (LVs or PCs for PLS and PCR, respectively) included in the model. The curves are either normalized (*A*) or raw (*B*), calculated across all responses [median ± interquartile range (IQR)]. Raw error is defined as 10-fold cross-validated negative Spearman correlation between measured and predicted responses to estimate out-of-sample error. Normalized error is defined as the remaining fraction of total possible reduction in raw error $\left(\frac{e - e_{min}}{0.0 - e_{min}}\right)$. (*C*–*E*) Raw latent variable signatures (gray) with mean ± SD (orange) across all responses, for the top three LVs (*C*, *D*, and *E*, respectively) ([Fig. S3](#)).

similar combinations of immune cell population frequencies are predictive of functional responses to diverse stimuli, such as IL-2–induced STAT5 phosphorylation and IFN-α–induced STAT1 phosphorylation. The fact that functional responses to many different stimuli can be predicted from immune cell population frequencies only suggests a previously unappreciated relationship between immune cell composition and functional competence in a given individual. Furthermore, the similarities between predictive signatures suggest that particular collective dependencies between immune cell populations are especially important in determining even diverse functional responses spanning multiple pathways and stimuli. However, not all responses are explained by the same three LVs, and we decided to study more carefully the responses not explained by these key combinations of immune cell frequencies.

The Response to Influenza Vaccination Is Predicted by Different Combinations of Immune Cell Populations Compared with Those That Predict Responses to Cytokines. In contrast to in vitro stimulation with one given cytokine, in vivo responses are complex and involve multiple pathways within several cell populations. One powerful way of querying such complex responses in humans is offered by studying immune responses to vaccination. Such responses involve multiple stimuli and signaling pathways acting in concert in vivo and synergies between these components that could result in predictive LVs different from those found to predict responses to cytokines above. However, responses to vaccination were measured only in one case for a sufficient number of individuals to enable statistically meaningful analysis using PLS. This vaccine response involved the antibody responses to the influenza B/Brisbane/60/2008 strain contained in a commercial vaccine (Fluzone; Sanofi-Pasteur), measured using the hemagglutination inhibition (HAI) assay (4).

There was no measurable response to this vaccine in 31% of individuals and the PLS method had much improved explanatory power when these nonresponders were removed ([Fig. S4A](#)). Hence, our LV description of immune cell composition is pre-

dictive of the degree of response in the event that one does occur, but not of which patients do not mount an immune response at all. Importantly, the first three LVs that predict the response to vaccination are distinct from those that are predictive of the responses to in vitro stimulation by individual cytokines, in line with our hypothesis that vaccination elicits a more complex set of stimuli. This may be why different synergies between immune cell populations (different LVs) determine vaccine responsiveness compared with stimulation with a single cytokine. Some of these synergies are well known in the field, such as the requirement for T-cell "help" to enable most antibody responses, but there may be many other dependencies revealed here that have been less well characterized previously.

To further explore why 30% of the cytokine-signaling responses were not predicted by the LVs that predicted the responses to all other responses to cytokine stimulation (Fig. 2), we clustered the similarity matrix (*SI Materials and Methods*, *Spectral Clustering*) within the first LV for all of the cytokine responses ([Fig. S4B](#)). This analysis revealed two clusters of responses ([Dataset S5](#)), the larger of which ("group 1") contained 123 responses and the smaller of which ("group 2") contained 43 responses. The response magnitudes in these groups were very different ([Fig. S4C](#)); specifically, the 43 responses in group 2 involved signaling responses of very low dynamic ranges, i.e., weak to no induction of phospho-STAT1, -3, or -5 in response to stimulation. In contrast, the 123 responses in group 1 showed stronger amplitudes and variation across these responses was predicted quantitatively by the same LVs. We again conclude from this observation that immune cell composition is predictive of the relative strength of a functional response upon stimulation.

Our findings are summarized in Fig. 3, which compares the LVs of the vaccination response to a selection of representative signaling responses from group 1. This representation highlights shared cell populations across predictive signatures, suggesting that these populations have larger weight in the specific combinations of immune cell populations that define the LVs for response to
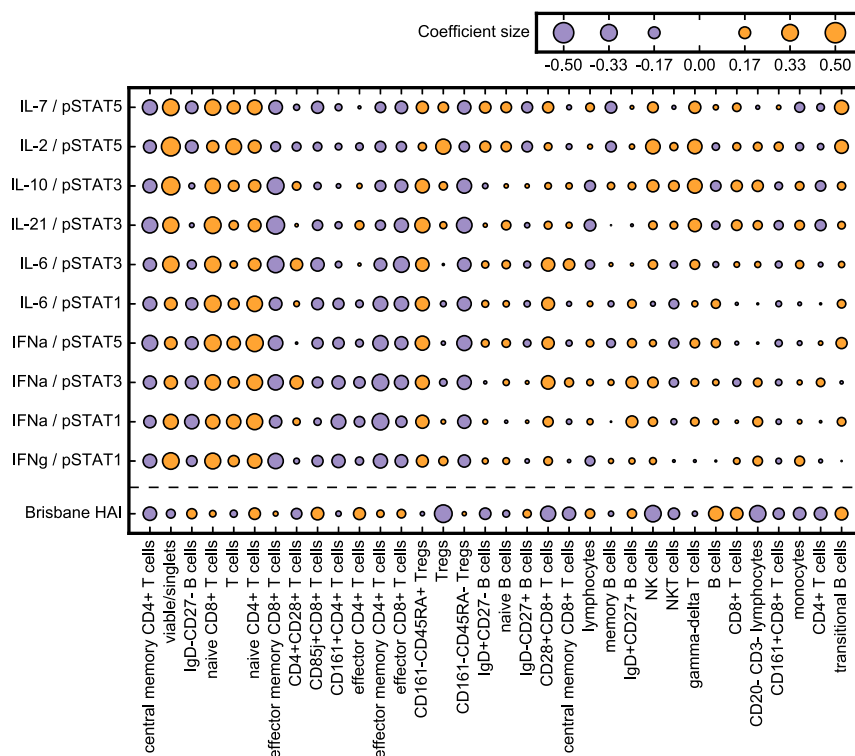
Fig. 3. Predictive signatures for diverse signaling responses are uniform across responses but distinct from that for vaccination by a single strain of influenza. Signatures show the top LV from group 1 responses (see text), averaged over responding cell type (top 10 rows), contrasted with that for the single flu-vaccination response (antibodies to influenza B/Brisbane/60/2008 strain; bottom row). Each solid circle represents the contribution of a particular immune cell population in the top LV that is predictive of functional response. Circle area is proportional to magnitude of the average LV1 signature coefficient, with positive or negative sign indicated by orange or purple, respectively (Fig. S4 and Dataset S5).

cytokine stimulation or vaccination. For cytokine stimulation, 12 subtypes tend to have the largest weights, most of which have the same sign across responses—a positive (negative) sign indicates that a high (low) frequency of that cell population is associated with strong functional responses. This is in contrast with the signature for the vaccination responses, in which a different immune cell signature was found with a different combination of immune cell populations being predictive of strong vaccine responsiveness.

Immunotypes Become More Heterogeneous with Age. The immunotypes of individuals can be visualized by projecting individuals onto the space defined by the top three LVs. Analyzing the distribution of individuals within this space can help us better understand human immune variation. The individuals studied correspond to a cloud of points distributed in a roughly conical shape. Younger individuals concentrate near the vertex of the cone and older individuals are distributed toward the base of the cone (Fig. 4A). Thus, older populations exhibit greater heterogeneity in their immunotypes than younger individuals.

The conical shape of the point cloud, and differences in immunotypes across populations of differing ages, can be quantified by graphing distributions of interpatient distances (defined as Euclidean distance on the top three LVs) in different age groups (Fig. 4B), revealing that the youngest 10% of individuals have a much narrower distribution than the older patients. Thus, an older person's immune response is not predicted very well by age, but rather by the individual's immunotype. Young people are similar enough in their immunotypes for their immune responses to be better predicted by age alone. Furthermore, the diversification appears to be quite rapid, where distributions of interprofile distance for the youngest 20% and 60% (age <21 y and <43 y, respectively) quickly approach the overall distribution (Fig. 4B). One

interpretation of this is that fewer significant new environmental exposures with an influence on immune variation are encountered after the mid-40s.

This result is consistent with a previous study in which immune cell frequencies were found to be more similar between younger (<20 y) than older (>60 y) monozygotic twins, presumably as a consequence of cumulative environmental influences occurring during the decades in between 20–60 y of age (4). However, it is important to emphasize that the distributions shown in Fig. 4B were calculated after excluding MZ or DZ twin pairs, as such pairs are expected to share many determinants of their immunotypes. Our findings thus show that all young individuals, and not just MZ twins, are more similar to one another than older individuals are, presumably as a consequence of the diverse cumulative influences shaping an individual's immune system over time.

The contributions of individual immune cell populations to the age-related changes (Fig. 4C) reveal an expected shift from naive to memory cells within adaptive lymphocyte populations, but also a lesser-known increase in CD161+ CD4+ T cells and NKT cells with age. In contrast, monocyte frequencies are stable throughout life (Fig. 4C). Finally, because the heterogeneity is observed specifically in the top LVs predictive of diverse immune system responses, it indicates that the diversification of immune cell composition with age in response to environmental influences has important effects on the functional responses elicited by an aging human immune system.

Collective Predictors of Immune Responses Are Shaped by Heritable and Nonheritable Influences. Brodin et al. (4) found that the majority of immune cell frequencies were determined by nonheritable influences. However, in defining a small set of combinations of immune cell measurements that predict functional responses, one
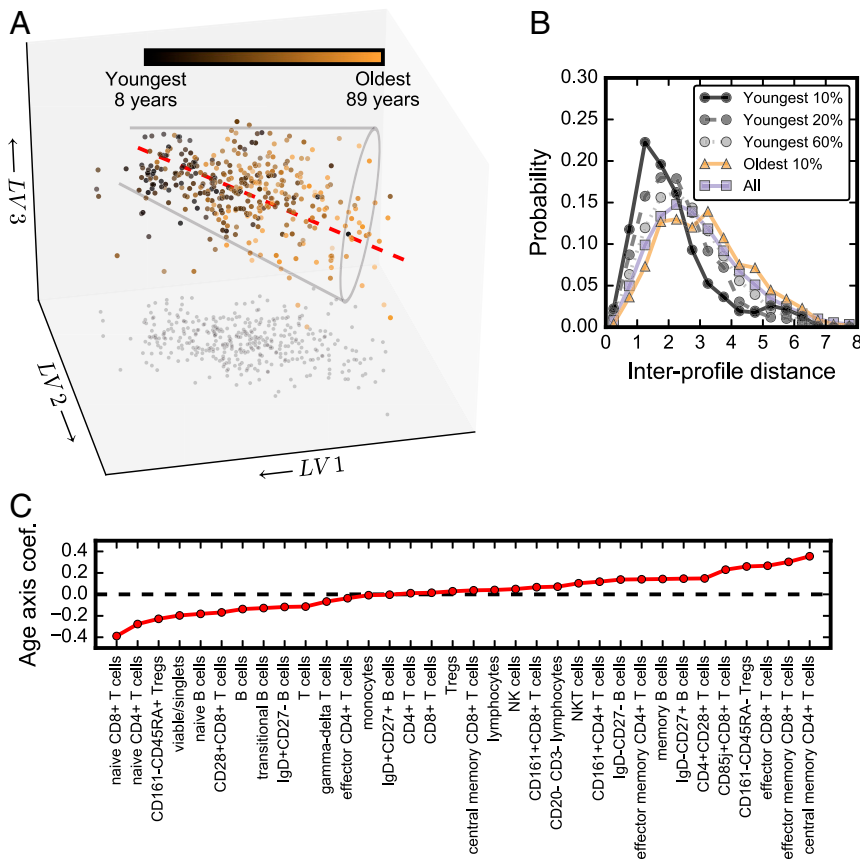
Fig. 4. Heterogeneity in immunotypes increases with age. This is defined by the top three LVs, averaged over group 1 responses (Dataset S5). (*A*) Individuals plotted in the immunospace defined by the first three PLS LVs. Individuals are colored in gradient from youngest (black) to oldest (orange) by rank order. The dashed red line indicates an "immunological age axis," the direction of gradient in age (found by logistic regression of age on the top three LVs). The plots contain grayscale "shadows" projected onto the LV1–LV2 plane to aid 3D visualization, as well as a transparent cone to highlight shape of point cloud. (*B*) Probability distributions of interpatient Euclidean distances defined on top three LVs for different age groups (youngest 10%, 20%, and 60% corresponding to ages <15.6 y, 20 y, and 44 y, respectively, and the oldest 10% corresponding to ages >75 y). Twin pairs are excluded from all three distributions. (*C*) Regression coefficients (normalized to unit vector) of immune cell populations that define the immunological age axis of *A*, where populations are ordered by size of coefficient.

confounds potentially heritable and nonheritable influences that shape these immunological parameters. Together with the assumption that twin pairs share many of the same environmental influences, at least early in life, we expect twin pairs to have more similar immunotypes than randomly selected pairs of individuals. Indeed, both MZ and DZ twin pairs have lower interindividual distances (Euclidean distance between immunotypes) compared with randomly chosen pairs of individuals in our cohort (Fig. 5); the random pairs of individuals were chosen with an age distribution comparable to the twin pair age distribution. The interindividual distances were lower in MZ twins than in DZ twins, illustrating the influence of both heritable and nonheritable factors for determining immunotypes (Fig. 5). This result confirms previous results and, importantly, also illustrates that the immunotypes we have defined on the basis of immune cell composition and functional responses provide an accurate representation of interindividual immune variation.

HCMV Seropositivity Is Associated with a Shift Toward an Older Immunotype Across All Age Groups. Given the dominance of nonheritable factors in shaping human immune systems, investigating the specific factors of importance and how they influence immunotypes is important. HCMV is known to affect the majority of immune system components (4). To this end, we reasoned that it would be interesting to compare the immunotypes of HCMV-seropositive and HCMV-seronegative individuals. Graphs analogous to Fig. 4*A* for HCMV serostatus (Fig. S5*B*) do not reveal any clear separation of individuals based on immune cell composition alone. In fact, if the LVs are used to correlate HCMV serostatus (using cross-validated logistic regression), many more LVs are required than those needed to describe the effects of aging (using cross-validated ridge regression), as shown in Fig. S5*C*. Fig. S5*D* indicates that the metadata alone (age, gender, HCMV serostatus)

are insufficient to accurately predict these functional responses, and adding them to the immune cell composition data does not improve model performance, suggesting that the information provided by the metadata is less predictive than the immunotypes we have defined. This is best exemplified by the great heterogeneity seen in older individuals (Fig. 4*A*), whose responses can be predicted by a person's immunotype, but not by age.

Several studies have suggested that HCMV may promote accelerated immunological aging, and it has been implicated as a driver of the immunosenescence in the elderly that is associated with an increase in all-cause mortality (32). HCMV serostatus has not been found to influence influenza vaccine responses in the elderly (32–36), but in young adults and mice, HCMV and mouse CMV (MCMV), respectively, enhance influenza immunity (33). To test whether HCMV seropositivity was reflected by a change in immunotype, we projected each individual's cell composition onto the line within the space defined by the top three LVs that best correlates with age (Fig. 4*A*). This line represents an "age axis" in this space. The distributions of individual cell compositions along this line are graphed in Fig. 6*A* (grouped by actual age and HCMV serostatus). Fig. 6*A* shows that across all age groups, the distribution of HCMV seropositive individuals is shifted upward along the age axis. Because this association is true for individuals with the same actual age, we conclude that HCMV seropositivity does induce a shift in immunotypes toward one associated with older individuals. The study of individuals of similar age eliminated the confounding effect of older individuals having a greater chance of exposure to the virus and more rounds of viral reactivation. This finding is important in relation to the ongoing discussion about the role of HCMV during immunosenescence (36) and suggestive of a much broader role for HCMV, visible also in individuals as young as adolescents when viewed from the perspective of their immunotypes.
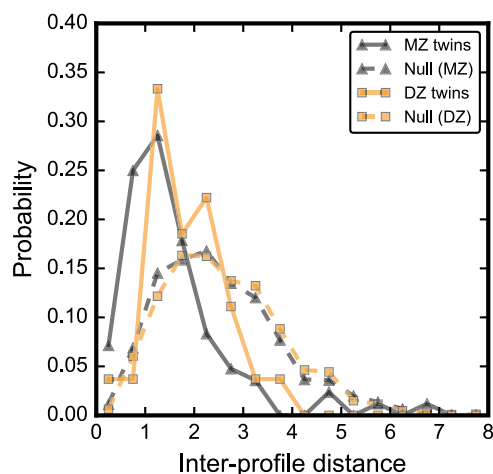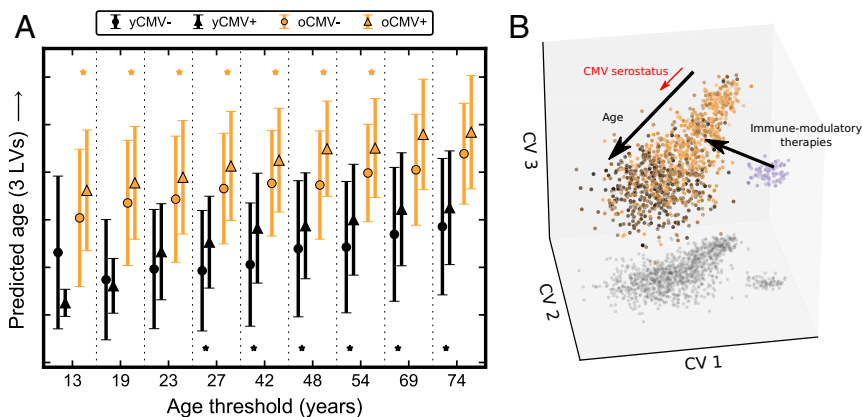
Fig. 5. Twin pairs have similar immunotypes. Shown are interindividual distances (Euclidean distance on the top three LVs from PLS) between MZ twin pairs and DZ twin pairs, along with appropriate null models. The null models are constructed from randomly chosen pairs of individuals in the cohort restricted to a similar age distribution as exhibited by the MZ twin pairs or DZ twin pairs, respectively, as described in *SI Materials and Methods, Twin Pair Immunotype Similarities Null Model*.

Discussion

Here we present a conceptual framework for analyzing human immune systems in terms of "immunotypes" within a space defined by axes that represent key combinations of immune cell frequencies. First, we find that individuals' immunotypes are distributed continuously across several healthy populations and are predictive of diverse functional immune responses as shown by in vitro cell signaling results, and these responses cannot be predicted from individual immune cell frequency measurements alone. Next, we show that specific factors, most notably age and HCMV serostatus, are associated with reproducible shifts in immunotype across unrelated, healthy individuals (shown schematically in Fig. 6B). Furthermore, by analyzing the distribution of individuals' immunotypes, we show that the immune systems of unrelated young individuals are more similar to each other, whereas older individuals are characterized by heterogeneous immunotypes. Our results suggest that an individual's immune cell composition may be globally modifiable by interventions to transform it from a less responsive to a more responsive immunotype.

Our findings are in accordance with recent data on the dominant influence of nonheritable factors in shaping human immune systems (4–6, 10). They further support previous findings suggesting that environmental influences exert a cumulative influence over time, leading to the divergence of immune phenotypes between MZ twins (4). The finding that collective sets of immune cell frequency measurements predict diverse functional responses is interesting and suggests that the composition of an individual's circulating immune system reflects a functional network of cell populations and that the balance between these cell populations determines the overall responsiveness of the system. It also suggests that modulating this balance and perturbing the network, for example by depleting one specific cell type, can have broad implications for the functional responsiveness of all remaining cells. This is an important observation given that such depleting strategies are routine practice in current medical practice. An example of a manifestation of this hypothesis is the clinical observation of delayed loss of neutrophils upon depletion of CD20-expressing B cells by an anti-CD20 antibody (Rituximab) (37). Another important implication of this is that the infusion of specific cell populations, such as tumor-specific lymphocytes or populations of stem cells infused for a range of clinical conditions, could impact the global network of immune cell populations with unpredictable consequences for the immune competence of the individual patient.

The relationship between immune cell measurements and metadata—age, gender, and HCMV serostatus—also adds to our understanding of human immune system variation. The fact that these metadata alone cannot accurately predict functional responses shows the need for collective indicators of immune health (immunotypes) such as those defined in Fig. 2 C–E. Second, a description of an aging immune system requires far fewer LVs than does gender or HCMV serostatus, suggesting that these characteristics of an individual play a more complex role in regulating the immune system. The fact that older individuals are increasingly heterogeneous with respect to their immunotypes also makes age insufficient as a predictor of immune responsiveness.

Despite the potential our conceptual framework has to pose and answer important questions in immunology, it is necessary to address possible improvements in our results and methods that can be addressed in future work. First, we emphasize that PLS regression is only one method for choosing the collective variables with which to define immunotypes. Extending this method using kernels to account for nonlinear interactions between immune parameters, in addition to introducing regularization into the LV decomposition (with a sufficiently large cohort size to allow

Fig. 6. HCMV seropositive individuals have immunotypes shifted toward that characteristic of older populations. (A) Individuals in the cohort are projected onto the immunological age axis (predicting age from the top three LVs) from Fig. 4A. The distributions of positions on this line are compared between HCMV seronegative and seropositive (circles and triangles, respectively) individuals for young and old (y and o, respectively) age groups defined at various age thresholds along the x axis (mean ± SD). Age thresholds were chosen as percentiles 10–90 in increments of 10 based on the ages of individuals in the Stanford cohort. Asterisks indicate statistical significance ($P < 0.05$ according to a two-tailed t test) between HCMV seropositive or seronegative individuals for the young (black) and old (orange) groups. (B) Schematic illustrating immunotypes, represented as a continuously distributed point cloud. An additional small cluster in purple indicates a hypothetical (as-yet undiscovered) disease-associated region of immunospace. Grayscale shadows are projected into the CV1–CV2 plane to aid 3D visualization. Arrows corresponding to age and HCMV serostatus emphasize that external factors exist that influence individuals' immunotypes throughout their lifetime. The arrow from the disease-associated to the "healthy" region of immunospace represents the potential for the proposed immunotype framework to guide the development of novel therapies that modulate immunotypes in a desired direction (Fig. S5).

unbiased fitting and evaluation of the regularization parameter), would likely lead to more robust definitions of immunotypes. Next, we acknowledge that predictions of our functional immune responses from immunotype, although statistically significant, are moderate correlations. This is likely due to a number of reasons, the most prominent being that the 34 immune cell populations (which are restricted to peripheral blood subsets) do not afford a complete definition of immunotypes, and thus we cannot expect perfect predictions. However, more extensive immune system measurements (serum cytokines, chemokines and growth factors, antibodies, including information from additional tissue compartments, etc.), as well as taking other factors, such as the symbiotic microbiota into account, may lead to even more reliable definitions of an individual's immunotype. Related to this, an expansion of immune measurements would also increase the potential for front loading of explanatory power into the top three LVs, covering a comparatively larger fraction of the global immune system variation in humans.

It is also important to note that our finding of a continuous distribution of immunotypes may also be sensitive to the specific immune cell populations measured in this study. An expansion of the immune measurements, as described above, could potentially reveal discrete clustering of individuals, especially if the additional immune parameters themselves exhibit discrete distributions. In addition, it is always the case that in clustering analysis, finding no evidence of discrete clusters does not prove their absence, although we are confident given our thorough analysis that discrete clusters are not present in these datasets. However, we also predict that this result could change if our cohorts were expanded to include, in addition to the healthy individuals studied here, individuals with perturbed immunotypes—e.g., patients treated with immunosuppressive drugs or individuals suffering from primary and secondary immunodeficiency syndromes with severe infectious diseases susceptibility and autoinflammatory symptoms may form a distinct cluster. We predict that these individuals may be qualitatively much different from the healthy individuals in the cohorts we studied, the discovery of which would be a next step for development of immune-modulatory therapies.

A possible criticism of the results we present here is that responses to cytokine stimulation in vitro may not be reflective of actual immune system competence that determines an individual's risk of immune-mediated disease or severe infectious diseases. However, several lines of evidence suggest that the broadly different signaling responses, mediated via different signaling pathways, are reliable correlates of global immune competence. For example, in a recent study, Shen-Orr et al. (11) found that defective responsiveness in the JAK-STAT pathway to the same type of in vitro stimuli was strongly predictive of cardiovascular disease in a cohort of elderly individuals.

It will be important to test the concepts revealed by studying the cohorts considered in this paper by carrying out similar analyses in additional cohorts for which both immune cell compositions (or other markers) and responses to stimuli have been measured. We anticipate that the concept of continuous immunotypes predictive of immune responses will be further validated. The precise definitions of the key combinations of cell types (LVs) that define immunotypes may be cohort dependent. Indeed, it will be interesting to study how the precise definition of the LVs depends upon diet, infection history, etc., by studying cohorts of different geographical and environmental contexts whose experiences in this regard are expected to be distinct from those of the Stanford cohorts described here. Finally, the conceptual framework we have described should be augmented by a mechanistic understanding of why and how a few specific combinations of immune cell populations define an individual's immunotype.

By addressing the above limitations in our method, we believe that our findings and the framework presented here could guide future developments of better immunomodulatory therapies de-signed to induce perturbations of an individual's immunotype rather than individual immune components (Fig. 6B). This framework could also be of particular importance for the development of improved vaccines in the future, taking immunotypes into account, perhaps personalizing adjuvants and vaccine formulations as well as optimizing the time for vaccination to improve responsiveness. Also within the field of cancer immunotherapy, the induction of antitumor immune responses could be improved by stimulating immunotypes rather than individual cells to improve responsiveness. As more data become available and immunotype definitions become more robust, modeling of necessary perturbations to elicit a desired response would become possible. We envision that such a framework would make the responsiveness of individual patients to such therapies more predictable because compensatory relationships between individual cell populations are taken into account in a way in which more traditional analyses, focusing on individual cell populations only, do not. Thus, the conceptual framework we have described may have important implications for predicting immunological health and the risk of disease.

Materials and Methods

Cohorts. In this work, we analyzed human immune cell frequency data from three cohorts. The Stanford cohort consisted of 398 individuals, both men and women, between ages 8 y and 89 y of age, sampled at the Stanford University clinical trials unit. The study protocol was approved by the Stanford University Administrative Panels on Human Subjects in Medical Research, and written informed consent was obtained from all participants. A total of 226 of these individuals were twin subjects from both MZ and DZ pairs from the Twin Research Registry at SRI International as previously described (4). A total of 23 additional subjects were included from a longitudinal study of aging in the immune system (11) and 149 previously unpublished subjects were included as well (combined in Datasets S1–S3). All demographic information of the subjects in the Stanford cohort is available in Dataset S1. The Stanford cohort comprises individuals from the following ClinicalTrials.gov study IDs: NCT01827462, NCT03020498, NCT03022396, NCT03022422, and NCT03022435. The Roederer et al. (10) cohort consists of 669 female twins sampled in the United Kingdom. The Carr et al. (6) cohort consists of 670 nontwin subjects, both men and women sampled in Belgium.

Immune Cell Population Frequency Measurements. All three cohorts analyzed PBMCs, using either flow cytometry or mass cytometry. The immune cell composition data for the Stanford cohort were generated by the Human Immune Monitoring Center, using mass cytometry and antibody panels (4, 11). All cells were classified by manual gating, using the FlowJo Software (ThreeStar Inc.), and gating schemes are presented in the original publications. The PBMCs from the Roederer et al. (10) and Carr et al. (6) cohorts were analyzed by a combination of consecutive flow cytometry panels and manual gating.

Functional Response Measurements. Frozen PBMCs were thawed and stimulated, with one of seven different cytokines (IL-2, IL-6, IL-7, IL-10, IL-21, IFN-α, or IFN-γ) for 30 min; fixed; and stained with antibodies to surface antigens (identifying eight responding cell populations: monocytes, B cells, CD4$^+$, CD4$^+$CD45RA$^+$, CD4$^+$CD45RA$^-$, CD8$^+$, CD8$^+$CD45RA$^+$, and CD8$^+$CD45RA$^-$ T cells) as well as intracellular phosphor-STAT1, -3, and -5, respectively, as described in detail previously (4). The 90th percentile fluorescence intensity of stimulated samples was compared with the 90th percentile in PBS-treated cells (unstimulated) as a fold-change ratio used for downstream analyses (Dataset S2). In addition to the cell-signaling responses, we also include an analysis of antibody responses to seasonal flu vaccines (years 2010–2012) (Dataset S3). The vaccine administered was the Fluzone quadrivalent vaccine (Sanofi-Pasteur Inc.) and the vaccine-induced antibody responses were measured in a strain-specific manner, using HAI assays, and fold changes day 28/day 0 were used for analyses as described in detail in ref. 4. The flu-vaccine data included eight strains: A/Texas/50/2012, B/Massachusetts/2/2012, B/Brisbane/60/2008, A/California/07/2009(H1N1), A/Perth/16/2009(H3N2), A/South Dakota/06/2007(H1N1), A/Uruguay/716/2007(H3N2), and Wisconsin. The A/Texas/50/2012, B/Massachusetts/2/2012, and Wisconsin strains were removed due to small sample size (Data Preprocessing).

Data Preprocessing. Of the 136 cell population frequencies measured in the Stanford cohort, 34 were considered in this study. We removed 89 cell populations because they were measured in only 10.3%, 29.1%, 39.4%, or 70.8% of the

398 patients. We also removed the following 13 subtypes because they either are rare or are the complement in a set of rare cell populations: plasmablasts, HLADR$^+$ NK cells, CD161$^+$CD45RA$^+$ Tregs, CD161$^+$CD45RA$^-$ Tregs, CD85j$^+$CD4$^+$ T cells, HLADR$^+$CD38$^+$CD4$^+$ T cells, HLADR$^+$CD38$^+$CD8$^+$ T cells, HLADR$^+$CD38$^-$CD4$^+$ T cells, HLADR$^+$CD38$^-$CD8$^+$ T cells, HLADR$^-$CD38$^+$CD4$^+$ T cells, HLADR$^-$CD38$^+$CD8$^+$ T cells, HLADR$^-$CD38$^-$CD4$^+$ T cells, and HLADR$^-$CD38$^-$CD8$^+$ T cells. To remove measurement dependence on measurement year (2009–2012), we applied the ComBat algorithm (38), an empirical Bayes method for removing nonbiological experimental variation that has been found to outperform other batch effect correction methods according to many metrics (39). Patient 398 was removed due to missing batch information, leaving a total of 397 patients studied from the Stanford cohort. A summary of the immune cell population frequencies and demographics data for the Stanford cohort is available in Dataset S1.

Functional responses having fewer than 34 patients (equal to the number of retained cell populations) were removed from consideration: three influenza strains (A/Texas/50/2012, B/Massachusetts/2/2012, and Wisconsin), as well as all 21 cytokine stimulation responses measured in the "Non-BT" cell subtype.

Clustering Methods. The Stanford, Roederer et al. (10), and Carr et al. (6) cohorts included composition measurements of nonidentical immune cell populations (see Table S1 for a list of cell subtypes in each cohort), thus precluding a combined clustering analysis of the full dataset. The "monocytes" cell population was removed from the Roederer et al. (10) cohort as it was missing in 7.6% of the 729 patients. Of the remaining 32 populations, 11 had missing data in at most 1.6% of patients. For the Carr et al. (6) cohort, patients containing >20% missing measurements were removed, leaving a total of 449 patients. Cell populations with >10% missing values across patients were removed from consideration. Of the remaining 32 populations, 5 had missing data in at most 4% of patients. Missing values in both cohorts were imputed to the average across all patients.

Clustering analysis was performed on the PC scores (projections of immune cell composition data onto the eigenvectors of the correlation matrix—this places all immune cell population frequencies on the same scale). Clustering of individuals was performed on the top PCs with a k-means clustering algorithm implemented using the sklearn.cluster.KMeans() class in scikit-learn v0.17 (40), using interpoint Euclidean distance in PC space. The quality of clusters from this algorithm was quantified using two metrics:

i) The silhouette coefficient (41) is defined for each data point, i, as $(b_i - a_i)/\max\{a_i, b_i\}$ in which a_i is the average pairwise Euclidean distance between data point i and all other points within the same cluster as i, whereas b_i is the lowest average pairwise Euclidean distance between data point i and all data points in any other cluster. The silhouette coefficient ranges from -1 to 1, with a high value indicating good clustering.

ii) The between-cluster fractional explained variance is defined as the variance in the data where each point is replaced by the average of the points in the same cluster, divided by the total variance, V^{tot}, in the data (equal to the sum of between-cluster and within-cluster variance, $V^{b/w}$ and $V^{w/i}$, respectively). More formally, let $\mathbf{x}_i = \left(x_i^{(1)} \quad x_i^{(2)} \quad \ldots \quad x_i^{(d)} \right)^{\mathsf{T}}$ be a vector of PC scores (PCs $1 - d$) for patient i and $\boldsymbol{\mu} = \sum_{i=1}^{N} \mathbf{x}_i/N$ be the mean over all N patients. Now, let the clustering algorithm assign each patient to one of C clusters, $c_i \in \{1, \ldots, C\}$, and let $\mathbf{X}_c$ be a matrix composed of the subset of rows of X corresponding to patients assigned to cluster c; i.e., the rows of $\mathbf{X}_c$ are the set $\{\mathbf{x}_i : c_i = c\}$. Finally, let $\boldsymbol{\mu}_c = \frac{\mathbf{1}^{\mathsf{T}} \mathbf{X}}{\#\{i : c_i = c\}}$ be the cluster average, where $\mathbf{1}$ is a vector of ones. Then, the between-cluster fractional explained variance, $V^{b/w, frac}$, is calculated as follows:

$$V^{tot} = \mathrm{tr}\left[(\mathbf{X} - \boldsymbol{\mu})^{\mathsf{T}} (\mathbf{X} - \boldsymbol{\mu}) \right]$$

$$V^{w/i} = \sum_{c=1}^{C} \mathrm{tr}\left[(\mathbf{X}_c - \boldsymbol{\mu}_c)^{\mathsf{T}} (\mathbf{X}_c - \boldsymbol{\mu}_c) \right]$$

$$V^{b/w, frac} = \frac{V^{Tot} - V^{w/i}}{V^{tot}} = \frac{V^{b/w}}{V^{tot}}.$$

Because these measures of cluster performance are difficult to interpret alone, they are compared with a null model formed by the following: (*i*) For each PC, i, in the original data, calculate the difference, a_i, between the 97.5th and 2.5th percentiles of the scores of that PC. This gives a robust-to-outliers measure of the spread in each PC. (*ii*) Choose a number of PCs, d, to include in the clustering. (*iii*) Generate a uniform random sample of profiles within the PC space contained within a d-dimensional ellipsoid with principal axes equal to the respective differences calculated in i. If x_i represents the randomly generated value for the PC i score, the null model satisfies

$$\sum_{i=1}^{d} \frac{x_i^2}{(a_i/2)^2} < 1.$$

This null model was chosen because the PC projections appear ellipsoid with apparent uniform density of points except at the outskirts (although with a fatter tail than a multivariate Gaussian). Thus, using a uniform density for the null model is conservative (i.e., it will tend to give a larger number of clusters than the real data).

These analyses are presented in Fig. S1, along with corresponding t-distributed stochastic neighbor embedding (t-SNE) plots, implemented using the sklearn.manifold.TSNE() class in scikit-learn v0.17 (40) with two components, perplexity of 30.0, and learning rate of 1,000.0. The t-SNE method for visualizing high-dimensional data overcomes some of the limitations of the linear PCA method (20).

PLS Regression. Formally, we sought to develop a set of regression models that separately predict the values of each functional response, $\mathbf{y}^{(k)}$, given the immune cell compositions, $\mathbf{X}$, of a group of individuals. Here, element $y_i^{(k)}$ of vector $\mathbf{y}^{(k)}$ represents the value of response k for individual i, whereas element X_{ij} refers to the immune cell population frequency j in individual i. Linear models assume that the responses and predictors are related through the functional form $\hat{\mathbf{y}}^{(k)} = \mathbf{X}\boldsymbol{\beta}^{(k)}$, where the "hat" distinguishes the predicted responses from the experimentally measured ones, and $\boldsymbol{\beta}^{(k)}$ is the regression vector for functional response k. The elements of the regression vector are the weights of each cell population that define a linear combination that best predicts the functional response.

For the regression models, the functional responses were first transformed using the log-modulus transformation (42) to preserve rank ordering of responses while minimizing the effect of large outliers. For each functional response, PLS (Dataset S4) is used to find a set of linear combinations (termed the LVs) of cell populations that have the highest covariance with the functional response. These LV signatures, $\mathbf{r}_j^{(k)}$, are the directions in the space of possible immune cell compositions that explain a particular functional response, ordered from best to least, as columns of the LV matrix $\mathbf{R}^{(k)}$. In this way it is possible that only a small number of such directions need to be identified to predict functional responses well. The advantages of using PLS are its simplicity, the avoidance of multicollinearity issues resulting from highly correlated immune cell populations (SI Materials and Methods, Regression Methods), and the conceptual advantage afforded by grouping cell populations together into latent variables that collectively explain a response of interest. We compared the performance of PLS with PCR, in which directions in the space of possible immune cell compositions are chosen to explain the greatest variance in immune cell composition across patients.

The regression models are evaluated using learning curves of 10-fold cross-validated (CV) negative Spearman correlation coefficients between observed and predicted response measurements for both PLS and PCR regression models. As described in ref. 43, cross-validation estimates out-of-sample model performance for smaller datasets by fitting a model on only a fraction of the available data and testing predictions on the remaining data left out. We used a 10-fold cross-validation scheme based on recommendations by Hastie et al. (43). The dataset is randomly partitioned into 10 groups and a PLS model separately fitted on each possible combination of 9/10 of those groups. The predictions on the left-out groups are compared with their experimental values, using a negative Spearman correlation coefficient. Learning curves are made by calculating CV error as a function of the number of LVs included in the model. The default error (when no LVs are used) is a correlation coefficient of zero. Error bars on learning curves are generally reported as SEs (44). However, because our measure of error combines errors over the entire dataset, we report error bars as SD of error calculated over 100 random instantiations of the CV learning curve (corresponding to different random 10-fold partitioning of the dataset).

We wish to verify that the PLS model is fitting meaningful correlations between both the immune markers and response functions and that improved model performance is not an artifact of the model-fitting procedure. To address these concerns, we evaluate the performance of a "null model," created by random reassignment of responses across patients. This randomization eliminates the relationship between a patient's immune cell composition and the value of the corresponding response, and the model-fitting procedure should correctly detect that there is no underlying structure to be learned from the data. We calculate 95% confidence intervals of the null model learning curve, using 100 instantiations of the null model to summarize how variable the estimates from the null model are. Responses for which the minimum error plus error bar come within 0.05% of the 95% confidence interval of the null

model learning curve are removed from further consideration. Six signaling responses (IL-6/Mono/STAT5, IL-7/CD4$^+$CD45RA$^+$/STAT1, IL-7/CD4$^+$CD45RA$^-$/STAT1, IL-7/CD4$^+$/STAT1, IFN-γ/CD8$^+$CD45RA$^-$/pSTAT3, and IFN-γ/CD4$^+$CD45RA$^+$/pSTAT3) and four of the five flu-vaccine responses [A/California/07/2009(H1N1), A/Perth/16/2009(H3N2), A/South Dakota/06/2007(H1N1), and A/Uruguay/716/2007(H3N2)] were so removed.

Finally, to indicate the amount of front loading of explanatory power into the top LVs, we also report a normalized error defined as follows: If the minimum observed error for a given functional response in either PLS or PCR is denoted e_{min}, and the raw model error is denoted e, the normalized error is defined as $e_{norm} = (e - e_{min})/(0.0 - e_{min})$, where an error of 0.0 indicates a degenerate model in which the mean of the response measurements is predicted regardless of an individual's immune cell composition.

As discussed in the main text, the PLS model for the HAI response to the influenza B/Brisbane/60/2008 strain was fitted on only the subset of individuals who had a measurement increase in HAI titer upon vaccination (31% of patients with a post- to prevaccination HAI titer ratio of 1 were not included in the fit). Furthermore, the fit benefits from further outlier moderation; seven patients with log-modulus HAI titer ratios greater than 4.0 (corresponding to HAI titer ratios >53.6) were adjusted to this 4.0 threshold for further outlier moderation. These seven individuals had HAI titer ratios equal to 64, 80 (for three individuals), 160, 320, and 640. The average HAI titer ratio for the remaining individuals was 11.3.

1. Maecker HT, McCoy JP, Nussenblatt R (2012) Standardizing immunophenotyping for the Human Immunology Project. *Nat Rev Immunol* 12:191–200.
2. Bandura DR, et al. (2009) Mass cytometry: Technique for real time single cell multi-target immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 81:6813–6822.
3. Tsang JS, et al.; Baylor HIPC Center; CHI Consortium (2014) Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* 157:499–513.
4. Brodin P, et al. (2015) Variation in the human immune system is largely driven by non-heritable influences. *Cell* 160:37–47.
5. Orrù V, et al. (2013) Genetic variants regulating immune cell levels in health and disease. *Cell* 155:242–256.
6. Carr EJ, et al. (2016) The cellular composition of the human immune system is shaped by age and cohabitation. *Nat Immunol* 17:461–468.
7. Rivera A, Siracusa MC, Yap GS, Gause WC (2016) Innate cell communication kick-starts pathogen-specific immunity. *Nat Immunol* 17:356–363.
8. Murphy K, Travers P, Walport M, Janeway C (2012) *Janeway's Immunobiology.* (Garland Science, New York), 8th Ed.
9. Chaiwatanatorn K, Lee N, Grigg A, Filshie R, Firkin F (2003) Delayed-onset neutropenia associated with rituximab therapy. *Br J Haematol* 121:913–918.
10. Roederer M, et al. (2015) The genetic architecture of the human immune system: A bioresource for autoimmunity and disease pathogenesis. *Cell* 161:387–403.
11. Shen-Orr SS, et al. (2016) Defective signaling in the JAK-STAT pathway tracks with chronic inflammation and cardiovascular risk in aging humans. *Cell Syst* 3:374–384.e4.
12. Aguirre-Gamboa R, et al. (2016) Differential effects of environmental and genetic factors on T and B cell immune traits. *Cell Reports* 17:2474–2487.
13. Thomas S, et al.; Milieu Intérieur Consortium (2015) The Milieu Intérieur study - an integrative approach for study of human immunological variance. *Clin Immunol* 157:277–293.
14. Duffy D, et al.; Milieu Intérieur Consortium (2014) Functional analysis via standardized whole-blood stimulation systems defines the boundaries of a healthy immune response to complex stimuli. *Immunity* 40:436–450.
15. De Jager PL, et al. (2015) ImmVar project: Insights and design considerations for future studies of "healthy" immune variation. *Semin Immunol* 27:51–57.
16. Nakaya HI, et al. (2011) Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol* 12:786–795.
17. Postow MA, Callahan MK, Wolchok JD (2015) Immune checkpoint blockade in cancer therapy. *J Clin Oncol* 33:1974–1982.
18. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD (2016) The future of cancer treatment: Immunomodulation, CARs and combination immunotherapy. *Nat Rev Clin Oncol* 13:273–290.
19. Jolliffe I (2014) Principal component analysis. *Wiley StatsRef: Statistics Reference Online*, eds Balakrishnan N, et al. (Wiley, Chichester, UK), pp 1–21.
20. van der Maaten L, Postma E, van den Herik J (2009) Dimensionality reduction: A comparative review. *J Mach Learn Res* 10:66–71.
21. Shekhar K, Brodin P, Davis MM, Chakraborty AK (2014) Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE). *Proc Natl Acad Sci USA* 111:202–207.
22. Amir AD, et al. (2013) viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat Biotechnol* 31:545–552.
23. Evans DM, et al. (2004) A major quantitative trait locus for CD4-CD8 ratio is located on chromosome 11. *Genes Immun* 5:548–552.
24. Strindhall J, et al. (2013) The inverted CD4/CD8 ratio and associated parameters in 66-year-old individuals: The Swedish HEXA immune study. *Age* 35:985–991.
25. Luz Correa B, et al. (2014) The inverted CD4:CD8 ratio is associated with cytomegalovirus, poor cognitive and functional states in older adults. *Neuroimmunomodulation* 21:206–212.
26. Galel SA, Lifson JD, Engleman EG (1995) Prevention of AIDS transmission through screening of the blood supply. *Annu Rev Immunol* 13:201–227.
27. Rosipal R, Krämer N (2006) Overview and recent advances in partial least squares. *Subspace, Latent Structure and Feature Selection*, eds Saunders C, et al. (Springer, Berlin), Vol 3940, pp 34–51.
28. Villarino AV, Kanno Y, Ferdinand JR, O'Shea JJ (2015) Mechanisms of Jak/STAT signaling in immunity and disease. *J Immunol* 194:21–27.
29. Sobolev O, et al. (2016) Adjuvanted influenza-H1N1 vaccination reveals lymphoid signatures of age-dependent early responses and of clinical adverse events. *Nat Immunol* 17:204–213.
30. Nakaya HI, et al. (2015) Systems analysis of immunity to influenza vaccination across multiple years and in diverse populations reveals shared molecular signatures. *Immunity* 43:1186–1198.
31. Pawelec G, Derhovanessian E, Larbi A, Strindhall J, Wikby A (2009) Cytomegalovirus and human immunosenescence. *Rev Med Virol* 19:47–56.
32. Furman D, et al. (2015) Cytomegalovirus infection enhances the immune response to influenza. *Sci Transl Med* 7:281ra43.
33. Looney RJ, et al. (1999) Role of cytomegalovirus in the T cell changes seen in elderly individuals. *Clin Immunol* 90:213–219.
34. Pawelec G, Derhovanessian E (2011) Role of CMV in immune senescence. *Virus Res* 157:175–179.
35. Sauce D, et al. (2009) Evidence of premature immune aging in patients thymectomized during early childhood. *J Clin Invest* 119:3070–3078.
36. Pawelec G, McElhaney JE, Aiello AE, Derhovanessian E (2012) The impact of CMV infection on survival in older humans. *Curr Opin Immunol* 24:507–511.
37. Voog E, Morschhauser F, Solal-Céligny P (2003) Neutropenia in patients treated with rituximab. *N Engl J Med* 348:2691–2694.
38. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8:118–127.
39. Chen C, et al. (2011) Removing batch effects in analysis of expression microarray data: An evaluation of six batch adjustment methods. *PLoS One* 6:e17238.
40. Pedregosa F, et al. (2011) Scikit-learn: Machine learning in Python. *J Mach Learn Res* 12:2825–2830.
41. Rousseeuw PJ (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 20:53–65.
42. John JA, Draper NR (1980) An alternative family of transformations. *Appl Stat* 29:190–197.
43. Hastie T, Tibshirani R, Friedman JH (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics (Springer, New York), 2nd Ed.
44. Murphy KP (2012) *Machine Learning: A Probabilistic Perspective*, Adaptive Computation and Machine Learning Series (MIT Press, Cambridge, MA).
45. de Jong S (1993) SIMPLS: An alternative approach to partial least squares regression. *Chemom Intell Lab Syst* 18:251–263.