RESOURCE

nature immunology

Social network architecture of human immune cells unveiled by quantitative proteomics

Jan C Rieckmann[1], Roger Geiger[2,3], Daniel Hornburg[1], Tobias Wolf[2,3], Ksenya Kveler[4], David Jarrossay[2], Federica Sallusto[2], Shai S Shen-Orr[4], Antonio Lanzavecchia[2,3], Matthias Mann[5] & Felix Meissner[1]

The immune system is unique in its dynamic interplay between numerous cell types. However, a system-wide view of how immune cells communicate to protect against disease has not yet been established. We applied high-resolution mass-spectrometry-based proteomics to characterize 28 primary human hematopoietic cell populations in steady and activated states at a depth of >10,000 proteins in total. Protein copy numbers revealed a specialization of immune cells for ligand and receptor expression, thereby connecting distinct immune functions. By integrating total and secreted proteomes, we discovered fundamental intercellular communication structures and previously unknown connections between cell types. Our publicly accessible (http://www.immprot.org/) proteomic resource provides a framework for the orchestration of cellular interplay and a reference for altered communication associated with pathology.

Distinct immune functions are executed by highly specialized cell types. The coordinated action of the immune system resembles a social network, which enables complex immunological tasks beyond the sum of the functions of individual isolated cell types. A major goal of immunology research is to understand how the context-dependent crosstalk of different cell types and the orchestration of their functions enable protection against disease[1–4]. However, the architecture and syntax by which biological messages, such as cytokines with pleiotropic functions, are exchanged between sometimes distant and mobile cell types is a central feature of the immune system that remains incompletely understood[5,6]. Moreover, the functions of transmitted messages vary depending on the cellular sender, receiver and pathophysiological state, making the interpretation of immune responses inherently challenging.

At the molecular level, intercellular signals are mostly communicated through proteins produced by sending cells that act on receptors of receiving cells. Current approaches focus on the characterization of interactions between individual cell types or cytokines, but new proteomics technologies promise to capture the complexity of intercellular communication comprehensively at the protein level. However, this promise has been challenging to fulfil because of the scarcity of certain subtypes of immune cells, the large dynamic range of the cellular proteome (>6 orders of magnitude) and the low concentration of secreted factors. Important insights into immune cell type composition and intracellular signaling networks have been gained by system-wide transcriptional approaches and antibody-based technologies[7–11]. Although relative measurements of transcripts can be very comprehensive, the correlation between mRNA and protein copy numbers can vary widely[12,13], especially for proteins with roles in intercellular crosstalk[14]. Quantitative, high-resolution mass spectrometry (MS)-based proteomics has developed rapidly and has now matured into a powerful technology that provides a unique opportunity for the system-wide characterization of cellular senders and receivers as well as the accurate quantification of transmitted messages[15–17].

To unravel the complex interactions between immune cells at the protein level, we combined advanced MS approaches developed in our laboratories and characterized 28 hematopoietic cell types that we sorted by flow cytometry from human donors. We accommodated the limited amount of rare immune cells by employing a single-run MS analysis approach[18] while keeping the greatest possible proteome depth. To interpret the dynamic cellular proteomes and secretomes, we developed a bioinformatics framework to assign cell-type-resolved immune functions and connect these via transmitted biological messages. This yielded a network topology from which we could deduce the intercellular information exchange in the immune system. By comparing with the current literature, we discovered systematically understudied intercellular connections and biological messages. Our results highlight different communication structures for myeloid and lymphoid immune cells, as well as on the level of individual proteins for messages with distinct paracrine functions. Exemplified by the cell-type- and context-dependent secretion of innate immune cells, our results provide a holistic picture of unique and shared biological messages with their associated inter- and intracellular information flow. Together, our findings define the social network architecture of immune cells and provide a systems biology reference framework of intercellular signaling. The quantitative, high-resolution proteomics immune cell compendium,

[1]Experimental Systems Immunology, Max Planck Institute of Biochemistry, Bayern, Germany. [2]Institute for Research in Biomedicine, Università della Svizzera italiana, Bellinzona, Switzerland. [3]Institute of Microbiology, ETH Zürich, Zürich, Switzerland. [4]Department of Immunology, Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel. [5]Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Bayern, Germany. Correspondence should be addressed to F.M. (meissner@biochem.mpg.de) or M.M. (mmann@biochem.mpg.de).

Received 31 August 2016; accepted 26 January 2017; published online 22 March 2017; doi:10.1038/ni.3693

including protein copy numbers, assignments of cell-type-resolved functions, intercellular communication structures and pairwise cell-type comparisons, is freely available via MaxQB[19] and an interactive online database (http://www.immprot.org/).

RESULTS

Proteome atlas of 28 distinct human hematopoietic cell types

We sorted 28 distinct human hematopoietic cell types from peripheral blood of healthy donors by flow cytometry (**Supplementary Fig. 1** and **Supplementary Table 1**). These comprised cells from seven major lineages, including granulocytes (GN), monocytes (MO), dendritic cells (DC), natural killer (NK), B cells (B), CD4 (T4) and CD8 (T8) lymphocytes, as well as erythrocytes and platelets. The two latter cell types were excluded from further analysis, as their proteomes differed markedly from all other proteomes because of their very specialized functions and lack of a nucleus (**Fig. 1a** and Online Methods). We analyzed the cellular proteomes in their steady state and for a subset (17 cell types) in activated states (**Supplementary Table 2**) in single runs by high-resolution MS using a quadrupole Orbitrap instrument[20] (Online Methods). Each proteome state was measured from three to four donors, generating a total of 175 immune cell proteomes. At a peptide and protein false discovery rate (FDR) of 1%, we identified more than 10,000 different proteins with an average of 7,500 proteins per measurement and 8,700 proteins per quadruplicate. For each major immune cell lineage, we identified an average of 9,500 proteins (**Fig. 1b**, **Supplementary Fig. 2a**,**b** and **Supplementary Table 3**). We performed quantification in a label-free format using the MaxQuant algorithms and the resulting intensity values served as the basis for relative and absolute proteome determination[21,22]. This demonstrated high inter-donor correlation for cell types ($r^2 > 0.97$) (**Supplementary Fig. 2c**).

Gene ontology (GO) category analysis revealed >80% coverage of proteins with known immune-related functions (**Fig. 1c** and **Supplementary Fig. 2d**). In comparison, two previous large-scale human immune cell type cataloguing efforts only covered about 60% of immune annotated and identified approximately 8,000 proteins or genes in total[8,23] (**Fig. 1c**,**d** and **Supplementary Fig. 2d**). Furthermore, the global proteome profiles are consistent with established cellular functional relationships[24] (**Supplementary Fig. 2e**). All marker proteins used for cell sorting were identified, except for the chemokine (C-C motif) receptor 4 (CCR4), and showed the expected expression profiles (**Supplementary Fig. 2f**).

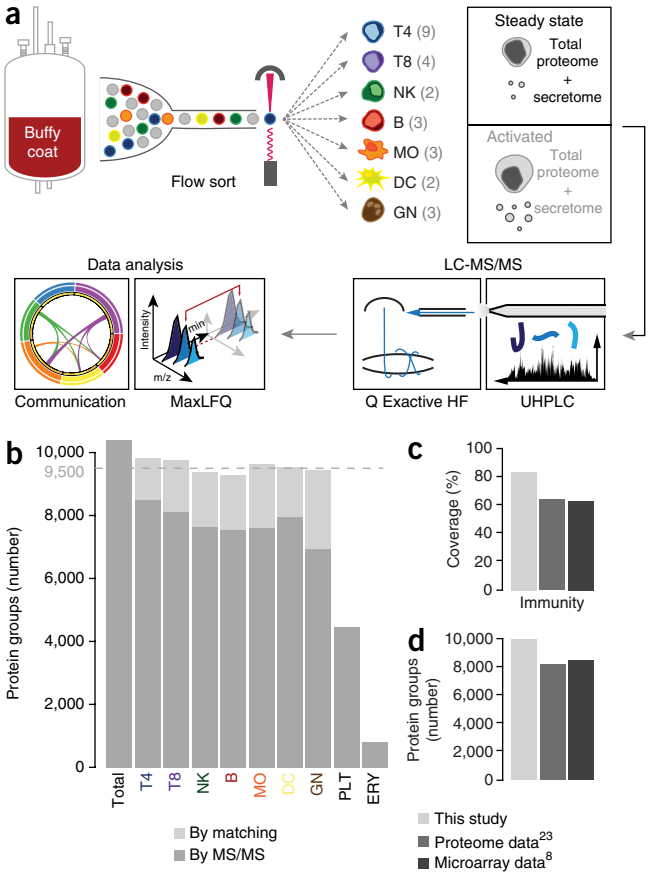

Figure 1 Comprehensive proteome atlas of 28 hematopoietic cell types by single shot LC-MS/MS analysis. (**a**) Schematic of experimental design and proteomics workflow. GNs, MOs, DCs, NK cells, B cells, and T4 and T8 lymphocytes were isolated by flow cytometry and analyzed by LC-MS/MS in steady and activated states. Numbers in brackets indicate the quantity of individual cell types of the indicated cellular lineage (for example, NK: $CD56^{bright}$ and $CD56^{dim}$ NK cells). ERY, erythrocytes; PLT, platelets. (**b**) Number of identified protein groups for each major cell lineage. MS/MS-based identifications and those transferred by 'match-between runs' are indicated in dark and light gray, respectively. (**c**,**d**) Comparison to proteome[23] and microarray data[8]. (**c**) Relative coverage of proteins with known immune function. (**d**) Comparison of the number of protein and gene identifications.

Immune-cell-type-resolved protein expression patterns

Principal component analysis (PCA) to compare the proteomes of the different immune cell types found a clear distinction between lymphoid and myeloid immune cell functions in the first principal component (PC1) (**Fig. 2a**). Bioinformatic enrichment analysis of functional GO terms along this axis[25] revealed that this separation could be attributed to 'T-cell regulation' versus 'early defense mechanisms' (**Supplementary Table 4** and Online Methods). The second dimension (PC2) differentiates 'microbial defense mechanisms' versus 'humoral immunity'.

T lymphocytes and NK cells formed a dense cluster, clearly revealing their close functional relationship. An independent PCA of T and NK cell types (**Fig. 2a**) resolved the effector classes from helper to cytotoxic in PC1 and lineage differentiation from naive to central or effector memory in PC2. Thus, the closest neighbors of NK cells are CD8 T effector memory (T_{EM}) cells, which both target virus-infected cells[26,27].

We selected the top differentially expressed proteins using the seven major lineages described above as labels in a supervised clustering analysis, which confirmed known lineage-specific marker proteins (**Fig. 2b**). These signature proteins were enriched for biological processes according to the functional attributes of the lineages, as expected (**Supplementary Table 5**), with some of these proteins being expressed more than 100-fold higher in particular cell types, such as CD3δ in T cells, CD79A in B cells and Toll-like receptor 2 (TLR2) in monocytes (**Supplementary Fig. 3a–c**). We used the fixed ratio between the total histone signal and cellular DNA to calculate protein copy numbers for the identified proteins[28] (**Supplementary Table 6**). Housekeeping proteins such as GAPDH or ACTB are present at between 10–50 million copies in immune cells, at the top of the ranked expression order (**Fig. 2c**). Lineage markers such as CD14 in classical monocytes or FCGR3B in neutrophils also reach high expression levels of 1–5 million copies. Other important lineage receptors, such as CD4 (T4), CD8A (T8) and CD79A (B cells), have copy numbers between 200,000 and 1 million (**Fig. 2c**,**d**). Although we found that some transcription factors were present in less than a 100 copies per cell, transcription factors as

a whole had a median of 10,000 copies. At the top of the distribution, IRF7 and IRF8, key transcriptional regulators of type I interferons, were quantified in plasmacytoid DCs (pDCs) with 200,000 and 1 million copies, respectively.

Cell types are commonly defined by exclusive or combinatorial expression of cell surface markers. We applied Lasso regression analysis[29] to investigate previously unknown cell surface receptor combinations for cell lineages as well as for individual cell types. This revealed

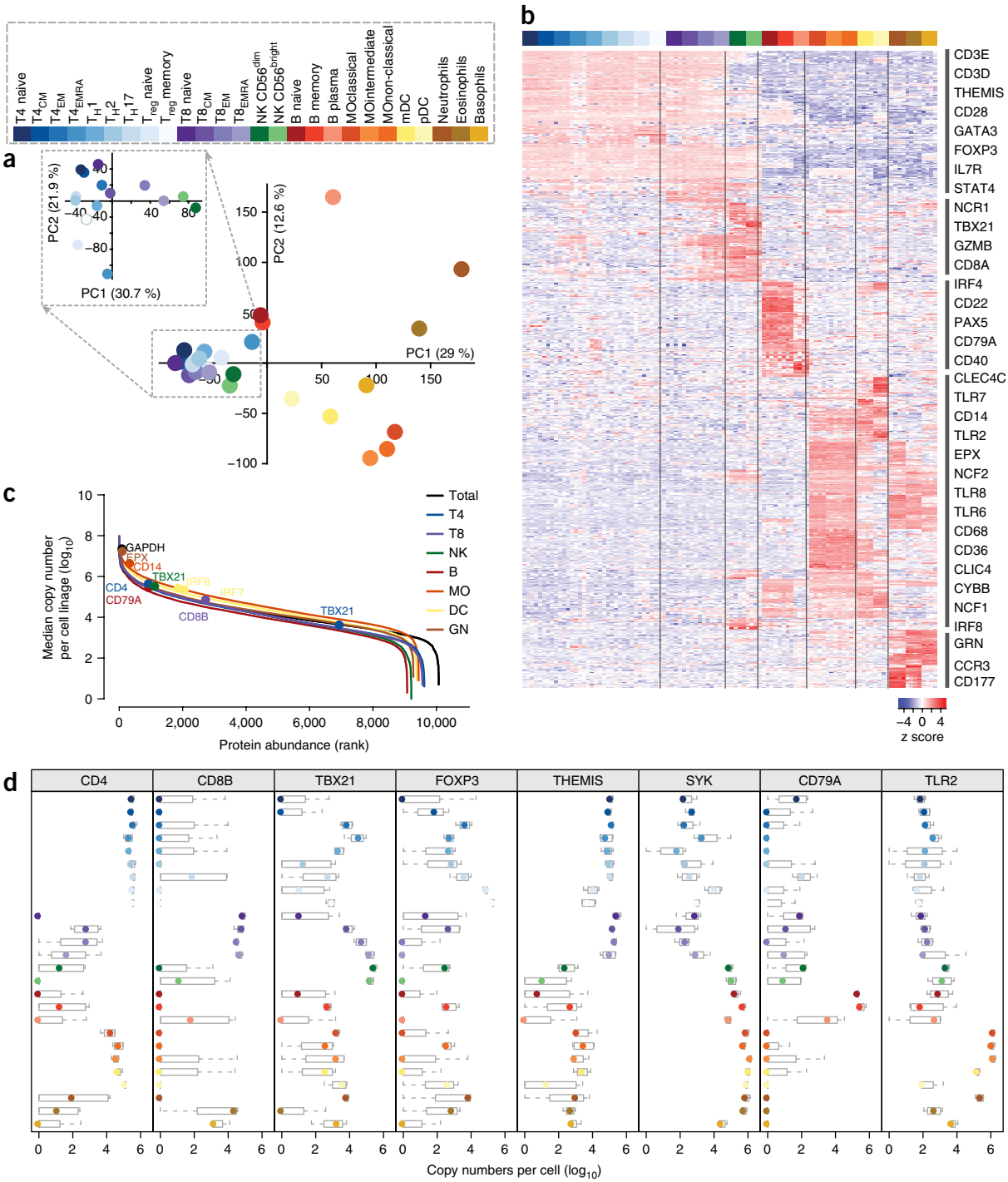

Figure 2 Immune cell relationships are defined by lineage-specific signature proteins. (**a**) Principal component analysis (PCA) of median protein abundances (ANOVA, Benjamini-Hochberg (BH) $P < 0.01$, df1 = 25, df2 = 78) for different immune cell types. Inset, PCA of T lymphocytes (T4 and T8) and NK cells. Different cell types are indicated by colors in the legend. (**b**) Heat map with the top 100 significantly differentially expressed proteins for each cell lineage (T4, T8, NK, B, MO, DC and GN) (two-tailed Welch's *t* test, FDR < 1%, $S_0 = 1$, $n = 4$ from independent donors). Rows were clustered using complete linkage with Euclidean distance and columns were ordered by lineage as shown in the legend. Protein names are examples of proteins present in the indicated clusters. (**c**) Dynamic range of the proteomes of measured cell lineages, based on median estimated copy numbers. (**d**) Copy number profiles of representative lineage marker proteins. Box-plot elements: center point, median; box limits, first to third quartile (Q1 to Q3); whiskers, from Q1–1.5 × interquartile range (IQR) to Q3+1.5 × IQR.

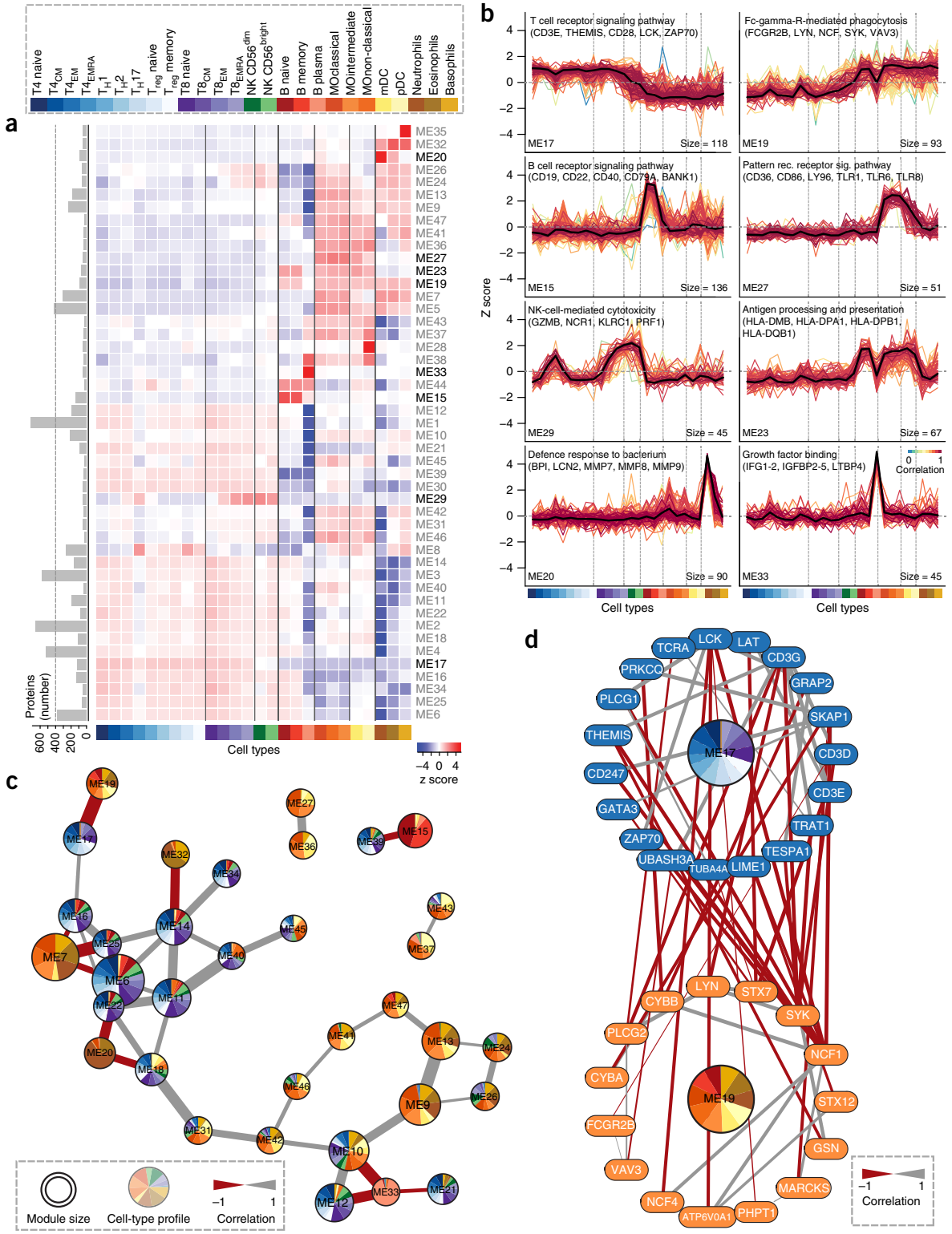

Figure 3 Functional modules of the immune system. (**a**) Average expression levels of 47 modules (ME1–47) defined by weighted gene cluster network analysis (WGCNA) across cell types. Bars on the left indicate the number of proteins in each module. Modules listed in bold are depicted in **b**. (**b**) Protein expression profiles of selected modules across cell types. Each line represents one protein. Average profile is shown in black. Module sizes are indicated at the bottom right. Specific functions of each module assigned by annotation enrichment analysis (Fisher exact test, BH FDR < 5%; **Supplementary Table 7**). (**c**) Immune system module network. Edges represent Pearson correlation coefficients. Only edges with absolute Pearson correlation > 0.8 are shown. Modules 1–5 and modules with no connecting edges are not shown. Pie chart colors correspond to cell types and the size of the slices corresponds to the module protein expression profile. Module size is represented by node size. (**d**) Network of proteins in module 17 (ME17, blue) and 19 (ME19, orange) annotated for the enrichment terms 'T cell receptor signaling pathway' and 'Fc-gamma-R-mediated phagocytosis', respectively. Edges represent Pearson correlation coefficient for each protein pair.

the chemokine receptor CX3CR1, identified on T cells with low proliferation, but high cytotoxic capacity[30], to be an lineage receptor for CD8 CD45RA$^+$ effector memory (T_{EMRA}) cells (**Supplementary Fig. 3d**). For rare cell populations, exclusive markers are advantageous for staining and sorting; we identified plasmalemma vesicle-associated protein (PLVAP) and multiple epidermal growth factor-like domains protein 10 (MEGF10) to be such unique cell surface markers for B cell plasma blasts (**Supplementary Fig. 3d**), and confirmed PLVAP expression by flow cytometry (**Supplementary Fig. 3e**).

Functional modules of the immune system

We defined cell-type-resolved functions of the immune system in an unbiased clustering approach[31] by grouping proteins with correlating expression patterns and assigning their functional properties by GO enrichment analysis (**Fig. 3a**,**b** and **Supplementary Table 7**). Among the 47 resulting functional modules (ME), module ME27 contains the TLR proteins TLR1, TLR6 and TLR8 and is enriched for 'inflammatory response' and 'pattern recognition receptors signaling pathway'. These properties are characteristic of monocytes and

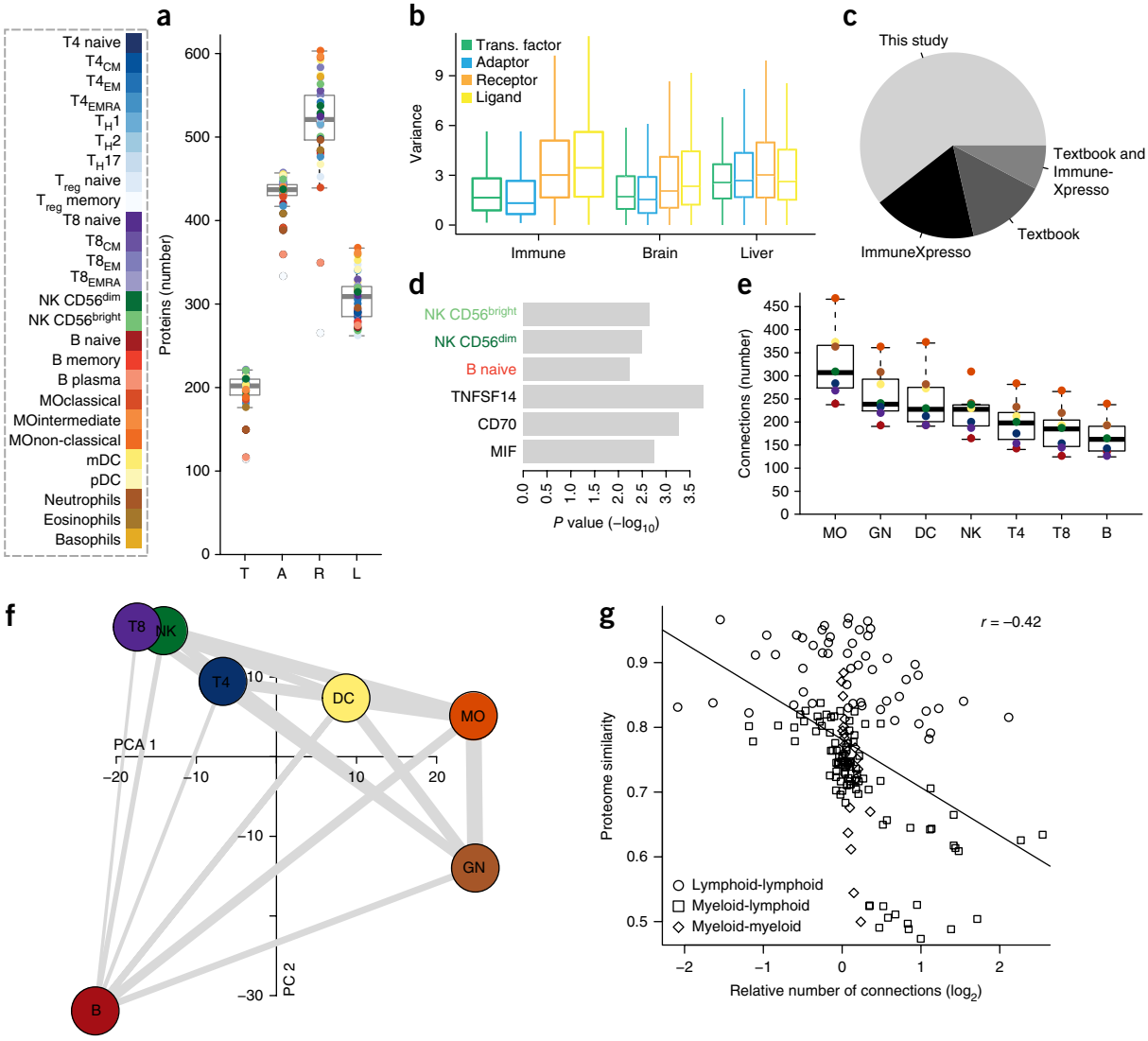

Figure 4 Communication network of immune cells. (**a**) Number of detected transcription factors (T), adaptor molecules (A), receptors (R) and ligands (L) per cell type. (**b**) Expression variances of all immune cell types according to the categories in **a** in steady state compared with cell-type-resolved brain and liver proteomes[38,39]. The variance for each protein was calculated from expression values across all cell types. Activated immune cell states as well as immune cell types from brain (microglia) and liver (Kupffer cells) were excluded from the analysis. Box-plot outliers are not shown in this panel. (**c**) Outgoing connections in our study compared with Textbook and ImmuneXpresso (Online Methods). The pie chart represents all outgoing connections from our study with connections already captured by the two different data sets indicated in slices. (**d**) Understudied cell types and cytokines from outgoing connections. Significantly enriched cell types and cytokines in our study compared to Textbook and ImmuneXpresso identified by Fisher's exact test with BH FDR < 5%. (**e**) Number of intercellular connections between cells of each lineage with all other lineages. (**f**) Immune cell lineage connectivity network. PCA of immune cell lineages proteomes filtered for receptors and ligands. Connecting line strength indicates the number of intercellular connections between lineages. (**g**) Pearson correlation coefficient of cell type median expression values and relative number of intercellular connections of cell type pairs in steady state. The number of interactions between each cell type pair was normalized by the number of receptor and ligand expressed on both cell types. Cell type pairs of the same lineage were excluded. Linear fit (solid line) and Pearson correlation (*r*) are shown (top right). Box-plot elements: center line, median; box limits, first to third quartile (Q1 to Q3); whiskers, from Q1–1.5 × IQR to Q3+1.5 × IQR; points, outliers.

dendritic cells, which specialize in sensing microbial- and danger-associated molecular patterns[32].

Other important modules include ME17, which is described by 'T cell receptor signaling pathway' and includes T-cell-specific proteins such as CD3ε and THEMIS. The specialized function of B lymphocytes, DCs and monocytes to present antigens to T lymphocytes is captured in ME23, which is highly enriched for MHC class II proteins. Module ME29 contains proteins that are highly abundant in NK, CD8 T effector memory (T_{EM}), T_{EMRA} and CD4 T_{EMRA} cells. Cytolytic proteins, such as granzyme B and perforin, are enriched in this module, suggesting that cytotoxic CD4 T cells belong to the $CCR7^{-}CD45RA^{+}$ memory compartment of CD4 T lymphocytes (**Supplementary Fig. 4a**). Low expression of CD27 and CD28 protein, characteristic of the CD4 T cell cytotoxic phenotype[33], support this finding. The transcription factors TBX21, EOMES, and HOPX are also present in ME29. Although the first two have been reported to be essential for the development of NK cells[34,35], the latter—already linked to T cell effector memory formation[36,37]—might also be involved in NK development or cytolytic immune function. Taken together, our proteome resource and assignment of proteins to functional modules recapitulates many known relationships of the immune system and holds the potential for the generation of new hypotheses.

All of the modules clustered in four groups, reflecting their functional lineage relationships (**Supplementary Fig. 4b**). T and NK cell modules, as well as monocyte and dendritic cell modules, overlapped, indicating shared functionalities; however, each lineage has its unique modules. Anti-correlated modules point toward unique and mutually exclusive cellular functions. For instance, ME20 in neutrophils and ME33 in plasma blasts are clusters that are not shared with any other cell type that are enriched for proteins in 'defense response to bacterium' or 'growth factor binding', respectively (**Fig. 3c**). The highest anti-correlating modules are ME17, representing proteins with adaptive immune functions enriched for 'T cell receptor signaling pathway' and 'lymphocyte differentiation' versus ME19, with proteins involved in innate immune functions such as 'phagocytosis' or 'response to fungus' (**Supplementary Fig. 4c,d**). These two arms of the immune system are also captured in ME25 versus ME7. Module-level anticorrelation could further be resolved at the protein level, and this showed near mutual exclusivity of cellular receptors (CD3ε versus CYBB), adaptors (LCK versus SYK) and transcription factors (STAT4 versus SPI1) (ME19 and ME17; **Fig. 3d**).

Architecture of intercellular communication

To assess which level of cellular signal transduction contributes most to cell-type specificity, we categorized proteins as transcription factors, adaptor molecules, receptors and secreted molecules and investigated how their expression differed between cell types (**Fig. 4a** and **Supplementary Fig. 5a**). The average expression variances of receptors and secreted ligands were nearly twice those of adaptors and transcription factors. These findings indicate that immune cells are more specialized intercellularly than intracellularly. This appears to be a special feature of the immune system, as we did not observe it in the cell-type-resolved proteomes of the brain and the liver[38,39] (**Fig. 4b**).

To further characterize the architecture of intercellular crosstalk, we compared the number of connections established by receptor-receptor or ligand-receptor interactions between cell types. Potential interactions recorded in the STRING database[40] are not cell-type resolved and would be compatible with a network of more than 4×10^5 connections between the 460 receptors and 300 secreted ligands identified in our study, counting interactions as different if they occurred between different cell types. On the basis of the proteomics

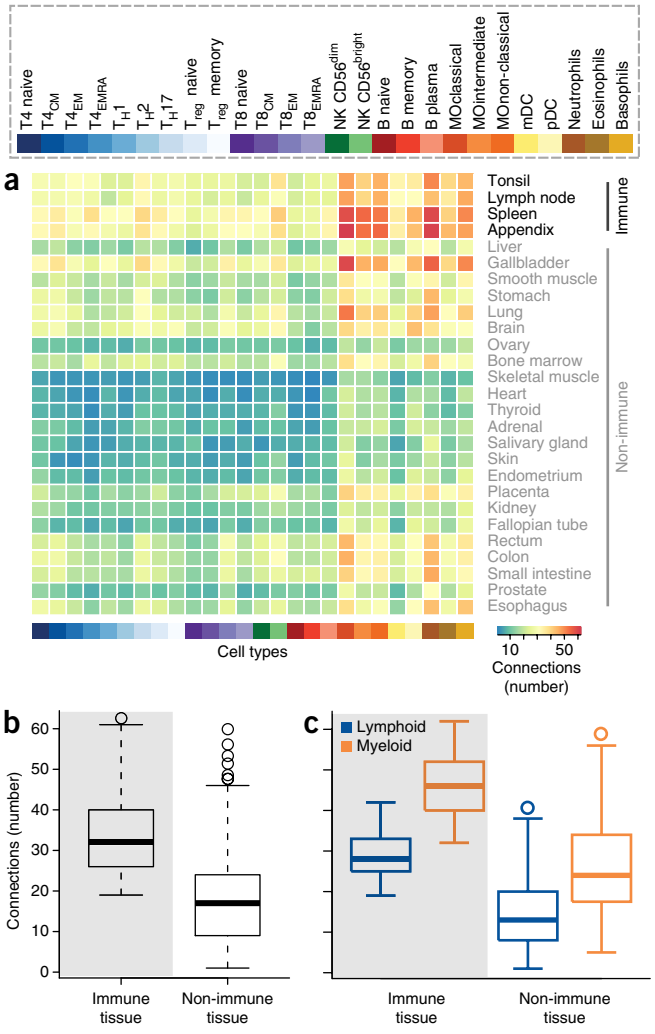

Figure 5 Intercellular connections between immune cells and tissues. (**a**) Number of intercellular receptor-receptor connections between human immune cell types and tissues. Tissue RNA sequencing data were retrieved from ref. 42. (**b**) Number of connections to immune or non-immune tissues from **a**. (**c**) Comparison of lymphoid and myeloid immune cell connections to tissues. Box-plot elements: center line, median; box limits, first to third quartile (Q1 to Q3); whiskers, from Q1−1.5 × IQR to Q3+1.5 × IQR; points, outliers.

data, we prioritized intercellular connections formed by proteins with high and reproducibly measured expression in the respective cell types. We further considered 'Textbook' knowledge[41] derived from a sophisticated analysis of current literature (Online Methods) to reduce our communication network to the 180,000 most confident connections of protein pairs in different cell types (**Supplementary Fig. 5b,c** and Online Methods). This analysis revealed a plethora of previously unknown cell-type-specific interactions for major cytokines (**Fig. 4c** and **Supplementary Fig. 5d**), as compared with those reported in 1.3 million published PubMed abstracts (ImmuneXpresso, Online Methods). Overrepresented in our communication network were outgoing connections in which NK and naive B cells function as senders as well as those involving tumor necrosis factor (TNF) family members TNFSF14 and TNFSF7 (CD70) and macrophage migration inhibitory factor (MIF) (**Fig. 4d** and **Supplementary Fig. 5e**). We validated one out-going and one in-going intercellular connection that are not described in the current literature. First, we found Resistin

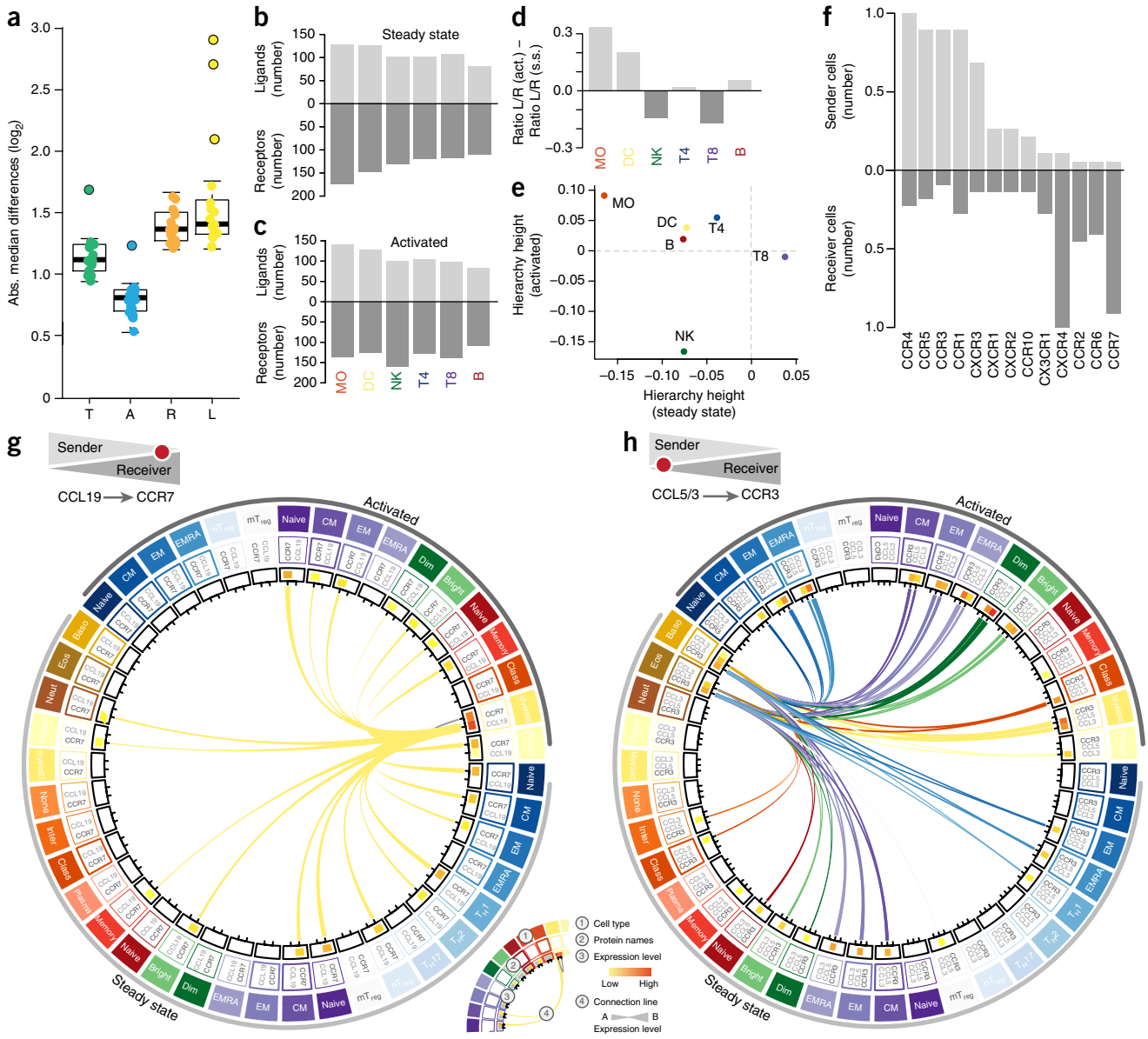

Figure 6 Dynamics of senders and receivers of biological information. (**a**) Absolute median *t* test expression differences for transcription factors (T), adaptor molecules (A), receptors (R) and ligands (L) between cells in steady and activated states. (**b**) Receptor and ligand count for the indicated cell types in steady state. (**c**) Receptor and ligand count for the indicated cell types in activated state. (**d**) Activation induced change of cellular sending and receiving capacity. Ratio of R/L (act., activated state) to R/L (s.s., steady state) are displayed. (**e**) In-out degree hierarchy height (h = (O – I) / (O + I)) of immune cell lineages in steady and activated state. (**f**) Number of cellular senders and receivers for all quantified chemokine receptors. The total numbers are divided by the maximum sender or receiver number, respectively. (**g**,**h**) Intercellular communication structures for chemokine receptors (CCRs) and their ligands (CCLs): CCR7-CCL19 (**g**) and CCR3-CCL5/CCL3 (**h**). Outer circle, cell types; middle circle, receptor and ligand gene names; inner circle, expression levels. Connecting lines indicate biological information flow with the line color corresponding to the cellular sender. Box-plot elements: center line, median; box limits, first to third quartile (Q1 to Q3); whiskers, from Q1–1.5 × IQR to Q3+1.5 × IQR; points, outliers.

(RETN), which is known to be secreted by monocytes, to also be secreted by activated memory B cells and confirmed this finding by ELISA (**Supplementary Fig. 5f**). Second, we found that the receptor CSF1R, which is known to be expressed on CD4 T cells, was also activated by IL-34 on CD4 memory T cells (**Supplementary Fig. 5g**). In comparison with RNAseq (**Supplementary Table 8**) and microarray data, we observed median correlation of approximately 0.5 for matched cell types and protein profiles (**Supplementary Fig. 6a–d** and **Supplementary Table 9**). Notably, proteome- and transcriptome-based intercellular communication networks were similar in size, but only shared 50% of their connections (**Supplementary Fig. 6e**). We conclude that the current knowledge on immune cell interplay is still incomplete, and our analysis revealed systematically understudied intercellular immune cell signaling paths.

Focusing on the major cell lineages (T4, T8, NK, B, MO, DC, GN), we observed that myeloid immune cells had more connections than lymphoid immune cells (**Fig. 4e**,**f**). Monocytes establish on average 300 connections with other immune cells, compared with T lymphocytes, which on average establish only 200 connections. Analyzing all cell types in this way established the general trend that

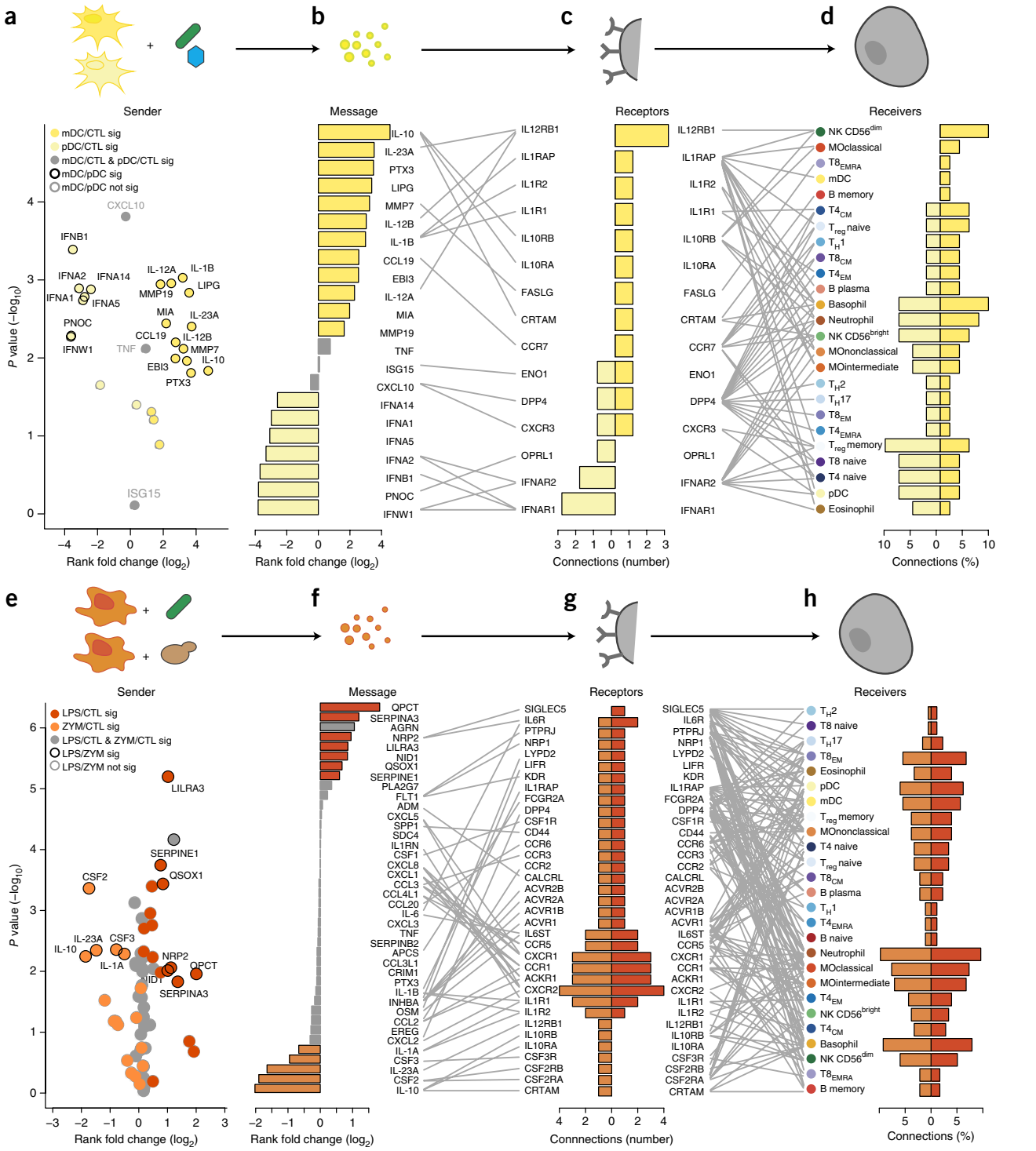

Figure 7 Cell-type- and context-dependent communication. (**a**) Significantly (two-tailed Welch's *t* test, FDR<5%, $S_0 = 1$, $n = 4$ independent donors) secreted proteins from activated mDCs versus pDCs plotted by rank of secretion. Only significantly secreted proteins compared with unstimulated cells and those proteins annotated as ligand are shown. (**b**) Same as in **a** ordered by rank. (**c**) Receptors interacting with the significantly secreted proteins from mDCs or pDCs. Receptors are ordered by number of ligands and cell-type specificity. Ligand-receptor pairs are connected with a gray line (StringDB score > 0.7). (**d**) Relative count of connections between secreted molecules from mDCs or pDCs with other immune cell types in their steady state. Receptor expression on receiving cell types are indicated with a connecting gray line. (**e**) Significantly (two-tailed Welch's *t* test, FDR < 5%, $S_0 = 1$, $n = 5$ independent donors) secreted proteins from cMOs activated either with LPS or ZYM plotted by rank of secretion. Only significantly secreted proteins compared to unstimulated cells and those proteins annotated as ligand are shown. (**f**) Data are presented as in **d** ordered by rank. (**g**) Receptors interacting with the significantly secreted proteins from cMO stimulated with LPS or ZYM. Receptors are ordered by number of ligands and cell-type specificity. Ligand-receptor pairs are connected with a gray line (StringDB score > 0.7). (**h**) Relative count of connections between cMOs and other immune cell types following either LPS or ZYM activation. Receptor expression on receiving cell types are indicated with a connecting gray line.

less-related cell types (as determined by proteome correlation), such as neutrophils to naive B cells, are more highly connected (**Fig. 4f**,**g**). Thus, the immune system links distinct cell-type-resolved functions via strong communication structures.

To further elucidate how immune cells connect with organs and tissues, we expanded our network using an atlas of 32 different human organs and tissues[42], focusing only on receptor-receptor intercellular interactions (Online Methods). Immune cells had significantly more connections to tissues related to the immune system (for example, lymph node, spleen, tonsil, and appendix) than to non-lymphoid tissues ($P < 10^{-15}$; **Fig. 5a**,**b**). Furthermore, myeloid immune cells established more connections with non-immune tissues than lymphoid immune cells (**Fig. 5c**). This presumably reflects cellular preference for tissue residence as phagocytic cells, which scan almost the entire body for foreign or altered self and T lymphocytes, reside primarily in the lymphatic system.

Dynamics of senders and receivers of biological information

Biological messages are dynamic and depend on cell type as well as context, and this is captured in our data set by the dynamic changes of immune cell proteomes following activation. We observed dominant changes at the level of receptors and ligands, which is the level that determines context-dependent adaptation of cellular communication behavior (**Fig. 6a** and Online Methods). To systematically evaluate cellular sender or receiver roles, we assessed the expression and secretion of protein ligands (out degree, O) as well as the expression of the corresponding receptors (in degree, I), respectively. Our analysis highlights the degree to which myeloid cells engage in intercellular communication, as these cells expressed more receptors and ligands than lymphoid cells (**Fig. 6b**,**c**). Following activation, MO and DCs showed an increase in their ligand-receptor ratio (L/R), whereas the ratio for cytolytic cell types (NK, T8) decreased (**Fig. 6d**). Thus, immune cells seem to employ distinct communication strategies: antigen-presenting cells (MO and DC) changed both quantity and diversity of their ligands and reduce their receptor diversity, whereas cytolytic cell types changed only the quantity of their ligands and increased their receptor diversity (**Fig. 6a–d**). We compared the hierarchy height ($h = (O - I) / (O + I)$)[43] of intercellular signal transduction before and after activation, where senders have $h > 0$ and receivers $h < 0$ (**Fig. 6e**). Following activation, antigen-presenting cells increased their hierarchy height. In particular, MO showed the most prominent change by being at the bottom in steady state, but then move to the top of the hierarchy after activation. Together with the functional module analysis described above, our proteomic analysis sheds light on the dynamic cellular properties of antigen-presenting cells that enable them to alert surrounding cells after sensing microbes or danger.

Apart from the global analysis of communication networks, our data allow the deduction of sender and receiver patterns with different biological functions. Using the established framework at the level of individual proteins, we observed two general types of communication structures for chemokine receptors and their ligands. Broadly expressed receptors, such as CCR7, received information from few sending cells, whereas receptors with cell-type-specific expression patterns, such as CCR3, received information from many cell types (**Fig. 6f–h**). These opposing communication structures are associated with different paracrine functions of transmitted messages. For example, diverse recipient cells expressing CCR7 are attracted to lymph nodes[44], whereas receptor ligation of CCR3 exclusively recruits polymorphonuclear phagocytes to inflamed areas[45]. Thus, our analysis establishes on a global scale that chemokines require specificity either on the sending or receiving end, a principle of directed asymmetric information exchange.

Cell-type- and context-resolved communication structures

To learn more about the syntax of intercellular communication, we compared cell-type- and context-dependent innate immune responses employing in depth and quantitative secretome measurements[14]. We analyzed the outcome of the biological messages by tracing the paracrine intercellular information flow of the secreted proteins. Using dendritic cells as an example for cell-type specificity, we found that, following broad TLR activation, myeloid DCs (mDCs) secreted many different inflammatory cytokines, whereas pDCs secreted primarily type I interferons (**Fig. 7a**,**b**, **Supplementary Fig. 7a**,**b** and **Supplementary Table 10**). Using our communication framework, we established that mDCs target numerous receptors and form an intercellular connection network with multiple channels to diverse recipient cells (**Fig. 7c**,**d**). Intracellularly, the targeted signaling pathways in recipient cells amplified the primary pro-inflammatory signature originating from mDCs by intracellular signal integration and converted it into a defined secondary signature by employing distinct JAK-STAT signaling pathways (**Supplementary Fig. 7c**). By contrast, pDCs fostered a defined single-layer antiviral response in recipient cells by engaging mainly with the interferon alpha receptor (**Fig. 7c**,**d** and **Supplementary Fig. 7c**).

To dissect context-dependent paracrine responses, we further compared responses of classical monocytes (cMOs, $CD14^+CD16^-$) to challenges that require distinct actions, such as fungal (zymosan, ZYM) versus bacterial (lipopolysaccharide, LPS) encounters. Key cytokines secreted following ZYM exposure included members of the colony-stimulating factor (CSF) and the anti-inflammatory IL-10 families that connect cMOs with myeloid cells, stimulating a strong intracellular NF-kB signature (**Fig. 7e**,**f**, **Supplementary Fig. 7d**,**e** and **Supplementary Table 11**). Conversely, LPS activation released serine protease inhibitors associated with tissue remodeling as well as cell-matrix interactions (**Fig. 7e**,**f**). Both activation events elicited a strong shared secretory program comprising almost 20 pro-inflammatory proteins, including interleukins, chemokines and TNF family members such as IL-6, CXCL8 or TNF (**Fig. 7e**,**f**). Based on the targeted receptors and their cell-type-specific expression, the shared pro-inflammatory program conveyed immune instructions directed toward multiple cellular recipients (**Fig. 7g**,**h**). It converged intracellularly with pleiotropic properties ranging from pro-inflammatory over chemoattractant to cell-death-inducing functions (**Supplementary Fig. 7f**).

DISCUSSION

Intercellular communication between diverse cell types orchestrates a plethora of physiological functions. However, a system-wide view of how cells exchange information had not been established because of technological limitations. We attempted to globally assess the dynamic communication structures of the immune system. We successfully measured the proteomes of all major human immune cell types present in the blood in steady and selected activation states at great sensitivity and quantitative depths. With more than 80% coverage of immune-system-annotated proteins and around 70% coverage for immune-relevant signaling molecules such as transcription factors, adaptor molecules, cell surface receptors and secreted molecules, we provide the most comprehensive cell-type-resolved human proteome resource to date. Some of the remaining proteins appear to be confined to non-circulating immune cells or to immune states that were not sampled. The immune system displays an almost infinite diversity in cell types and activation states, and achieving completeness in both dimensions is particularly challenging for human cells[46].

By combining total proteome with secretome measurements and a bioinformatics framework, we constructed a social network of human immune cells and deduced the logic of immune cell interplay. This enabled the discovery of potential new cell type markers, protein associations to functional modules and intercellular connections. As a result of the imperfect correlation of proteome to transcriptome, communication networks constructed from RNA data may insufficiently capture intercellular signaling on the level of proteins. Our proteomics-based analysis revealed that the immune cell communication network is more complex than what is captured by the current literature. Common messages are used between many more cell types, highlighting promising research directions, such as NK cells, which appear to have additional functions in orchestrating immune responses. As examples for the discovery potential of our analysis, we validated one previously unknown outgoing connection (B cells → RETN) and one ingoing connection (IL-34 → T4 memory cells). However, as a result of the automated extraction of intercellular connections from the current literature (Online Methods) and its mapping to our proteomics communication network, inaccuracies in total as well as individual understudied connections can be expected. Given that connections in our network are defined by annotated evidence (Online Methods), individual pairwise interactions may not be captured correctly. Although false positives may derive from incorrect or outdated annotations, false negatives may derive from not discovered or not deposited interactions as well as by undetected proteins. To avoid overestimation, we discarded 50% of the possible interactions of the communication network and report only those with high confidence using 'Textbook' knowledge as reference. To evaluate the probability of individual intercellular connections, however, further information on cell type abundance and tissue distributions could be considered in the future.

Our results demonstrate that immune cell signaling is more diverse inter- than intracellularly in comparison with other organs such as brain and liver. To adapt to extracellular cues, intercellular immune signaling is further specifically tuned by the regulation of ligand and receptor expression. Thus, proteomics highlight how immune cell diversity and plasticity together shape protective immunity. We found that antigen-presenting cells increased their capacity to send information in response to changing conditions, forming the top of the intercellular signaling hierarchy. In contrast, cells with cytotoxic functions decreased their hierarchy height. Dependent on the biological messages that are sent, however, different communication structures are employed: as a general underlying principle, we found that communication was restricted to a limited number of sending or receiving cell types for any given cytokine, presumably to achieve specificity in intercellular signal transduction at a minimum cost of resources. On the level of individual cells, intercellular communication structures may differ from those obtained from sorted cell populations. Single-cell technologies to study cell-to-cell crosstalk on the protein level may further increase our understanding of communication diversity and heterogeneity in the future.

According to our data and analysis, the current classification of immune cell identity based on expression of lineage markers, mainly transcription factors and cell surface receptors, would gain resolution by including secreted ligands and corresponding intercellular connections. Moreover, the interpretation of biological messages transmitted during disease, that is, cocktails of cytokines in case of complex immune responses, could be improved by integrating information on targeted cell types and their activated intracellular signaling programs. Current experimental approaches are largely focused on a single or a few different proteins, which makes it challenging to dissect the logic of inflammatory programs in redundant versus cooperative immunological signals. By contrast, the proteomics approach naturally deals with these challenges and provides a systems-wide, interconnected network of healthy immune functions, representing a reference to identify deviating communication patterns associated with pathology.

The immune system's diversity of cellular components and complementation of cell-type-resolved competences comprises features immanent to social communities. Considering heterogeneity, plasticity and networking of cell types, the immune system is tremendously complex. Given the mobility of the cellular players, it presents an ideal example for studying how physiological functions are orchestrated by intercellular communication. Our analysis reveals some of the concepts that are shared between the immune system and other types of social networks. These include communication structures between individual players with defined hierarchy and logic aimed at achieving complex tasks.

Our study should be considered a starting point for the investigation of intercellular information exchange at higher cellular and contextual resolution. An expansion of the communication network to tissue resident immune cells and their direct interactions would be particularly exciting, although accessibility and the number of cells obtainable from healthy human donors remain a challenge. Future studies should evaluate whether the sender-receiver principles of cellular communication deduced here can be extended to other organs and multicellular organisms. The availability of further system-wide information that can also be gained by proteomics, such as extracellular protein-protein interactions or post-translational modifications in intracellular signaling networks, could further refine the architecture of our communication network.

In conclusion, our findings establish a social immune cell network on the level of proteins and provides an experimental and conceptual framework to deduce fundamental communications structures and interpret context and cell-type-dependent immune dynamics.

METHODS

Methods, including statements of data availability and any associated accession codes and 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 R. Scheltema for MS assistance, K. Mayr, I. Paron, S. Gabriele, S. Dewitz and M. Dodel for technical assistance, M. Wierer and P. Sinitcyn for support with RNAseq analysis, M. Oroshi and C. Schaab for computer and database support, A. Zychlinsky for critical review of the manuscript, and J. Geddes, A. Frauenstein, M. Phulphagar and L. Kühn for helpful discussions. The work was funded by the Max Planck Society for the Advancement of Science. R.G. was supported by a grant from the Swiss SystemsX.ch initiative, evaluated by the Swiss National Science Foundation. A.L. is supported by the Helmut Horten Foundation.

AUTHOR CONTRIBUTIONS

J.C.R., R.G., D.J. and T.W. performed flow and cell culture experiments. J.C.R. developed and implemented the bioinformatics methods. J.C.R. and F.M. conceived the data analysis and interpreted the data. D.H. assisted in data analysis and implemented the website. S.S.S. and K.K. assisted in data analysis and provided the Textbook and ImmuneXpresso data sets. F.M. and M.M. conceived the study. F.M., A.L. and F.S. supervised the experiments. F.M., J.C.R. and M.M. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Subramanian, N., Torabi-Parizi, P., Gottschalk, R.A., Germain, R.N. & Dutta, B. Network representations of immune system complexity. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **7**, 13–38 (2015).
2. Rivera, A., Siracusa, M.C., Yap, G.S. & Gause, W.C. Innate cell communication kick-starts pathogen-specific immunity. *Nat. Immunol.* **17**, 356–363 (2016).
3. Hotson, A.N. *et al.* Coordinate actions of innate immune responses oppose those of the adaptive immune system during Salmonella infection of mice. *Sci. Signal.* **9**, ra4 (2016).
4. Dobrin, R. *et al.* Multi-tissue coexpression networks reveal unexpected subnetworks associated with disease. *Genome Biol.* **10**, R55 (2009).
5. Frankenstein, Z., Alon, U. & Cohen, I.R. The immune-body cytokine network defines a social architecture of cell interactions. *Biol. Direct* **1**, 32 (2006).
6. Shen-Orr, S.S. *et al.* Towards a cytokine-cell interaction knowledgebase of the adaptive immune system. *Pac. Symp. Biocomput.* **2009**, 439–450 (2009).
7. Heng, T.S., Painter, M.W. & Immunological Genome Project, C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9**, 1091–1094 (2008).
8. Novershtern, N. *et al.* Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).
9. Bendall, S.C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
10. Parnas, O. *et al.* A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686 (2015).
11. Paul, F. *et al.* Transcriptional heterogeneity and lineage commitment in myeloid progenitors. *Cell* **163**, 1663–1677 (2015).
12. Schwanhausser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
13. Jovanovic, M. *et al.* Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. *Science* **347**, 1259038 (2015).
14. Meissner, F., Scheltema, R.A., Mollenkopf, H.J. & Mann, M. Direct proteomic quantification of the secretome of activated immune cells. *Science* **340**, 475–478 (2013).
15. Altelaar, A.F., Munoz, J. & Heck, A.J. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **14**, 35–48 (2013).
16. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).
17. Meissner, F. & Mann, M. Quantitative shotgun proteomics: considerations for a high-quality workflow in immunology. *Nat. Immunol.* **15**, 112–117 (2014).
18. Nagaraj, N. *et al.* System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol. Cell Proteomics* **11**, M111 013722 (2012).
19. Schaab, C., Geiger, T., Stoehr, G., Cox, J. & Mann, M. Analysis of high accuracy, quantitative proteomics data in the MaxQB database. *Mol. Cell Proteomics* **11**, M111 014068 (2012).
20. Scheltema, R.A. *et al.* The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Mol. Cell. Proteomics* **13**, 3698–3708 (2014).
21. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
22. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526 (2014).
23. Kim, M.S. *et al.* A draft map of the human proteome. *Nature* **509**, 575–581 (2014).
24. Jojic, V. *et al.* Identification of transcriptional regulators in the mouse immune system. *Nat. Immunol.* **14**, 633–643 (2013).
25. Cox, J. & Mann, M. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics* **13** (Suppl. 16), S12 (2012).
26. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nat. Immunol.* **9**, 503–510 (2008).
27. Zhang, N. & Bevan, M.J. CD8(+) T cells: foot soldiers of the immune system. *Immunity* **35**, 161–168 (2011).
28. Wisniewski, J.R., Hein, M.Y., Cox, J. & Mann, M. A 'proteomic ruler' for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics* **13**, 3497–3506 (2014).
29. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).
30. Bottcher, J.P. *et al.* Functional classification of memory CD8$^+$ T cells by CX3CR1 expression. *Nat. Commun.* **6**, 8306 (2015).
31. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
32. Bieber, K. & Autenrieth, S.E. Insights how monocytes and dendritic cells contribute and regulate immune defense against microbial pathogens. *Immunobiology* **220**, 215–226 (2015).
33. Appay, V. The physiological role of cytotoxic CD4$^+$ T cells: the holy grail? *Clin. Exp. Immunol.* **138**, 10–13 (2004).
34. Gordon, S.M. *et al.* The transcription factors T-bet and Eomes control key checkpoints of natural killer cell maturation. *Immunity* **36**, 55–67 (2012).
35. Townsend, M.J. *et al.* T-bet regulates the terminal maturation and homeostasis of NK and Valpha14i NKT cells. *Immunity* **20**, 477–494 (2004).
36. Bezman, N.A. *et al.* Molecular definition of the identity and activation of natural killer cells. *Nat. Immunol.* **13**, 1000–1009 (2012).
37. Albrecht, I. *et al.* Persistence of effector memory Th1 cells is regulated by Hopx. *Eur. J. Immunol.* **40**, 2993–3006 (2010).
38. Sharma, K. *et al.* Cell-type- and brain-region-resolved mouse brain proteome. *Nat. Neurosci.* **18**, 1819–1831 (2015).
39. Azimifar, S.B., Nagaraj, N., Cox, J. & Mann, M. Cell-type-resolved quantitative proteomics of murine liver. *Cell Metab.* **20**, 1076–1087 (2014).
40. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
41. Dembic, Z. *The Cytokines of the Immune System: The Role of Cytokines in Disease Related to Immune Response* (Elsevier Science, 2015).
42. Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
43. Gerstein, M.B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489**, 91–100 (2012).
44. Forster, R., Davalos-Misslitz, A.C. & Rot, A. CCR7 and its ligands: balancing immunity and tolerance. *Nat. Rev. Immunol.* **8**, 362–371 (2008).
45. Ponath, P.D. *et al.* Molecular cloning and characterization of a human eotaxin receptor expressed selectively on eosinophils. *J. Exp. Med.* **183**, 2437–2448 (1996).
46. Kim, C.C. & Lanier, L.L. Beyond the transcriptome: completion of act one of the Immunological Genome Project. *Curr. Opin. Immunol.* **25**, 593–597 (2013).

ONLINE METHODS

Blood samples and cell sorting. Blood from healthy donors (~450 ml) was obtained from the Swiss Blood Donation Center of Basel and Lugano, and used in compliance with the Federal Office of Public Health (Authorization no. A000197/2 to F.S.). Peripheral blood mononuclear cells (PBMC) were isolated using Ficoll-Plaque Plus (GE Healthcare) sedimentation. Cells were positively selected using magnetic microbeads (Miltenyi Biotec). Sorting was performed with a FACSAria flow cytometer (BD Biosciences) and cell types were sorted to >98% purity (**Supplementary Fig. 1** and **Supplementary Tables 1** and **12**). Cells were either collected in complete medium or for proteome measurement washed in PBS, snap frozen and stored at −80 °C. Erythrocytes were isolated by removing the fraction with lower density as previously described[47]. The erythrocyte packed phase at the bottom after the density centrifugation was washed twice with ice cold PBS, snap frozen and then stored at −80 °C until further use. Platelets were isolated by OptiPrep density gradient purification[48]. 3 ml whole blood was layered over 5 ml of the density solution (5 ml of 1.320 g/ml 60% iodixanol stock solution; OptiPrep density gradient medium, Sigma-Aldrich) and mixed with 22 ml diluent (0.85% NaCl, 20 mM HEPES-NaOH, pH 7.4, 1 mM EDTA). The sample was centrifuged at 350 *g* for 15 min at 20 °C in a swinging bucket rotor with the brake switched off. Platelets were harvested from the broad turbid band below the interface. Platelets were washed twice with PBS and stored at −80 °C until further use.

***In vitro* activation assays.** Flow sorted cells were cultured in RPMI-1640 medium supplemented with 2 mM glutamine, non-essential amino acids (0.1 mM of each amino acid), 1 mM sodium pyruvate, penicillin (50 U/ml), streptomycin (50 µg/ml; all from Invitrogen) and 5% (v/v) human serum (Swiss Blood Center). T lymphocytes were stimulated for 48h with plate-bound anti-CD3 (5µg/ml, clone TR66) and anti-CD28 (1µg/ml CD28.2, BD Biosciences) and expanded for 48 h with IL-2 (500 U/ml). NK cells were stimulated with the NK Cell Activation/Expansion Kit, according to the manufacturer's instructions (130094483, Miltenyi). In brief, NK cells were co-cultured in the presents of CD2 and CD355 (NKp46) coated beads together with IL-2 (500 U/ml) for up to 14 d. B cells were stimulated using a cocktail of Goat F(ab′)2 anti-human Lambda/Kappa (each 2 µg/ml), F(ab′)2 Goat anti-mouse IgG Fc (3 µg/ml), CpG (100ng/ml), anti-CD40 (6µg/ml), for 96h. Monocytes and dendritic cells were stimulated with the TLR4 agonist lipopolysaccharide (LPS, 100ng/ml) and TLR7/8 agonist Resiquimod (R848, 2.5 µg/ml) for 12 h (activation conditions are summarized in **Supplementary Table 2**). After stimulation, cells were washed with PBS and shock frozen until further use. For secretome measurements, cells were cultured as described above, with the adjustment of exchanging the complete medium after 2 h of priming with medium lacking serum and phenol red. Our data show that with this strategy we also capture proteins that are released early (such as many chemokines or TNF) while maintaining optimal cell viability. Cells were pelleted at 3,000 rpm for 5 min. Supernatants were removed carefully while leaving ¼ culture medium surplus to reduce cell contamination. Cell debris were removed by an additional centrifugation step at full speed for 10 min. Supernatants were shock frozen and stored at −80 °C until further use.

Cytokine analysis. Secretion of resistin (RETN) was measured in culture supernatants using the human resistin ELISA kit from Sigma Aldrich.

CD4 T cell activation with IL-34. Total CD4 memory T cells were flow sorted ($CD4^+CD45RA^+$) and activated as described above with the alteration that IL-34 was added to the culture medium (1, 10, 100 ng/ml). After 96 h cells were washed twice with PBS, shock frozen and subjected to total proteome analysis.

Sample preparation for MS analysis. Cell pellets and supernatants were lysed in 8 M urea and 2 M urea, respectively, 10 mM HEPES (pH 8), 10 mM DTT and sonicated at 4 °C for 15 min (level 5, Bioruptor, Diagenode). Alkylation of reduced cysteines was performed in the dark for 30 min with 55 mM iodacetamide (IAA) followed by a two-step proteolytic digestion. Samples were digested at 21–24 °C with LysC (1:50, w/w, Wako) for 3 h. Cell lysates were adjusted to 2 M Urea with 50 mM ammoniumbicarbonate and then both cell lysates and supernatants were digested with trypsin (1:50, w/w, Promega) at 21–24 °C overnight. The resulting peptide mixtures were acidified and loaded on C18 StageTips (EmporeTM, IVA-Analysentechnik). Peptides were eluted with 80% acetonitrile (ACN), dried using a SpeedVac centrifuge, and resuspended in 2% ACN, 0.1% trifluoroacetic acid (TFA), and 0.5% acetic acid. For MACS enriched cell pellets lysis was performed in 4% SDS, 10 mM HEPES (pH 8, Biomol), 10 mM DTT. Cells were heated at 95 °C for 10 min and sonicated at 4 °C for 15 min (level 5, Bioruptor, Diagenode). Proteins were precipitated with acetone at −20 °C overnight and resuspended the next day in 8 M urea, 10 mM Hepes (pH 8). Proteolytic digestion was carried out as described above. Chemicals were purchased from Sigma-Aldrich unless stated otherwise.

LC-MS/MS. Peptides were separated on an EASY-nLC 1000 HPLC system (Thermo Fisher Scientific) coupled online to the Q Exactive HF mass spectrometer via a nanoelectrospray source (Thermo Fisher Scientific)[20]. Peptides were loaded in buffer A (0.5% formic acid) on in house packed columns (75-µm inner diameter, 50 cm length, and 1.9 µm C18 particles from Dr. Maisch GmbH, Germany). Peptides were eluted with a nonlinear 180-min gradient of 5–60% buffer B (80% ACN, 0.5% formic acid) at a flow rate of 250 nl/min and a column temperature of 55 °C. Operational parameters were real-time monitored by the SprayQC software[49]. The Q Exactive HF was operated in a data dependent mode with a survey scan range of 300–1,650 m/z and a resolution of 60,000 to 120,000 at m/z 200. Up to the ten most abundant isotope patterns with a charge >1 were isolated with a 1.8 Thomson (Th) isolation window and subjected to higher-energy collisional dissociation (HCD) fragmentation at a normalized collision energy of 26. Fragmentation spectra were acquired with a resolution of 15,000 at m/z 200. Dynamic exclusion of sequenced peptides was set to 30 s to reduce repeated peptide sequencing. Thresholds for ion injection time and ion target values were set to 20 ms and 3E6 for the survey scans and 55 ms and 1E5 for the MS/MS scans, respectively. Data were acquired using the Xcalibur software (Thermo Scientific).

LC-MS/MS data analysis. MaxQuant software (version 1.5.3.2) was used to analyze MS raw files[21]. MS/MS spectra were searched against the human Uniprot FASTA database (Version May 2013, 90,507 entries) and a common contaminants database (247 entries) by the Andromeda search engine[50]. Cysteine carbamidomethylation was applied as fixed and N-terminal acetylation, deamidation at NQ, and methionine oxidation as variable modifications. Enzyme specificity was set to trypsin with a maximum of 2 missed cleavages and a minimum peptide length of 7 amino acids. A false discovery rate (FDR) of 1% was applied at the peptide and protein level. Peptide identification was performed with an allowed initial precursor mass deviation of up to 7 ppm and an allowed fragment mass deviation of 20 ppm. Nonlinear retention time alignment of all measured samples was performed in MaxQuant. Peptide identifications were matched across all samples within a time window of 1 min of the aligned retention times[51]. A library for 'match between runs' in MaxQuant was built from duplicate and additional single shot MS runs from MACS enriched cell types. Protein identification required at least 1 'razor peptide' in MaxQuant. A minimum ratio count of 1 was required for valid quantification events via MaxQuant's Label Free Quantification algorithm (MaxLFQ)[22]. We enabled FastLFQ with a minimum of three and an average of six neighbors. Data were filtered for common contaminants and peptides only identified by side modification were excluded from further analysis. We observed lower identification rates for plasma blasts and neutrophils compared to all other cell types. In the case of plasma blasts the lower identification rate derives from the low cell count. Neutrophil identification rates are below average because 1) neutrophils contain many proteases and as a result peptides are cleaved at different positions than arginine or lysine (the expected residues from tryptic digestions) and 2) have high abundant proteins that result in unusually broad peptide peak widths, interfering with the detection of other ions.

Transcription profiling. Total RNA was isolated from flow sorted cell types (as described in Blood samples and cell sorting) using the RNeasy Plus Mini Kit from Qiagen. RNA samples from four donors were pooled. RNA was analyzed using an Agilent 2100 Bioanalyzer system (Agilent Biotechnologies). Only samples with RIN > 8 were subjected to sequencing at the Genomics Core Facility (GeneCore), Heidelberg using an Illumina Hiseq2000 sequencer. Processed reads were mapped to the human genome (GRCh37) using the

doi:10.1038/ni.3693

PERSEUS software (v. 1.5.5.5). RPKM values were calculated by normalizing to the maximum gene length. A cutoff value of 1 RPKM was used as a limit for detection.

Data filtering, imputation of missing values and copy number estimation. Data analysis and visualization was performed using the Perseus software[52] and the R statistical framework[53,54]. Before imputing missing values, protein identifications were required to have more than 50% valid values in at least one group of replicates. The remaining missing values were imputed by a normal distribution with a s.d. of 30% in comparison to the deviation of the measured values and a down-shift of the mean by 2.2 s.d. to simulate the distribution of low abundant proteins. Copy numbers were estimated using the proteomic ruler approach by using the fixed ratio between the total histone signal and cellular DNA mass[28].

Comparison to proteome and microarray data. Microarray data were retrieved from http://www.broadinstitute.org/dmap/home[8]. We employed the R Bioconductor package 'mygene' to match Uniprot identifier to the provided Entrez identifier[55]. Raw files from[23] were processed with MaxQuant applying identical cut offs for protein identifications as stated above. Only MACS enriched immune cell protein identifications were used for comparison. The total number of identifications and coverage of immune system relevant annotations was determined by matching all three data sets to a human reference genome (20,591 entries, Perseus v1.5.2.12) based on Uniprot identifiers (**Fig. 1c**,**d**).

Lineage-specific protein signatures and cell-type marker selection. Proteomes were compared by three different groupings: (1) pairwise, (2) cell lineage (T4, T8, NK, B, MO, DC, and GN) versus rest, (3) and individual cell types versus all other proteomes. Proteomes of activated cell types were excluded from cell lineage comparisons and platelets and erythrocytes from all comparisons. Proteins that significantly differed in abundance were identified by a parametric two-tailed Welch's *t*-test with a permutation-based false discovery rate (FDR) of 5% and a S_0 parameter of 1 (ref. 56), if not stated otherwise (**Fig. 2b**). In addition, we employed Lasso regression analysis[29] to identify cell-type-specific markers for cell sorting (**Supplementary Fig. 3d**). The pool of proteins was reduced to cell surface receptors.

Assignment of functional modules. Proteins were clustered to functional modules using weighted gene co-expression network analysis (WGCNA)[31]. The proteome data was reduced to proteomes of cell types in their steady state and selected for proteins with a significantly different abundance pattern between cell types (ANOVA, Benjamini-Hochberg (BH) FDR $P < 5 \times 10^{-5}$, df1 = 25, df2 = 78). Standard parameters were changed to a power of 14, 'signed' network, average clustering, and a minimum module size of 20. The algorithm assigned the 6,982 proteins to 47 modules containing 23–725 proteins.

Annotation enrichment analysis. Protein modules, cell type signatures as well as selected principal components were functionally characterized by annotation enrichment analysis. For all three cases we used proteins annotations form the Gene Ontology (GO)[57], KEGG[58], and Uniprot Keywords[59] databases. For the first two cases, enrichment scores were determined using Fisher exact test and in the latter 1D annotation enrichment analysis was performed[25]. Both tests we corrected for multiple hypotheses using a Benjamini-Hochberg false discovery rate of 5%, if not stated otherwise.

Categorization of intracellular signaling levels. Based on protein annotations combinations, we defined four levels of signal transduction: transcription factor (T), adaptor molecule (A), receptor (R) and secreted molecules/ligands (L). 'T' were characterized by being localized in the nucleus (GOCC: 'nucleus', 'nuclear part'), DNA binders (GOMF: 'DNA binding') and transcription regulators (GOMF: 'transcription regulatory region DNA binding', 'sequence-specific DNA binding transcription factor activity'), but at the same time were not involved in DNA repair (GOMF: 'damaged DNA binding'; GOBP 'DNA repair'). 'A' were required to interact with at least one receptor (StringDB score > 0.4). In addition, they had to be localized in the cytoplasm (GOCC: ('cytoplasm', 'cytosol', 'cytoplasmic part') and involved in cell signaling (GOBP: 'signaling', 'signal transduction'). 'R' were defined solely by UniProt sequence features. Proteins were required to have either an' extracellular' or 'GPI-anchor' topology domain. 'L' were required to be secreted (Keywords: 'Secreted', 'Signal') or interact with a receptor (StringDB score > 0.4) and be localized in the extracellular region (GOCC: 'extracellular space', 'extracellular region'). Furthermore, for 'R' such as TNF, we define a specific category. 'R' with cytokine activity (GOMF: 'cytokine activity', 'chemokine activity', 'growth factor activity', 'hormone activity'; UniProt Keywords: 'Cytokine', 'Hormone', 'Growthfactor') were assigned 'L'.

Intercellular interaction network. To define intercellular connections we retrieved protein interaction data from the String database[40]. From the different score levels available we selected only interactions with experimental evidence and a score greater than 0.4. In addition, we established a twofold ranking for pairwise intercellular protein interactions using our proteome data. First, individual protein expressions were divided by their maximum expression level in order to weigh protein expression across all cell types independent of their abundance. Second, we computed a significance score by counting how often a protein was significantly (FDR 5%) more abundant in one cell type compared to all other cell types and normalized it by the maximum count. The product of normalized expression and normalized significance was used to compute connection ranks for all proteins. The final intercellular connection score between protein pairs was calculated as the mean of the individual connection ranks of the two proteins. We excluded self-loops as well as L-L interactions. To avoid intercellular connection between proteins of the same receptor complex, we annotated receptor complexes using GOCC terms and the protein complex database PCDq[60] and merged proteins of the same complex to one entity. This resulted in a network of about 4×10^5 possible receptor-ligand or receptor-receptor connection between the 26 different cell types in their steady and activated states. We evaluated our intercellular connection scores by comparing their distribution to a manually annotated immune cell-cytokine connection network 'Textbook' (http://www.immunexpresso.org/, see below). We observed that with increasing intercellular connection score the number of connections covered by the textbook fits a sigmoid curve and set intercellular interaction score cutoff at the inflection point (0.47) (**Supplementary Fig. 5b**,**c**). Together, this resulted in a high confidence immune cell communication network comprising approximately 180,000 interactions, containing roughly 80,000 R-R and 100,000 R-L interactions. The networks were visualized with the R package 'circlize'[61]. Furthermore, we extend our interaction network by applying the described interaction framework features to body tissues[42]. For the transcriptomes, we applied the same scoring but restricted the network to receptor-receptor interactions.

Transcriptome-proteome correlation and module analysis. Data sets were filtered for cell types present in both data sets (T4.naive - TCELLA6, T4.CM - TCELLA8, T4.EM - TCELLA7, T8.naive - TCELLA2, T8.CM - TCELLA4, T8.EM - TCELLA3, T8.EMRA - TCELLA1, NK.bright - NKA2, NK.dim - NKA3, B.memory - BCELLA2, B.naive - BCELLA1, MO.classical - MONO2, mDC - DENDA2, pDC - DENDA1, Neutrophil - GRAN3, Eosinophil - EOS2, Basophil - BASO1). We calculated Pearson correlation coefficients for protein profiles across all matched cell types. WGCNA was performed with ANOVA significant proteins (Benjamini-Hochberg (BH) FDR $P < 5 \times 10^{-5}$). Standard parameters were changed to a power of 14, 'signed' network, average clustering, and a minimum module size of 20. Highly correlating modules were merged at a cutHeight of 0.2 or 0.3 for the microarray and proteome data set, respectively. This assigned the 5,782 genes of the microarray data set to 18 modules and 5,974 proteins of the proteome data set to also 18 modules. Module similarity was determined by two measures: Pearson correlation of the module Eigengenes and relative gene overlap between the modules. Intercellular communication network for both data set were constructed as described above but restricting the networks to shared cell types only.

Comparison to Textbook and ImmuneXpresso. To identify novel connections between immune cell pairs, we compared our intercellular interaction network to a comprehensive literature-based interaction network (http://www.immunexpresso.org/). This literature-based network consists of two data sets 'Textbook' (1) and 'ImmuneXpresso' (2). The 'Textbook' contains semi-manual annotated cell-cytokine or cytokine-cell interactions retrieved from[41].

'ImmuneXpresso' has an identical data structure, but the interactions are automatically extracted from PubMed abstracts by a Natural Language Processing engine. For comparisons, the data sets were matched by gene names and cell types. In case cell types were not identical, the parent cell types were matched.

Intracellular interaction network. We used the bioinformatics framework described above in 'Intercellular interaction network' to construct intracellular protein signaling networks for each immune cell type. To this end, we assigned to each receptor the closest adaptors (path length = 1, degree > 10) and the closest transcription factors (path length < 3, degree > 5). We grouped receptors according to their downstream signaling molecules to identify at which layer of signal transduction input signals from ligands present in secretomes converge (**Supplementary Fig. 7c**,**f**).

Secretome analysis. iBAQ intensity values[12] were median normalized and missing values imputed as described in 'Data filtering, imputation of missing values and copy number estimation'. Significant secreted proteins were identified by two-tailed Welch's *t* test comparing stimulated with unstimulated cells with a permutation-based false discovery rate (FDR) of 1% and a S_0 parameter of 1 for monocytes (**Supplementary Fig. 7d**,**e**). For mDC the FDR was set to 5% and for pDC to 15% (**Supplementary Fig. 7a**,**b**). To compare the context depended activations, significant secreted proteins were ranked to account for abundance biases due to different agonists or cell types. For example, TNF is one of the most abundant proteins secreted from monocytes after ZYM and LPS stimulation. However, ZYM treated monocytes secrete significantly more TNF than LPS treated monocytes. However, this abundance difference might be minor at different time points or LPS concentrations. Therefore, secretome results are presented as rank-based secretion (**Fig. 7a**,**b**,**e**,**f**)

Data availability. All data generated during this study are included in this published article and deposited online (ProteomeXchange Consortium, data set identifier PXD004352). Detailed descriptions for all submitted data tables can be found in **Supplementary Note 1**. The data is publicly accessible under http://www.immprot.org or MaxQB (http://maxqb.biochem.mpg.de/mxdb/). Additional published data include, microarray data of immune cells (http://www.broadinstitute.org/dmap/home), transcriptome data of tissues and organs (10.1126/science.1260419), protein interaction data (http://string-db.org/, version 10.0), and protein complex annotations (http://h-invitational.jp/hinv/pcdq/). Raw data and the MaxQuant output tables have been deposited to the ProteomeXchange Consortium via the PRIDE[62] partner repository with the data set identifier PXD004352 (https://www.ebi.ac.uk/pride/archive/projects/PXD004352). The data is publicly accessible under http://www.immprot.org/. The data that support the findings of this study are either available in a public repository, attached as supplementary tables or can be requested from the corresponding authors.

47. Brosseron, F. *et al.* Stepwise isolation of human peripheral erythrocytes, T lymphocytes, and monocytes for blood cell proteomics. *Proteomics Clin. Appl.* **6**, 497–501 (2012).
48. Trichler, S.A., Bulla, S.C., Thomason, J., Lunsford, K.V. & Bulla, C. Ultra-pure platelet isolation from canine whole blood. *BMC Vet. Res.* **9**, 144 (2013).
49. Scheltema, R.A. & Mann, M. SprayQc: a real-time LC-MS/MS quality monitoring system to maximize uptime using off the shelf components. *J. Proteome Res.* **11**, 3458–3466 (2012).
50. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
51. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
52. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
53. RStudio, I. *Shiny: Easy web applications in R* (2014).
54. Team, R.D.C. *R: a Language and Environment for Statistical Computing* (2008).
55. Gentleman, R.C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
56. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121 (2001).
57. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
58. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).
59. Magrane, M. & UniProt, C. UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* **2011**, bar009 (2011).
60. Kikugawa, S. *et al.* PCDq: human protein complex database with quality index which summarizes different levels of evidences of protein complexes predicted from h-invitational protein-protein interactions integrative dataset. *BMC Syst. Biol.* **6** (Suppl, 2), S7 (2012).
61. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
62. Vizcaino, J.A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41**, D1063–D1069 (2013).

doi:10.1038/ni.3693