nature immunology

The cellular composition of the human immune system is shaped by age and cohabitation

Edward J Carr[1,8], James Dooley[2,3,8], Josselyn E Garcia-Perez[2,3], Vasiliki Lagou[2–4], James C Lee[5,6], Carine Wouters[3], Isabelle Meyts[3], An Goris[4], Guy Boeckxstaens[7], Michelle A Linterman[1,9] & Adrian Liston[2,3,9]

Detailed population-level description of the human immune system has recently become achievable. We used a 'systems-level' approach to establish a resource of cellular immune profiles of 670 healthy individuals. We report a high level of interindividual variation, with low longitudinal variation, at the level of cellular subset composition of the immune system. Despite the profound effects of antigen exposure on individual antigen-specific clones, the cellular subset structure proved highly elastic, with transient vaccination-induced changes followed by a return to the individual's unique baseline. Notably, the largest influence on immunological variation identified was cohabitation, with 50% less immunological variation between individuals who share an environment (as parents) than between people in the wider population. These results identify local environmental conditions as a key factor in shaping the human immune system.

Enormous progress has been made in understanding the cellular and molecular components of the immune system. The key tool in this progression has been the use of mouse models, especially genetically modified models, which have allowed the functional dissection of the myriad of interconnecting components. Despite this progress, or perhaps because of it, some have argued convincingly that the focus of immunology should return to the human context[1] and include a comprehensive analysis of the full spectrum of immunological diversity and the causes thereof.

Several studies have embraced this call for investigation of human immunological diversity and have found that genetic factors account for ~25–50% of measured immunological variation[2–4]. A recent twin study indicated that at least half of immune trait variance is explained by nongenetic factors[3]. Similarly, the mean heritability of immune traits reported by a quantitative trait locus (QTL) study in healthy Sardinians was 41% (ref. 2). The ImmVar project, which tested for gene-expression QTLs in circulating human immune cells, estimated that ~22% of the variance in gene expression is explained by genetic factors[4]. Together, these studies suggest that the immunoprofile of the healthy population is governed in a large part by nongenetic factors. As nongenetic factors predominate in (and are arguably more amenable to) clinical manipulation, it is important to identify and quantify the key factors that shape the human immune landscape.

Using hypothesis-based approaches, researchers have identified several nongenetic factors that influence the landscape of the human immune system. Chronic infections, in particular latent infection with herpes virus, are associated with a panoply of immunological changes, and discordance for cytomegalovirus (CMV) seropositivity in monozygotic twin pairs results in weaker pairwise correlations for many immune parameters[3]. Aging represents another important nongenetic impact on the immune system, with potent effects on both the innate and adaptive arms of the immune response[5]. Within the adaptive system, aging is associated with a decline in naive T cells[6]. In mice, this is due to thymic involution; however, in humans, loss of naive $CD4^+$ T cells is driven primarily by failure of peripheral replication of naive cells[7,8], again demonstrating the importance of assessing human immunology directly rather than relying on the mouse (and, far too often, on a single inbred strain).

In the present study, we profiled the immune system of 670 healthy human volunteers, aged 2 to 86 years, to provide a description of the population-level heterogeneity in the cellular composition of the circulating immune system. Through the targeted recruitment of subcohorts with longitudinal sampling before and after severe immunological challenge, we determined that the immunological diversity between individuals is highly robust, with an elastic return to the unique steady state of the individual following immunological challenge. We found that coparenting profoundly reduced the immunological variation between two people, suggesting that environmental influences drive convergence as well as diversity within the human immunological profile.

RESULTS

Elasticity of the human cellular immune system

To investigate diversity in the composition of the human immune system, we developed an immune phenotyping platform quantifying

[1]Lymphocyte Signaling and Development ISP, Babraham Institute, Cambridge, UK. [2]Translational Immunology Laboratory, VIB, Leuven, Belgium. [3]Department of Microbiology and Immunology, University of Leuven, Leuven, Belgium. [4]Department of Neurosciences, University of Leuven, Leuven, Belgium. [5]Cambridge Institute for Medical Research, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK. [6]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK. [7]Department of Experimental Medicine, University of Leuven, Leuven, Belgium. [8]These authors contributed equally to this work. [9]These authors jointly directed this work. Correspondence should be addressed to A.L. (adrian.liston@vib.be).

Received 26 September 2015; accepted 9 December 2015; published online 15 February 2016; doi:10.1038/ni.3371

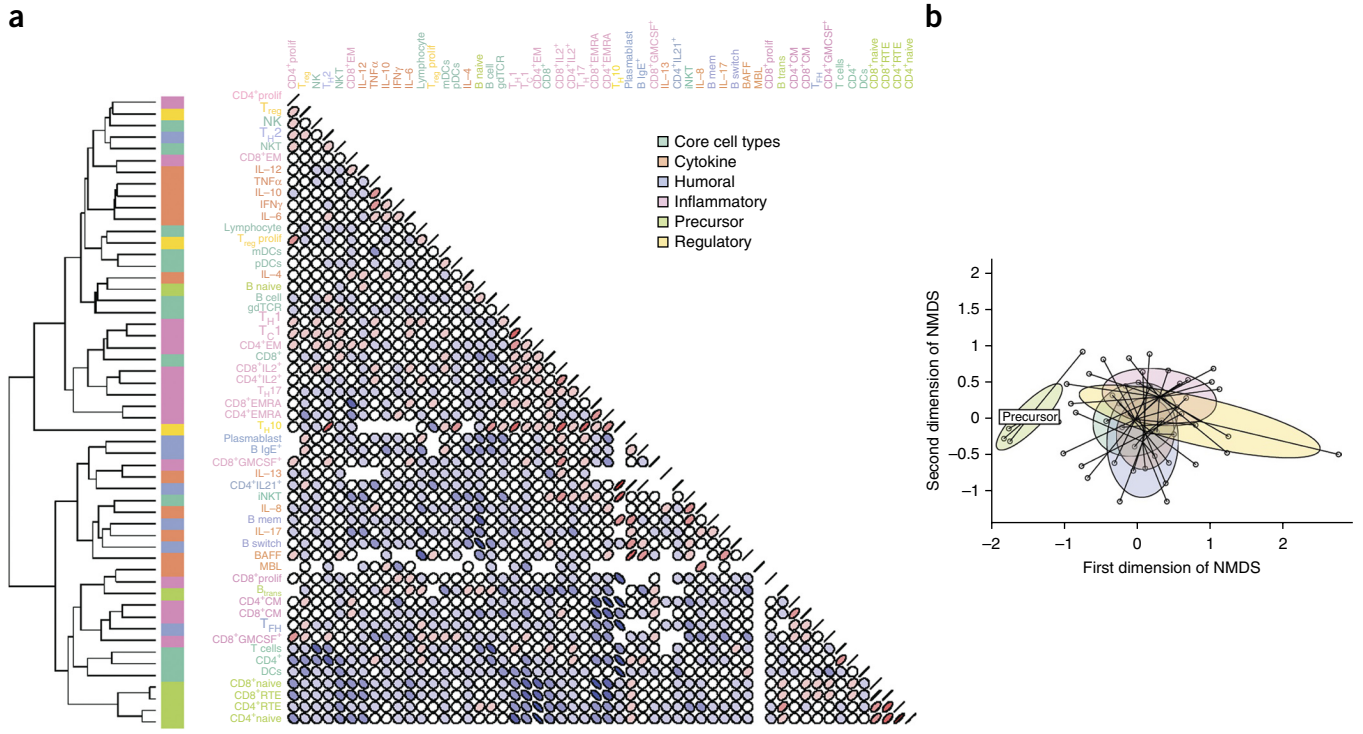

Figure 1 Data-driven analysis of immunological variation reveals biologically meaningful cocorrelations between individual immune parameters. (**a**) Left, dendrogram of immune parameters based on the complete data set of the most recent samples from 638 individuals, generated by hierarchical clustering on Euclidean distances of Spearman correlations between each parameter. Right, correlation plots using pairwise Spearman correlation coefficients between immunological parameters. Coefficients are described by the angle of ellipse (left-leaning, negative; right-leaning, positive) and color (blue, negative; red, positive). Manually annotated thematic groups of immune parameters are indicated by the color bar next to the dendrogram. (**b**) Nonmetric MDS (NMDS) of pairwise Spearman correlations. The data set is reduced from 54 immune parameters to a two-dimensional representation. Points indicate individual immune parameters; outlined areas represent thematic groups.

54 distinct immunological parameters by flow cytometry and serum analysis, with a focus on cellular subsets within the adaptive immune system. After optimization, we recruited 638 healthy Belgian individuals, ranging from 2 to 86 years of age and free from self-reported gastrointestinal, autoimmune or inflammatory disease, for immune profiling (**Supplementary Table 1**). Of these, 140 individuals were recruited as pairs (70 sets of coparents). We sampled 177 individuals at multiple time points (with an average of 6 months between samplings) to allow the measurement of longitudinal variation. Within this longitudinal cohort, we targeted 50 individuals who were planning to travel to areas where they would be at high risk of developing gastroenteritis and obtained pre- and post-travel samples from them. In total, we assessed 921 samples from 638 individuals over 3 years.

As an exploratory analysis of the data set, we examined the degree and structure of the variation. Substantial variation was observed in all immunological parameters measured (**Supplementary Table 2**). To determine whether there were underlying patterns within the variation, we performed unsupervised hierarchical clustering (**Fig. 1a**). The strongest clustering was observed between parameters manually annotated as 'precursor' cell types, with recent thymic emigrant (RTE) $CD4^+$ T cells, RTE $CD8^+$ T cells, naive $CD4^+$ T cells and naive $CD8^+$ T cells forming a single cluster (**Fig. 1**). This cluster was robust, being identified through iterative reclustering (**Supplementary Fig. 1**). Parameters within the other manually annotated groups (humoral, inflammatory, regulatory, core cell types, cytokines) were distributed throughout the hierarchy. Using multidimensional scaling (MDS), we identified, in a data-driven manner, cocorrelations between precursor populations (**Fig. 1b**). As precursor parameters separated from the rest of the parameters along the first dimension, this represents the largest source of variability between immune parameters, suggesting more coordinated biological control of these particular subsets. These data demonstrate that there is a high degree of variation in the immunological profiles of healthy individuals, with the largest component of the variation being a coregulated change in the frequency of naive or precursor cell types. Activated cell types and products, by contrast, demonstrated only minor coregulation (with several biologically relevant exceptions, such as between type 1 helper T (T_H1) and type 1 cytotoxic T (T_C1) cells).

To determine whether immunological variation represented a dynamic process of change within individuals or a spectrum of stable equilibria among individuals, we used data from the longitudinal subcohort of 177 individuals in an ANOVA model with two independent variables—the volunteer's unique identifier and the sample time point (**Fig. 2**). Each immune parameter (**Supplementary Table 2**) was used as the response variable in these models (**Fig. 2**). The majority of the variation in each parameter was explained by a model of stable intraindividual immune profiles over longitudinal sampling (**Fig. 2a**), with a median R^2 of 0.84 across the parameters (range 0.5–1.0) and 60% of models maintaining statistical significance (Bonferroni-corrected $P < 0.05$) after correction for multiple testing (**Fig. 2b**). Variation between repeat samples of individuals, by contrast, contributed very little to the observed total variation, with a median proportion of R^2 of 0.017 (range, 0.004–0.066), whereas interindividual variation explained a much larger effect, with a median proportion of R^2 of 0.983 (range, 0.934–0.996) (**Fig. 2c**). Thus, of the total variation observed within our data set, the majority (84%) can be explained

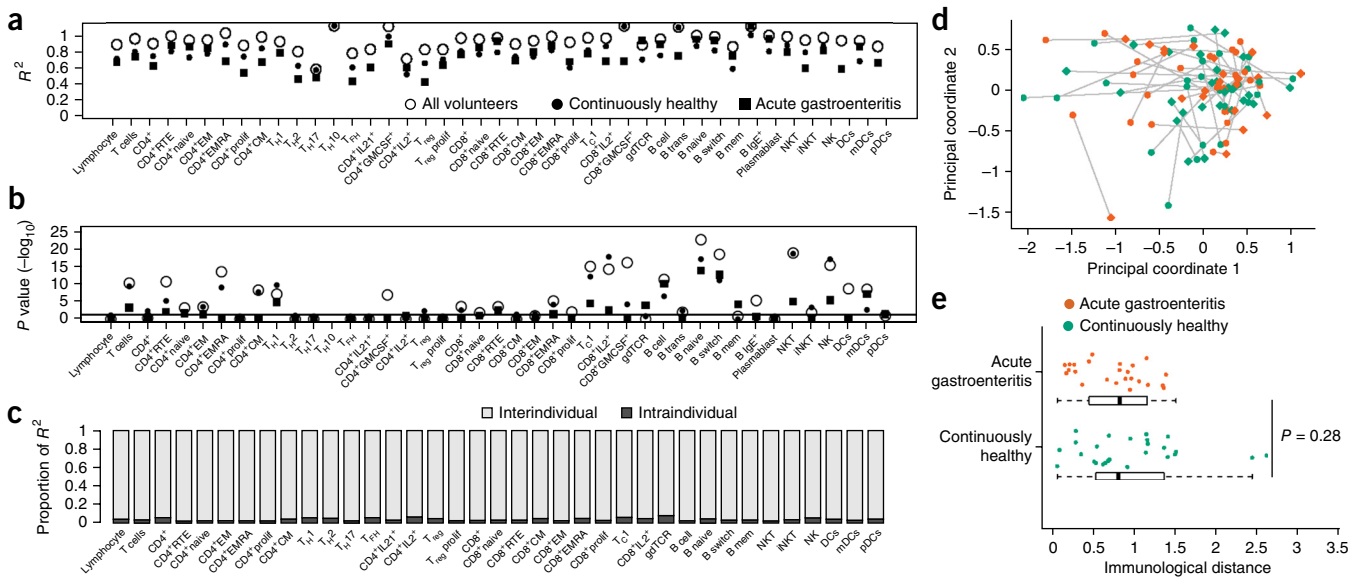

Figure 2 The human immune system is robustly maintained in multiple stable equilibriums. 177 individuals were sampled at least twice, allowing a dissection of inter- versus intraindividual variation. (**a**,**b**) Linear models showing R^2 values (**a**) and $-\log_{10}$ of Bonferroni-adjusted P values (**b**) for each immune parameter, based on several samples from each individual. Open circles represent models built using data from 638 individuals, 177 of whom were sampled several times (up to three visits), for a total of 921 visits. Filled circles represent models using only individuals ($n = 152$) who were continuously healthy between visits; filled squares represent models using only individuals ($n = 24$) who experienced acute gastroenteritis between visits. (**c**) Proportion of R^2 values from all volunteers attributable to interindividual or intraindividual differences. (**d**) MDS of pre- and post-travel visits. Each individual's first and second visits are shown as circles and diamonds, respectively; gray lines indicate immunological distance ($n = 50$ individuals; 100 visits). Green indicates continuously healthy individuals ($n = 26$); orange indicates individuals with intervening acute gastroenteritis ($n = 24$). (**e**) Quantification of the immunological distance between the first and second visits for continuously healthy individuals ($n = 26$) versus individuals with intervening acute gastroenteritis ($n = 24$). Two-tailed Mann-Whitney test was used to compare the immunological distances. Boxes and center lines represent interquartile range (IQR) and median, respectively. Whiskers extend to the farthest data point, up to a maximum of 1.5× IQR. Data beyond these bounds are plotted as points.

by a model that includes both interindividual and intraindividual variation, with just 1.4% attributable to the intraindividual variation between visits. This stability in the cellular immune profile of individuals over an extended sampling period is consistent with that observed by other flow cytometry–based studies[2,3,9] and is consistent with a diverse set of stable equilibria observed among individuals.

Having established that the relationship between immune subsets within an individual is highly stable over time, we sought to determine whether this stability was elastic or fragile. Within our data set we included a subcohort of 50 individuals who were sampled before and after travel to a developing nation, where there is an elevated risk of gastrointestinal infection. Of this subcohort, 24 individuals developed acute gastroenteritis while abroad (**Supplementary Fig. 2a**). Of these cases, 22 were classified as having moderate or classic gastroenteritis, with a median duration of 2 d of diarrhea. Individuals were asked not to use antibiotics during this gastrointestinal challenge unless clinically indicated (87.5% did not use antibiotics), to allow a natural immune response to take course. Given the activation of the immune system during infection, as well as the importance of the gut microbiome[10], this experimental design allowed the determination of whether a combined immunological and microbiological disturbance would act as a 'reset' on the immunological landscape, with individuals stabilizing at an alternative equilibrium point after the resolution of infection. To test this hypothesis, we repeated our analysis of intraindividual variation by segregating the population into individuals that were continuously healthy during the sampling period and individuals that experienced acute gastrointestinal challenge between samples. We observed no substantive effect for gastrointestinal interlude on any immunological parameter, with modest changes in R^2 (**Fig. 2a**) driven by reduced numbers in subsets, as indicated by the accompanying reductions in P values (**Fig. 2b**). Notably, IL-17-producing helper T (T_H17) cells were unaffected in this analysis, despite compelling evidence for microbiome interaction in this cell type[11] (**Fig. 2a**,**b** and **Supplementary Fig. 3**). Having failed to identify single immune parameters altered by immunological perturbation, we performed MDS on paired samples to test whether severe gastrointestinal infection had an effect via the cumulative effect of minor changes on multiple parameters (**Fig. 2d**). In this analysis, multiple dimensions (each representing one of 54 immune parameters) were reduced to two dimensions. Samples (from different visits or different individuals) that appear closer together are more similar; those farther apart are more dissimilar. The data from the individuals affected by diarrhea did not separate from the rest of the data, and the 'immunological distance' between longitudinal samples from this group was no greater than that of individuals who were continuously healthy (**Fig. 2e**). Even among the subset of patients with the longest duration of gastroenteritis (≥4 d) we did not observe an increase in immunological distance (**Supplementary Fig. 2**). These analyses demonstrate not only that the human immune system exists in a diverse set of stable equilibria but also that these equilibria are maintained after immunological and microbiological disturbances, with each individual returning to steady state after the resolution of infection. This result does not exclude functional or numerical changes within the response clones (expansion, conversion to a memory phenotype) but instead refers to the overall cell subset structure of the immune system. One possible explanation for this is that antigen-specific memory cells remaining after immunological challenge are few in number[12–14], and the intrinsic biases in the individual that make up the prior

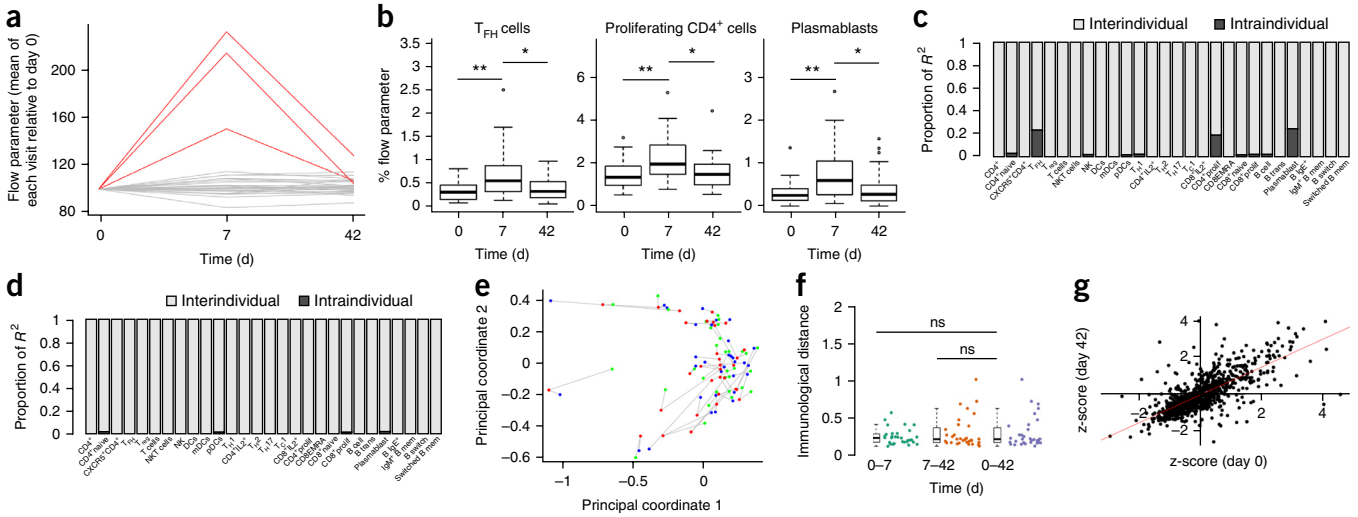

Figure 3 Immunological equilibrium shows elasticity after influenza vaccination. (**a**) Immunological parameters in prevaccination and post-vaccination (days 7 and 42) samples from 32 English individuals after phenotyping, normalization to day 0 values and assessment for change using paired *t*-tests. Unchanged variables are shown in gray; significantly modified variables ($P < 0.0018$) are shown in red. (**b**) Box plots for each significant immune parameter from **a**. $*P < 0.001$; $**P < 0.0001$, uncorrected, two-tailed paired *t*-test. Boxes and center lines, interquartile range (IQR) and median, respectively; whiskers, 1.5× IQR. (**c**,**d**) Proportion of R^2 values from all volunteers attributable to either interindividual or intraindividual differences assessed for day 0 and 7 (**c**) or day 0 and 42 (**d**). (**e**) MDS of vaccination time points. Points represent individuals' samples from day 0 (green), 7 (red) and 42 (blue); gray lines indicate immunological distance ($n = 32$ individuals and 96 visits). (**f**) Quantification of the immunological distances between days 0 and 7; days 7 and 42; and days 0 and 42 for each volunteer ($n = 32$), with paired *t*-test. ns, not significant. Boxes and center lines, interquartile range (IQR) and median, respectively. Whiskers, farthest data point or 1.5× IQR. (**g**) *z*-scores for each volunteer and parameter at days 0 and 42. Correlation analysis indicates the line of best fit.

immune status also apply to newly expanded clones. Thus, although individual responding T cell and B cell clones may change markedly as a response to activation, the functional landscape into which they assimilate remains intact.

In the real-world context, study of gastrointestinal infection has several limitations, including the inability to take a peak-infection sample and variation in the sampling schedule and infection. To overcome these issues, we initiated an independent cohort to assess how a defined immunological stimulus—influenza vaccination—affects the immune landscape. Thirty-two healthy English individuals, between 53 and 64 years of age, were recruited during the 2014–2015 winter influenza vaccination season. Volunteers were sampled before intramuscular vaccination with the standard seasonal inactivated influenza vaccine and at 7 and 42 d after vaccination. Samples were then analyzed on a parallel immune phenotyping platform, which replicated the variation structure and substructure present in the Belgian cohort (**Supplementary Fig. 4**). Analysis of individual immune parameters indicated that most parameters were unchanged throughout the study (**Fig. 3a** and **Supplementary Fig. 5**). The exceptions were circulating follicular T helper–like ($cT_{FH\text{-}like}$) cells, proliferating $CD4^+$ ($Ki67^+$) T cells and plasmablasts, all of which showed marked increases at day 7 and had returned to baseline at day 42 (**Fig. 3b**). To determine whether the immunological challenge of vaccination disrupted the immune landscape of the volunteers, we built a longitudinal ANOVA model. Between day 0 and day 7, almost all variation could be accounted for by interindividual variation, with the exception of the three vaccination-response parameters, where sample time point (intraindividual variation) was significant (**Fig. 3c**). With the resolution of the vaccine response (assessing days 0 and 42), even the vaccination response parameters showed no substantial time point (intraindividual) variation (**Fig. 3d**), demonstrating that even perturbed parameters rebounded to prechallenge settings. This is consistent with systems-vaccinology studies that describe an alteration in the gene-expression profile of peripheral blood samples in the first 2 weeks after vaccination, followed by a return to baseline state[12,15,16]. To demonstrate the global robustness of the immunological landscape, we used MDS. This analysis indicated that samples were clustered by individual rather than time point (**Fig. 3e**), with low immunological distances between samples (**Fig. 3f**). To test whether the return to baseline was a population-level process or whether individuals returned to a unique baseline, we calculated *z*-scores for each immunological parameter and individual at day 0 and day 7. We observed a strong correlation ($R^2 = 0.74$, $P < 2 \times 10^{-16}$) between time points (**Fig. 3g**), which indicates that individuals retained their relative interindividual differences after vaccination. These results, based on an independent immunological challenge, show substantial stability in human immune cell subsets and indicate an elastic ability to respond to antigen then return to a baseline state.

Age and cohabitation affect the immunological landscape

Having established the diversity of elastic stable equilibria in the human immune system, we sought to determine the underlying biological drivers. As variability was greatest in the precursor cell populations, we first investigated the effect of age on immune profile. Our data set included substantial numbers of both children (<18 years; $n = 40$) and older people (>65 years; $n = 54$). We observed strong relationships, both positive and negative, between immune parameters and age (**Fig. 4a–c**). With a threshold of adjusted $P < 0.01$ and $r < -0.35$ or $r > 0.35$ (**Supplementary Fig. 6** and **Supplementary Table 2**), three immune parameters had a negative relationship with age, and seven had a positive relationship with age. Proportions of $CD4^+$ RTE cells (**Fig. 4d**), transitional B cells (**Fig. 4e**) and $CD8^+$ RTE cells (**Fig. 4f**) decreased in a linear fashion as age increased, consistent with an age-dependent reduction in thymus and bone marrow activity. We observed positive correlations between several inflammatory populations and age—namely T_H1 cells, $CD4^+IL\text{-}2^+$ cells, T_C1 cells, $CD8^+IL\text{-}2^+$

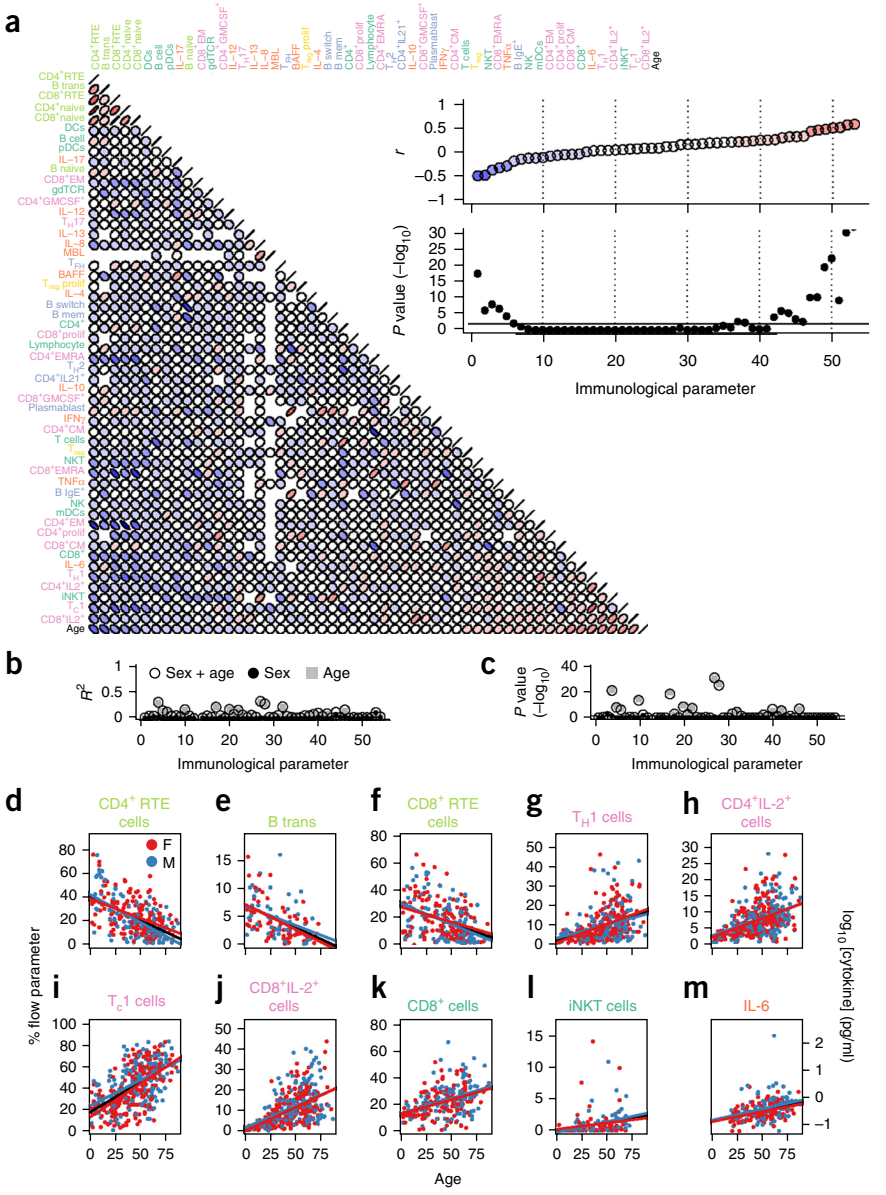

Figure 4 Age is a major determinant of immunological equilibrium. (**a**) Pairwise Spearman *r* values between age and immune parameters. Coefficients are described by the angle of ellipse (left-leaning, negative; right-leaning, positive) and color (blue, negative; red, positive). Inset, Spearman's *r* (top) and $-\log_{10}$ Bonferroni-corrected *P* values (bottom) plotted against age. (**b**,**c**) R^2 values (**b**) and $-\log_{10}$ (Bonferroni-corrected *P* values) (**c**) for each immune parameter for models incorporating sex (filled circles), age (gray squares) or both (open circles) as the independent variable(s). (**d**–**m**) Individual scatter plots for each immune parameter significantly associated with age: $CD4^+$ RTE cells ($P = 3 \times 10^{-18}$) (**d**), transitional B cells ($P = 8 \times 10^{-7}$) (**e**), $CD8^+$ RTE cells ($P = 1 \times 10^{-8}$) (**f**), T_H1 cells ($P = 3 \times 10^{-20}$) (**g**), $CD4^+IL\text{-}2^+$ T cells ($P = 5 \times 10^{-23}$) (**h**), T_C1 cells ($P = 5 \times 10^{-31}$) (**i**), $CD8^+IL\text{-}2^+$ T cells ($P = 8 \times 10^{-33}$) (**j**), $CD8^+$ T cells ($P = 8 \times 10^{-11}$) (**k**), *i*NKT cells ($P = 6 \times 10^{-10}$) (**l**) and serum IL-6 ($P = 7 \times 10^{-11}$) (**m**), plotted as percentage of the flow parameter (**d**–**l**) or $\log_{10}$ of cytokine concentration (**m**) against age, with linear regression lines shown for women (red), men (blue) and all subjects (black). F, female; M, male.

cells and invariant natural killer T (*i*NKT) cells—and an age-associated increase in $CD8^+$ T cells (**Fig. 4g–l**). Despite the presence of significantly more ($P = 3 \times 10^{-20}$) T_H1-associated cells with age, serum interferon-γ (IFN-γ) did not reach our threshold for correlation coefficient significance ($r = 0.18$ (95% CI 0.06–0.29); adjusted $P = 5.2 \times 10^{-3}$). We did, however, observe that the proinflammatory cytokine IL-6 significantly increased with age (**Fig. 4m**). Together, these data demonstrate that age is a major contributor to the immune profiles of healthy individuals, with downregulation of precursor populations and an upregulation of T_H1-associated inflammatory populations with age. Notably, the data from two extremes of age (<18 years and >65 years) did not show larger variances than the data from the central ages. This observation suggests that—at least for these immune parameters—pediatric and geriatric immunology are not 'special cases' with different rules.

There are well-defined differences between sexes in aging-related diseases. We therefore assessed whether any of the immune parameters were associated with sex when controlling for age. We found that sex added no explanatory power to a model already including age, with essentially no effect on the variance of each immune parameter (median difference in $R^2 = 0.006$) (**Fig. 4b**,**c**). The only significant effect of sex on the immune profile was on $CD4^+$ T cell numbers, which were higher in women than in men (median $CD4^+$ T cell frequency = 17.8% in women and 14.6% in men; adjusted $P = 1.3 \times 10^{-3}$), consistent with previous reports[17].

To extend the analysis of physiological influences on immune profiles, we assessed body mass index (BMI), anxiety and depression in a subset of adult (age >18 years) volunteers. BMI provided significant associations with two immune parameters—proportion of $CD4^+IL\text{-}2^+$ cells (adjusted $P = 0.03$) and serum IL-6 concentration ($P = 0.007$)—but their effect sizes were modest, with an R^2 of 0.06 and 0.09, respectively (**Fig. 5a**,**b**). Adding complexity is the tendency of BMI to increase with age (**Fig. 5c**). To control for this, we built age and BMI into our model. For all the age-BMI immune associations with the exception of IL-6, BMI was the minor contributor to the variation, with age having a much greater role (**Fig. 5d**). By contrast, serum concentrations of IL-6, an adipose-associated cytokine, were influenced equally by age and BMI (**Fig. 5d**). Many other changes associated with BMI in the literature[18–21] were not observed in our data set. This discrepancy may be due to inadequate controlling for age in prior studies (the relative strength of associations observed here suggests that using age as a category rather than as a linear variable would lead to substantial over-estimation of the effect of BMI) or a relative lack of individuals at the extremes of the BMI scale in the current study. To investigate a potential neuroimmunological connection, we sought to determine whether anxiety and depression scores[19] altered immune parameters, but observed no substantial or significant effects (**Fig. 5e**). Together, these data indicate that BMI, depression and anxiety do not substantively alter immune equilibrium

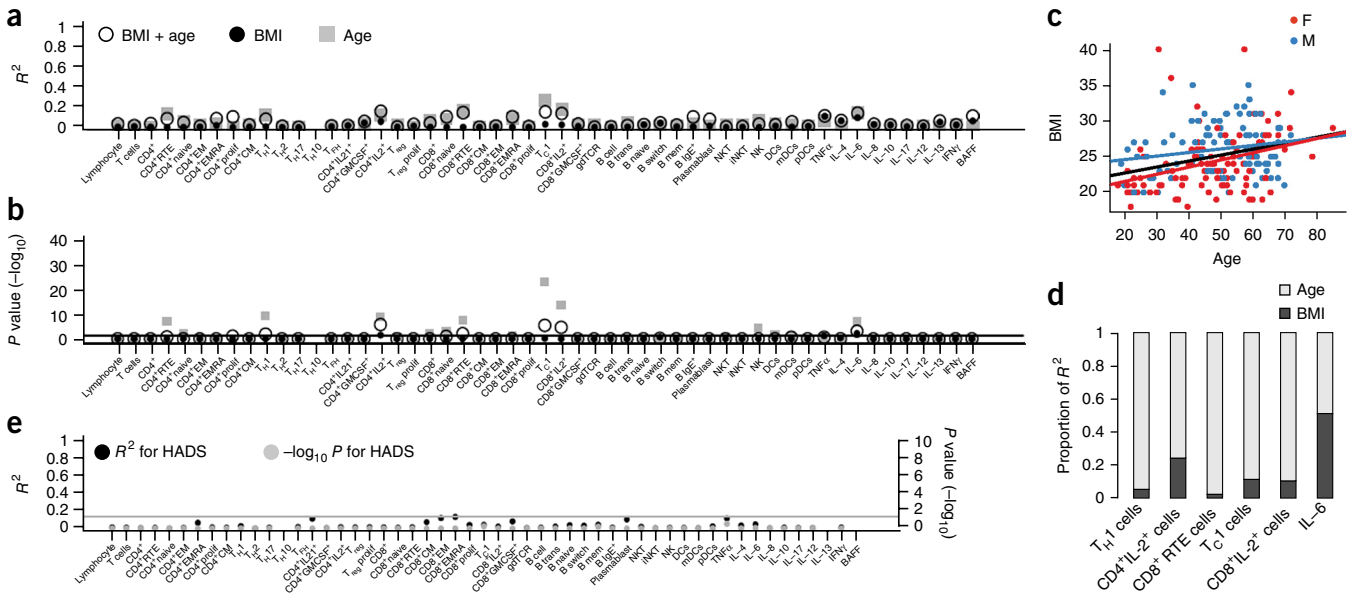

Figure 5 Immunoprofile is not influenced by BMI, depression or anxiety. (**a**,**b**) R^2 (**a**) and $-\log_{10}$ of Bonferroni-corrected P values (**b**) for each immune parameter for models incorporating BMI (filled circles; $n = 213$ individuals), age (filled squares; $n = 367$ individuals) or both BMI and age (open circles; $n = 213$ individuals) as the predictor variable(s). Analysis excluded children (<18 years). (**c**) Relationship between BMI and age. F, female; M, male. (**d**) Relative R^2 contributions of age and BMI to immune parameters that were significant in a model including both BMI and age (adjusted $P < 0.05$). (**e**) R^2 and $-\log_{10}$ of adjusted P values for each immune parameter for a model incorporating Hospital Anxiety and Depression Scale (HADS) anxiety depression scores ($n = 235$ individuals).

in healthy individuals beyond the association of BMI with serum IL-6 concentrations.

Finally, we sought to determine whether cohabitation has an impact on immune equilibrium. We sampled 70 parental pairs (140 adults aged 18–65 with one or more children living at home) to determine whether a shared environment altered immunoprofiles. Our hypothesis was that the immunoprofiles of a parental pair would be more closely related than a random *in silico* male-female pairing. To test this assertion, we used MDS to reduce the diversity of the immune system to two dimensions (**Fig. 6a**) and linked each pair. We measured the distance between mother and father for all parental pairs (**Fig. 6b**) and compared this to the distance between randomly generated pairings (**Fig. 6b**). We found that there was a significant and substantial (~50%) reduction in immune variability between parental pairs compared to the randomized pairs, suggesting that a shared environment drives a convergence between immunoprofiles. This effect was independent of age, as exclusion of the age-related parameters identified did not change the result (**Supplementary Fig. 7**). Likewise, measurement of cytokine production after *in vitro* stimulation of peripheral blood mononuclear cells (PBMCs) showed significant convergence between the profiles of coparents (**Supplementary Fig. 8**), demonstrating that the effect extends to alternative methodologies. These results demonstrate that although the immunological equilibrium point is robust and stable in an individual, two individuals in a cohabitation relationship converge toward a single immunological equilibrium point.

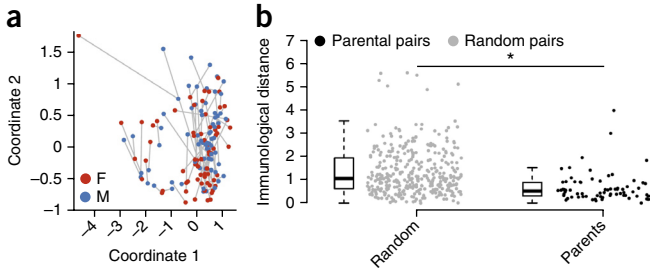

Figure 6 Parenthood shapes the immune system toward a shared equilibrium. (**a**) MDS compression of immune profile correlations in 70 pairs of biological parents with a child still living at home. Gray lines indicate immunological distance between partners. F, female; M, male. (**b**) Immunological distance, as measured by MDS, between parental pairs and random male-female pairs. $^*P = 8 \times 10^{-11}$ (two-tailed Mann-Whitney test). Boxes and center lines, interquartile range (IQR) and median, respectively; whiskers, farthest data point or 1.5× IQR.

DISCUSSION

Nongenetic factors are estimated to account for ~50–75% of immunological variation between healthy individuals[2–4], yet a thorough understanding of the causative factors at play has been lacking. Through the use of a systems-immunology approach and targeted subcohorts of healthy individuals, we were able to assess the main nongenetic factors of immunological variation. Among age, BMI, sex and psychological factors, an individual's age was the most important influence on their immunological landscape. This result complements longstanding observations that immune function (response to vaccination, infection, cancer immunosurveillance) deteriorates with age. The reduction in T cell precursors may be explained by thymic involution[22]; however, the highly concordant decrease in transitional B cells suggests a common root cause, such as impaired bone marrow function[23]. The age-dependent increase in T_H1-associated populations is notable, as T_H2 and T_H17 cell populations did not show an association with age, which indicates that this effect is specific to the T_H1 arm of the immune system. Overall, the combined effect of age and genetics seem to account for most of the variability in the ten age-associated traits identified. Here we observed a median age-dependent R^2 for these traits of 0.13, and the median published heritability (R^2 for genetic effects)[3] is 0.67. We estimate that ~80% of the variation in age-related immune parameters is explicable by age and genetic factors, with some of the remaining variability potentially

explained by interactions between age and genetics. It is intriguing that age-related immune parameters have higher heritability estimates than the rest of the immunoprofile[3].

The immunological effect of BMI is perhaps best regarded as a minor acceleration in the normal immuno-aging process. The notable exception to the negligible effect of BMI was its positive association with serum IL-6 concentrations. One possibility is that this IL-6 is 'nonimmune' in origin, as monocytes from aged individuals secrete less IL-6 in response to Toll-like receptor (TLR) ligation[5]. Although the precise cellular origin of increased IL-6 concentrations with age and high BMI is unknown, a plausible candidate is vascular smooth muscle cells (VSMCs). VSMCs from aged mice and nonhuman primates produce more IL-6 than those from younger controls[24,25], and obesity increases the inflammatory phenotype of VSMCs[26]. Given the complex biological function of IL-6 (ref. 27), this effect could partially account for the alteration of clinical outcomes that obesity has on diseases such heart failure.

We found very little effect of sex on the immune landscape. This finding is at odds with the longstanding observation that autoimmune diseases are generally more common in women (at premenopausal ages) than in men and that vaccine responses are more robust in women[28], although the sex difference in vaccine response varies greatly[28–30]. Notably, sex-based differences are more limited at the cellular level than at the molecular level[31]. The incomplete correlation between gene signature and cell type suggests that the discrepancy can be resolved by a model where high diversity in molecular expression is largely compensated for at the cellular level.

One of the most surprising results from our study was the degree to which immune profiles were more similar between parents than unrelated pairs. This suggests that a shared environment acts in some way to bring immunoprofiles toward a convergent equilibrium. There are many plausible biological mechanisms for this. For example, cohabiting individuals show convergence in microbiomes[32], an effect that extends to pet dogs[33]. Cohabiting individuals in a relationship have more similar microbiomes[34], possibly via direct transmission[35], which would make the microbial convergence even stronger for parents. In our study this shared environment includes a shared vector (the child), the significance of which requires further investigation. Beyond the bacterial components of the microbiome, close proximity allows the transmission of viral pathogens, including CMV, which was found to influence twin concordance in more than half of their immune traits[3]. A shared environment (including, presumably, socioeconomic status) will also bring shared behaviors, a process called spousal concordance; a person's diet (which can also influence gut microbiome)[36], smoking[37], alcohol intake[38], exercise levels and even control of chronic diseases such as hypertension[39] are all likely to be influenced by a partner's attitudes toward them. Perhaps partner choice also influences response to immunity and immune pathologies. We also note, but decline to comment further, that this 'parenthood effect' is a much stronger influence on the immune system than acute and untreated gastroenteritis.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all volunteers for their participation; the Cambridge BioResource staff for help with volunteer recruitment; members of the Cambridge BioResource SAB and Management Committee for support of this study; and A. Nuygen, D. Franckaert, D. Danso-Abeam and L. Van Eyck (KUL-VIB) for technical assistance. Supported by the National Institute for Health Research Cambridge Biomedical Research Centre; the European Research Council (Start Grant IMMUNO to A.L. and Start Grant TWILIGHT to M.A.L.). the NIHR (E.J.C.), the Biotechnology and Biological Sciences Research Council (M.A.L.), Research Fund KU Leuven (OT/11/087 to A.G.), Research Foundation Flanders (G073415N to A.G.) and the Wellcome Trust (105920/Z/14/Z to J.C.L.).

AUTHOR CONTRIBUTIONS

E.J.C. analyzed the data and drafted the manuscript. J.D. and J.E.G.-P. performed the experiments. V.L. analysed the vaccination cohort. J.C.L., C.W., I.M., A.N. and G.B. designed and recruited subcohorts. M.A.L. designed and supervised the vaccination study, contributed to the data analysis and the manuscript. A.L. designed and supervised the study and drafted the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Davis, M.M. Immunology taught by humans. *Sci. Transl. Med.* **4**, 117fs2 (2012).
2. Orrù, V. *et al.* Genetic variants regulating immune cell levels in health and disease. *Cell* **155**, 242–256 (2013).
3. Brodin, P. *et al.* Variation in the human immune system is largely driven by non-heritable influences. *Cell* **160**, 37–47 (2015).
4. De Jager, P.L. *et al.* ImmVar project: Insights and design considerations for future studies of "healthy" immune variation. *Semin. Immunol.* **27**, 51–57 (2015).
5. Shaw, A.C., Goldstein, D.R. & Montgomery, R.R. Age-dependent dysregulation of innate immunity. *Nat. Rev. Immunol.* **13**, 875–887 (2013).
6. Jamieson, B.D. *et al.* Generation of functional thymocytes in the human adult. *Immunity* **10**, 569–575 (1999).
7. den Braber, I. *et al.* Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity* **36**, 288–297 (2012).
8. Johnson, P.L., Yates, A.J., Goronzy, J.J. & Antia, R. Peripheral selection rather than thymic involution explains sudden contraction in naive CD4 T-cell diversity with age. *Proc. Natl. Acad. Sci. USA* **109**, 21432–21437 (2012).
9. Tsang, J.S. *et al.* Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* **157**, 499–513 (2014).
10. Caballero, S. & Pamer, E.G. Microbiota-mediated inflammation and antimicrobial defense in the intestine. *Annu. Rev. Immunol.* **33**, 227–256 (2015).
11. Ivanov, I.I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).
12. Tsang, J.S. *et al.* Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* **157**, 499–513 (2014).
13. Lucas, M. *et al. Ex vivo* phenotype and frequency of influenza virus-specific CD4 memory T cells. *J. Virol.* **78**, 7284–7287 (2004).
14. Miller, J.D. *et al.* Human effector and memory $CD8^+$ T cell responses to smallpox and yellow fever vaccines. *Immunity* **28**, 710–722 (2008).
15. Bucasas, K.L. *et al.* Early patterns of gene expression correlate with the humoral immune response to influenza vaccination in humans. *J. Infect. Dis.* **203**, 921–929 (2011).
16. Nakaya, H.I. *et al.* Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.* **12**, 786–795 (2011).
17. Amadori, A. *et al.* Genetic control of the CD4/CD8 T-cell ratio in humans. *Nat. Med.* **1**, 1279–1283 (1995).
18. Spiegelman, B.M. & Hotamisligil, G.S. Through thick and thin: wasting, obesity, and TNF-α. *Cell* **73**, 625–627 (1993).
19. Spielmann, G., Johnston, C.A., O'Connor, D.P., Foreyt, J.P. & Simpson, R.J. Excess body mass is associated with T cell differentiation indicative of immune ageing in children. *Clin. Exp. Immunol.* **176**, 246–254 (2014).
20. Martín-Romero, C., Santos-Alvarez, J., Goberna, R. & Sanchez-Margalet, V. Human leptin enhances activation and proliferation of human circulating T lymphocytes. *Cell. Immunol.* **199**, 15–24 (2000).
21. Damluji, A.A. *et al.* Association between anti-human heat shock protein-60 and interleukin-2 with coronary artery calcium score. *Heart* **101**, 436–441 (2015).
22. Dooley, J. & Liston, A. Molecular control over thymic involution: from cytokines and microRNA to aging and adipose tissue. *Eur. J. Immunol.* **42**, 1073–1079 (2012).
23. Franckaert, D. *et al.* Premature thymic involution is independent of structural plasticity of the thymic stroma. *Eur. J. Immunol.* **45**, 1535–1547 (2015).
24. Song, Y. *et al.* Aging enhances the basal production of IL-6 and CCL2 in vascular smooth muscle cells. *Arterioscler. Thromb. Vasc. Biol.* **32**, 103–109 (2012).
25. Csiszar, A. *et al.* Age-associated proinflammatory secretory phenotype in vascular smooth muscle cells from the non-human primate *Macaca mulatta*: reversal by resveratrol treatment. *J. Gerontol. A Biol. Sci. Med. Sci.* **67**, 811–820 (2012).
26. Van de Voorde, J., Boydens, C., Pauwels, B. & Decaluwe, K. Perivascular adipose tissue, inflammation and vascular dysfunction in obesity. *Curr. Vasc. Pharmacol.* **12**, 403–411 (2014).

27. Fontes, J.A., Rose, N.R. & Cihakova, D. The varying faces of IL-6: from cardiac protection to cardiac failure. *Cytokine* **74**, 62–68 (2015).
28. Giefing-Kröll, C., Berger, P., Lepperdinger, G. & Grubeck-Loebenstein, B. How sex and age affect immune responses, susceptibility to infections, and response to vaccination. *Aging Cell* **14**, 309–321 (2015).
29. Furman, D. *et al.* Systems analysis of sex differences reveals an immunosuppressive role for testosterone in the response to influenza vaccination. *Proc. Natl. Acad. Sci. USA* **111**, 869–874 (2014).
30. Tsang, J.S. *et al.* Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* **157**, 499–513 (2014).
31. Whitney, A.R. *et al.* Individuality and variation in gene expression patterns in human blood. *Proc. Natl. Acad. Sci. USA* **100**, 1896–1901 (2003).
32. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
33. Song, S.J. *et al.* Cohabiting family members share microbiota with one another and with their dogs. *eLife* **2**, e00458 (2013).
34. Lax, S. *et al.* Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* **345**, 1048–1052 (2014).
35. Kort, R. *et al.* Shaping the oral microbiota through intimate kissing. *Microbiome* **2**, 41 (2014).
36. Carmody, R.N. *et al.* Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* **17**, 72–84 (2015).
37. Falba, T.A. & Sindelar, J.L. Spousal concordance in health behavior change. *Health Serv. Res.* **43**, 96–116 (2008).
38. Graham, K. & Braun, K. Concordance of use of alcohol and other substances among older adult couples. *Addict. Behav.* **24**, 839–856 (1999).
39. McAdams DeMarco, M. *et al.* Hypertension status, treatment, and control among spousal pairs in a middle-aged adult cohort. *Am. J. Epidemiol.* **174**, 790–796 (2011).

ONLINE METHODS

Participant selection and ethical approval. All participants were self-reported to be Caucasian, and all sampling was conducted in Belgium, except for the vaccination cohort, who were Caucasian and sampled in England. All individuals or their legal guardians gave written informed consent, and the study was approved by the Ethics Committee of the University Hospitals Leuven. The influenza vaccination study protocol was approved by the Health Research Authority, National Research Ethics Service committee South Central, Hampshire A, UK (REC reference 14/SC/1077). Demographic and clinical data were collected through a questionnaire. Exclusion criteria were cancer, autoimmunity and gastrointestinal complaints. Individuals that were parents of the same child were considered to be cohabiting. Individuals planning a trip to developing countries in South America, Africa or Asia were considered to be at elevated risk of gastrointestinal infection. Acute gastroenteritis was defined as self-reported diarrhea (passage of unformed stools), with or without nausea, vomiting, abdominal cramps, pain, fever or blood in stools (using ROME-III criteria). Gastroenteritis was classified as mild if no additional symptoms were reported, moderate if at least one additional symptom was present, and classic if at least three diarrheal episodes occurred per day and at least one additional symptom was present. Anxiety and depression were assessed using the Hospital Anxiety and Depression Scale (HADS) and the Patient Health Questionnaire (PHQ-15). For recruitment to the vaccination study, potential participants were excluded if they had already received the 2014–2015 seasonal influenza vaccination; had a previous adverse reaction to any vaccination or a known allergy to any components of the vaccine; were taking immune-modulating medication; or were pregnant or breastfeeding. In total, 32 healthy donors between ages 53 and 64 years were recruited from Cambridge BioResource as part of the vaccination study during the 2014–2015 winter. Participants were administered the inactivated influenza vaccine (split virion) BP vaccine (Sanofi Pasteur) by intramuscular injection in the right deltoid.

Blood sampling and PBMC isolation. Blood samples from Belgian participants were collected in heparin tubes and rested at 22 °C for 4 h before separation of serum and PBMCs using lymphocyte separation medium (LSM, MP Biomedicals). PBMCs were frozen in 10% DMSO (Sigma) and stored at −80 °C for a maximum of 10 weeks. For the vaccination cohort, research nurses at the Cambridge BioResource collected blood samples into EDTA-coated tubes on the day of vaccination (before administration of the vaccine) and 7 d and 42 d after vaccination. PBMCs were isolated using 15 mL Histopaque-1077 (Sigma), frozen in FBS supplemented with 10% DMSO (Sigma) overnight at −80 °C then stored in liquid nitrogen before analysis by flow cytometry.

Flow cytometry phenotyping. Thawed cells were stained with antibodies as listed in **Supplementary Table 3**. Ki67 and Foxp3 staining was performed after treatment with fixation-permeabilization buffer (eBioscience). Cytokine staining was performed after *ex vivo* stimulation for 5 h in 50 ng/ml PMA (Sigma) and 500 ng/ml ionomycin (Sigma) in the presence of GolgiStop (BD Biosciences). Stimulated cells were surface stained, fixed and permeabilized with Cytofix/Cytoperm (BD), before staining for cytokines. Additional cells were stimulated for 72 h for supernatant assessment by MSD (see below). Data were acquired on a BD FACSCantoII and analyzed with FlowJo (Tree Star). The vaccination cohort data were acquired on a BD LSRFORTESSA and analyzed using FlowJo (Tree Star).

Serological assessment. Plasma samples collected were stored at −80 °C. Circulating levels of mannan-binding lectin (MBL) were quantified in plasma using the MBL Oligomer ELISA Kit form Bioporto (Copenhagen, Denmark). Circulating levels of B cell activation factor (BAFF) were measured using a human BAFF Quantikine ELISA (R&D Systems). Cytokine plasma concentrations were quantified by electrochemiluminescence immunoassay using the V-Plex human Proinflammatory panel MSD (Meso Scale Discovery) plates. All reagents and standards were provided by each manufacturer. Samples and standards were prepared according to each manufacturer's instructions.

Data handling. Data (phenotypic, flow cytometric and serological) were collated and stored in Microsoft Excel. All data analysis was performed using R version 3.1.0 (ref. 40) via RStudio IDE (https://www.rstudio.com/) version 0.98.1102. Figures were drawn using knitr (https://www.rforge.net/doc/packages/knitr/knitr-package.html), which produces PDF output via LaTeX.

ELISA and MSD data were preprocessed as follows: any experimental value below the lower limit of detection for the assay was replaced with the lower limit of detection of that cytokine. After this step, all ELISA and MSD data were $\log_{10}$-transformed. The flow cytometry data were used as percentages as exported from FlowJo. No data were excluded from analysis. Interquartile range, median, number of missing values for each immune parameter (flow cytometry–derived and serologically derived parameters together) are shown in **Supplementary Table 2**. For nonlongitudinal analyses only the most recent sample from each individual was used. The original data (phenotypic, flow cytometric and serological) are available to download as XLS or RData files (**Supplementary Data Sets 1**–**4**). To generate the random distribution in study of parental pairs (**Fig. 6**) each male in the parental data set was computationally paired in a random fashion with five females from the parental data set.

Statistical analysis. All sample collection, data acquisition and data processing was performed in a blinded fashion before statistical analysis. Spearman's rank correlation coefficient was used throughout for pairwise correlation comparisons. Euclidean distances were calculated from correlation matrices (using Spearman) as preprocessing for MDS (using either cmdscale in base R, or monoMDS from the vegan package (https://cran.r-project.org/web/packages/vegan/index.html) or hierarchical clustering (using heatmap in stats package)[40]. Consensus clustering was performed using the R package ConsensusClusterPlus, an R implementation of the original algorithm[41]. Correlation plots were drawn using the plotcorr function in the ellipse package (https://cran.r-project.org/web/packages/ellipse/ellipse.pdf). Linear regression modeling was performed using base R function (lm), with the proportions of R^2 calculated with the relaimpo package[42]. Lm will provide differing statistical models depending upon the input data. For most of the modeling shown, a continuous variable (immune parameter) is modeled using a categorical variable (such as subject identifier or visit number). In this situation, lm provides an ANOVA model (for example, in **Fig. 2a**–**c**: immune parameter ~ subject identifier + visit number). If both variables are continuous, linear regression modeling is used. Two-group comparisons were made using two-tailed Mann-Whitney tests, except for the paired vaccination data, where a paired *t*-test was used to compare samples from the same individual at different time points. Bonferroni correction was used for multiple testing, as implemented in base R.

Code availability. Original data sets are provided as both Excel (**Supplementary Data Sets 1** and **2**) and RData files (**Supplementary Data Sets 3** and **4**). The code used to produce the figures is provided as both PDF (**Supplementary Code 1** and **Supplementary Code 2**) and R markdown files (**Supplementary Code 3** and **Supplementary Code 4**).

40. R Development Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2014).
41. Monti, S., Tamayo, P., Mesirov, J. & Golub, T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* **52**, 91–118 (2003).
42. Grömping, U. Relative importance for linear regression in R: the package relaimpo. *J. Stat. Softw.* **17**, 1–27 (2006).