Organization of the autoantibody repertoire in healthy newborns and adults revealed by system level informatics of antigen microarray data

Asaf Madi[a,1], Inbal Hecht[b,c,1,2], Sharron Bransburg-Zabary[a,b], Yifat Merbl[d], Adi Pick[b], Merav Zucker-Toledano[a,e], Francisco J. Quintana[d], Alfred I. Tauber[f], Irun R. Cohen[d,2], and Eshel Ben-Jacob[b,c,2]

[a]The Sackler School of Medicine and [b]The Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel; [c]The Center for Theoretical Biological Physics, University of California San Diego, La Jolla, CA 92093; [d]Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel; [e]Pediatric Department, Dana Children's Hospital, Tel-Aviv Sourasky Medical Center, Tel-Aviv 64239, Israel; and [f]Department of Medicine, School of Medicine, Boston University, Boston, MA 02118

The immune system is essential to body defense and maintenance. Specific antibodies to foreign invaders function in body defense, and it has been suggested that autoantibodies binding to self molecules are important in body maintenance. Recently, the autoantibody repertoires in the bloods of healthy mothers and their newborns were studied using an antigen microarray containing hundreds of self molecules. It was found that the mothers expressed diverse repertoires for both IgG and IgM autoantibodies. Each newborn shares with its mother a similar repertoire of IgG antibodies, which cross the placental but its IgM repertoire is more similar to those of other newborns. Here, we took a system-level approach and analyzed the correlations between autoantibody reactivities of the previous data and extended the study to new data from newborns at birth and a week later, and from healthy young women. For the young women, we found modular organization of both IgG and IgM isotypes into antigen cliques—subgroups of highly correlated antigen reactivities. In contrast, the newborns were found to share a universal congenital IgM profile with no modular organization. Moreover, the IgG autoantibodies of the newborns manifested buds of the mothers' antigen cliques, but they were noticeably less structured. These findings suggest that the natural autoantibody repertoire of humans shows relatively little organization at birth, but, by young adulthood, it becomes sorted out into a modular organization of subgroups (cliques) of correlated antigens. These features revealed by antigen microarrays can be used to define personal states of autoantibody organizational motifs.

immune development | immune holography | immune network | network cliques

The immune system is a complex network of molecules, cells and organs that act together to maintain and repair the body and protect it against foreign invaders (1–8). These tasks involve the sensing and processing of foreign antigens together with internal signals that disclose the state of the body leading to immune responses, learning and memory. The repertoire of antibodies reflects the past history of immune responses to foreign antigens and prepares the immune system for efficient future responses to pathogens bearing these antigens. Antibodies binding to molecules of the body itself—autoantibodies—are associated with the pathologic inflammatory processes that cause autoimmune diseases (1, 2). However, autoantibodies in healthy individuals, in contrast to pathogenic autoantibodies, are thought to function in body maintenance and healing. Indeed, different individuals in similar physiological states share some common features of their autoimmune repertoires—a feature that has been termed the immunological homunculus (2, 3). Therefore, it has been proposed that the global pattern of autoantibodies can reveal various network states of the immune

system and provide some insights about the state of the individual (9, 10).

A new technology, the antigen microarray chip, enables simultaneous measurements of the reactivity of hundreds of antibodies (9, 11–13)(see also section I in *SI Text*). An immune microarray consists of various antigens covalently linked to the surface of a glass slide. A drop of blood serum (or other body fluid) is tested for antibody reactivities by measuring antibody binding to each antigen spot using fluorescence labeling. Note that the binding of antibodies to a spotted antigen cannot tell us about the stimulus that induced the antibodies, and it cannot define the affinity or the specificity of any particular antibody or collective of antibodies. Indeed, a positive antigen-binding signal probably reflects a polyclonal mixture of antibodies binding to a variety of structural epitopes exposed by each spotted antigen.

The parallel measurement of reactivities to hundreds of different antigens poses a computational challenge: how to extract relevant and meaningful information from a vast amount of data. Computationally, this problem is similar to that encountered in the parallel measurements of the expression of hundreds to thousands of genes by gene microarrays, for which many advanced computational methods have been developed. Generally speaking, these methods belong to two main classes: (*i*) Supervised methods that are aimed to identify classifier genes—genes that enable the distinction between known groups of subjects; (*ii*) Unsupervised methods that are aimed to identify the existence of significantly distinct groups of subjects and the classifier genes. The latter are usually based on cluster analysis. Thus, supervised (14–17) and unsupervised methods (18–29) developed for gene-expression studies can be adopted for the analysis of antigen chip data (9, 11, 13, 30).

Here, we studied the antibody reactivity data of 10 pairs of mothers and newborns previously reported by Merbl et al. (30), to which we added additional data from 5 females, neither pregnant nor post-labor, and from 8 newborns for whom we had samples at birth and at day 7. In contrast to Merbl et al., the present analysis was carried out from the perspective of the antibody repertoire as a complex functional network of immu-

Fig. 1. The ordered (by the dendrogram algorithm) antibody reactivity matrices. The results show the reactivities of 305 antigens of the IgM and IgG isotypes for samples taken from the 10 mother-cord pairs. The colors (green low to red high) indicate the reactivity levels. The similarity distances calculated by the dendrogram clustering algorithm for the subjects and antibodies are indicated by the blue trees in the *x* and *y* directions, respectively. (*A*) The ordered reactivity matrix of the IgM isotype. Note that in this matrix the cords (subjects 1–10) and the mothers (subjects 11–20) form two groups. The dendrogram distance trees show high similarity between the cords and lower (longer dendrogram distances) between the mothers. (*B*) Dendrogram reactivity matrix of the IgG isotypes. Note that each mother-cord pair forms a separate group.

noglobulins. Inspired by the ideas of Jerne (31–33), we adopted a correlation-based system-level approach to extract hidden information about functional relations between autoantibody-antigen reactivities.

More specifically, we computed the matrices of *antibody correlations* and the matrices of *subject correlations* rather than the reactivity matrices as is usually done. The correlation matrices were analyzed using the functional holography (FH) method of Baruchi et al. (34), originally devised for analyzing recorded brain activity. This method was recently shown to be useful in the analysis of gene-expression data: it can help identify operons, relations between genes within operons (gene order, gene separation by noncoding segments and start and end regions), and the functional relations between operons (35). The capacity of the FH method to successfully reveal such motifs between genes led us to test whether this method can also identify new functional relations and network motifs in natural antibody repertoires.

Reactivity Matrices

The reactivity matrices of the 10 mother-cord pairs are shown in Fig. 1*A* for IgM and Fig. 1*B* for IgG. The rows in these matrices are the antibody reactivity profiles (reactivities of one specific antigen for the different subjects) and the columns represent the immune profiles of each subject. The matrices are reordered using a dendrogram clustering algorithm (36, 37).

The dendrogrammed reactivity matrix for the IgM isotype in Fig. 1*A* shows that the cords (subjects 1–10) and the mothers (subjects 11–20) form two distinct groups, as was reported earlier (30). The separation into two distinct groups results from the fact that the average level of antibody reactivities in the mothers' samples is significantly higher than that in the cord samples as can be clearly seen in Fig. 1*A*. In the next section we show, by analyzing the subject-normalized correlations, that the mothers have very different immune profiles from one another and from the cords, but all newborns share a similar profile.

The dendrogrammed matrix of normalized IgG reactivities (Fig. 1*B*) shows that each mother-cord pair [(1,11), (2,12) . . . (*i*,10 + *i*) . . . (10,20)] forms a unique subgroup. This pairing is expected because it is known that IgG antibodies are transferred from the mother to her fetus through the placenta. In the next section we show that the IgG immune profile of each mother is

very similar to that of her newborn and very different from the other mothers.

Immune State Features Revealed in the Subject Correlation Matrices

The correlation matrices, both between the antibody reactivity profiles and the subject immune profiles, were calculated by Pearson's formula (38):

$$C(i, j) = \frac{\langle (X_i(n) - \mu_i)(X_j(n) - \mu_j) \rangle_n}{\sigma_i \sigma_j} \qquad [1]$$

More specifically, in the subject correlation matrix (of size 20*20), the elements $C(i,j)$ represent the correlations between the columns X_i and X_j for subjects (i) and (j) in the reactivity matrices. Similarly, in the antigen correlation matrices (of size 305*305), the elements $C(i,j)$ represent the correlations between the rows in the reactivity matrices. Following Baruchi et al. (36), we also performed collective normalization of the correlations using the metacorrelations $MC(i,j)$—the Pearson correlation between rows (i) and (j) of the correlation matrix after reordering. In the reordering process, the elements $C(i,i)$ and $C(j,j)$ are removed from the calculation. The correlation vector for (i) is $\{C(i,j), C(i,1), C(i,2),\ldots\}$, and for ($j$) it is $\{C(j,i), C(j,1), C(j,2),\ldots\}$. In other words, the metacorrelation is a measure of the similarity between the correlations of antigen (i) with all other antigens and the correlations of antigen (j) with all other antigens. Using the metacorrelations coefficient, the normalized correlation matrix $A(i,j)$ is given by:

$$A(i, j) = C(i, j) \cdot MC(i, j) \qquad [2]$$

In this section, we discuss the subject correlation matrices shown in Fig. 2. The antigen correlation matrices are presented in the *SI Text* and discussed further below.

The IgM correlation matrix reveals the separation of the mothers and cords into two distinct clusters (Fig. 2*A*). The formation of the two distinct groups is related to the fact that the immune profiles of all of the cords are very similar (as the cord-cord correlations are >0.9), whereas the similarities between the mothers' profiles are low, and the similarities between the mothers' and cords' profiles are lower. The aforementioned information becomes clearer after the collective normalization, as is evident in the normalized subject correlation matrix shown in Fig. 2*C*.

The high similarities between the IgG profiles of each cord and its mother are quite transparent when the corresponding correlation matrix is inspected (Fig. 2*B*): The high correlations between the immune profiles of the cord-mother pairs are reflected by the two red subdiagonals. The collective normalization further signifies features and can reveal additional information, as seen in the normalized correlation matrix shown in Fig. 2*D*. For example, the specific mother-cord pair (4, 14) stands out as having lower similarity compared with the other pairs.

The Congenital IgM Immune Profile Is Shared; the Adult IgM Profiles Are Individual

To identify hidden features associated with immune network development, we performed dimension reduction of the correlation matrices employing the PCA algorithm. This (39) and other dimension-reduction algorithms (40, 41) made it possible to extract a large amount of relevant information embedded in the matrices (see SI), and yet present it in a graphical way, which is naturally limited in the number of dimensions. We found that reduction to three dimensions extracts most of the relevant information (>85% of the information as described in section II of *SI Text* and in Fig. S1), similar to what was found for brain activity data (34) and gene expression data (35). Therefore, we present the immune response information in a 3-dimensional PCA space, whose axes are the three leading principal vectors
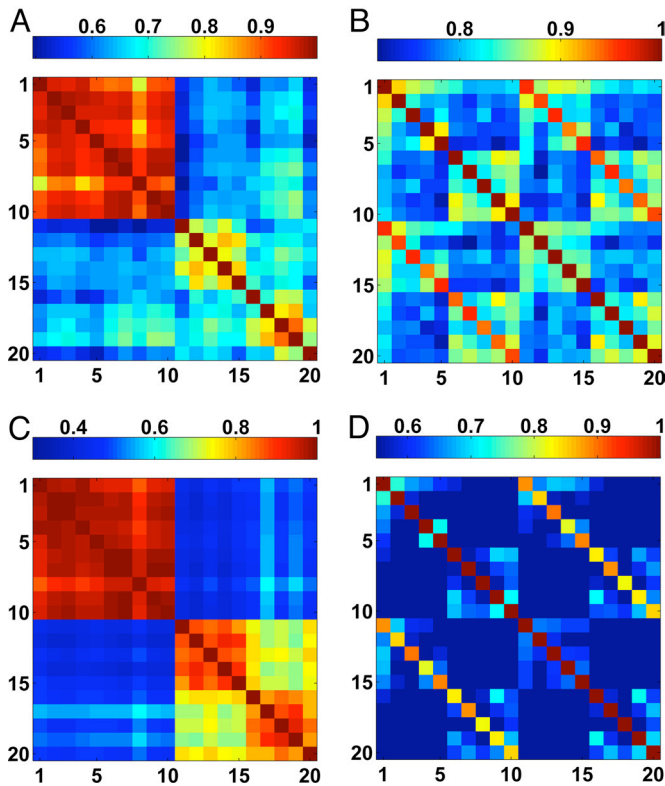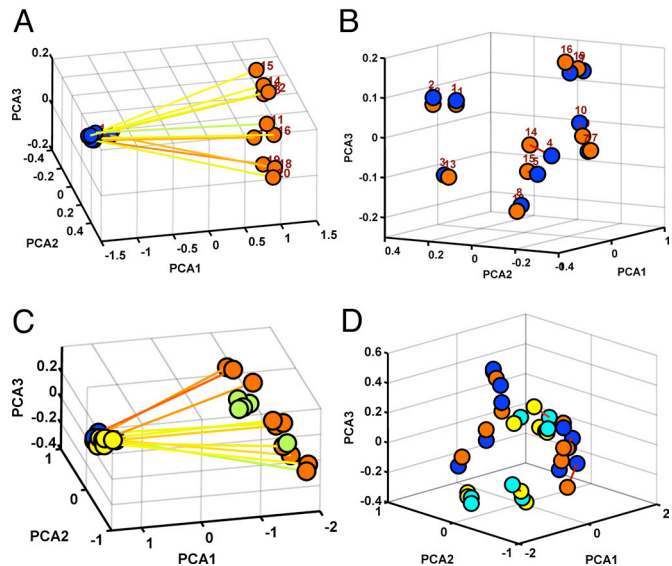
Fig. 2. The subject correlation matrices in color code representation. (*A* and *B*) The subject correlation matrices for the IgM and IgG isotypes, respectively. (*C* and *D*) The corresponding normalized matrices using the method described in the text. The matrices are ordered according to the cords (1–10) and the mothers (11–20). We note that the seeming separation of the mothers into two subgroups (the two subclusters in the mothers cluster in Fig. 2*C*, for mothers 1–5 and 6–10) appear to the eye more substantial because of the color code; the cross correlations between the two subgroups are ≈0.7 and they reflect experimental variations in storage and between microarrays which do not affect the results of the analysis.

Fig. 3. Holographic representation of the immune states for the IgM isotype (*Left*) and IgG isotype (*Right*). (*A*) The subject holographic network in the PCA space for IgM antibodies of the 10 pairs of mothers (red) and their cord samples (blue). Each mother-cord pair is linked with color coded line (blue low to red high) according to the values of the correlation between them (Fig. 2*A*). (*B*) The same as *A* for the IgG isotype. (*C*) The IgM holographic network including the 10 mother-cord pairs, the additional 5 females (green) and the additional 8 newborns at day 1 (yellow). (*D*) The IgG network including the 10 mother-cord pairs and the 8 newborns at day 1 (yellow) and day 7 (cyan).

computed by the PCA algorithm. Each subject is placed in this space according to its three eigenvalues for the three leading principal vectors. Note that subjects that manifest high normalized correlations will be placed in close vicinity in the PCA space.

Some relevant information can be lost in the dimension reduction process. To retrieve this information, as well as information that might be lost in the collective normalization process, we link each pair of subjects by lines, color coded according to the original (nonnormalized) correlations (34, 35, 42). The results (Fig. 3; see also Movies S1–S4 for more viewing angles and section VIII in *SI Text* for description) are subject networks, or manifolds, of linked nodes in the PCA space that provide a holographic presentation of the functional relations between the immune states of the different subjects. We termed this graphical presentation Immune Holography (IH), to signify its ability to capture the characteristics of the system as a whole in a clear, graphical way.

Fig. 3 *A* and *B* shows the subject networks for the IgM and IgG isotypes, respectively. The IgM subject network of mothers and cords (Fig. 3*A*) shows that the cords (blue nodes) and the mothers (red nodes) form two distinct subclusters. These two clusters have very different geometries: The cord manifold is very compact, indicating that the cords share a very similar IgM profile. In contrast, the maternal manifold is widely spread, indicating that each mother has her own individual IgM profile that is significantly different from the others. A more quantitative analysis can be found in section III of *SI Text*

and in Table S1. More specifically, we used the Tensors of Gyration (45) to quantify the differences between the manifolds, and found significant differences in the structures of the clusters of the mothers versus the cords, as described in detail in section III in the *SI Text*. The compact cord cluster of IgM reactivities suggests the existence of a universal congenital IgM immune profile. The spread manifolds for the mothers imply that this congenital profile develops through the personal life experience of each individual, leading to the development of diverse individual immune profiles during maturation (43). The mature profiles are substantially different from the congenital profile and from each other.

The IgG immune manifold of mothers and cords (Fig. 3*B*) is composed of 10 subclusters, one for each mother-cord pair. This organization indicates the high similarity between the IgG profiles of each newborn and his/her mother. This behavior is in accordance with the fact that >90% of the mother's IgG repertoire can be found in the cord blood (44), as maternal IgG antibodies cross the placenta and enter the fetal blood circulation. The paired clusters are widely spread, indicating that each mother and her newborn manifest a unique IgG immune state. These findings imply that, unlike the universally shared congenital IgM immune state, the IgG immune state at birth is individual and is largely set by the maternal IgG state.

Fig. 3*C* includes data of the combined dataset of the IgM antibodies of the combined dataset of the 10 mother-cord pairs, an additional set of 5 females and additional 8 newborns; the analysis reveals a compact manifold that includes all of the newborns and a dispersed manifold that includes the adult female subjects. These findings further support the existence of a universal congenital immune profile and show that each individual generates his/her own immune profile during immune development.

Analyzing the IgG antibodies of the extended dataset (Fig. 3*D*), we found that each of the newborns forms an individual cluster of autoantibody reactivities at days 1 and 7 and that the individual clusters are spread in the PCA space as are the

IMMUNOLOGY

mother-cord pairs. The fact that each of the newborns forms a personal cluster of IgG autoantibodies at days 1 and 7 indicates that the IgG autoimmune state is quite stable over the first week of life, despite colonization of the gut, upper respiratory tract and skin with normal flora. However, closer inspection of the autoantibody profiles between day 1 and 7 (section VI in *SI Text* and Fig. S8) captures some weak changes in the immune profiles of some of the newborns between day 1 and day 7 that might be associated with buds of early immune development during the first week.

Antigen Cliques—Modular Organization of Mature Immune Networks

Further analysis disclosed the formation of subgroups, or cliques, of antibodies with highly correlated reactivities in the adult autoantibody repertoires. We discovered such subgroups of antigen reactivities by analyzing the mothers' IgM and IgG data using the approach described in detail in section IV of *SI Text*, which is based on the search for gene operons by analyzing correlations in gene expression data (ref. 35; see also section V of *SI Text*, Figs. S6 and S7, and Table S3). It should be noted that the term clique is used here similarly to its use in the study of network modules in network theory: A clique is a group whose members associate regularly with each other on the basis of common processes and functions. This notion differs from that used in graph theory, where a clique refers to a graph in which every vertex is connected to every other vertex in the graph. We chose to use this term taking the perspective of the immune system as a dynamic network.

Our method to identify immune cliques involves four stages (for more detail see also Figs. S2–S5). First was standard-deviation (STD) filtering. Because we look for groups of antigens with higher intergroup correlations, the STD of their correlations is expected to be relatively high. So we compute the distribution of the correlation STD and select the antigen reactivities at the higher part of the distribution. Second was dendrogram selection. We apply a dendrogram clustering algorithm on the matrix of normalized correlations of the antibodies selected in stage 1. This algorithm computes the correlation distances between all of the antibodies, and we select those with correlation distances below a certain threshold. Third was identification of additional members of the clique. We calculate a representative immune profile for each of the identified immune cliques, which yields the averaged reactivity of the group. Then we search for additional antigen reactivities that manifest high correlations with the representative profiles. Fourth was identifying weaker cliques. Once the first set of antibody groups (the ones that have the strongest correlations) are found, we repeat the search process on the matrix of all antigen reactivities that were not selected at the first round.

The strong autoantibody groups we determined for the maternal IgM and IgG reactivities are shown in Fig. 4. In Fig. 4 *A* and *B*, we show the normalized correlation matrices for 45 and 57 antigens that compose the strong cliques for the IgM (4 groups) and IgG (3 groups) isotypes, respectively. The organization of these subgroups in the corresponding IH space is shown in Fig. 4 *C* and *D*. As can be seen from these figures, the subgroups of antigen reactivities form well separated clusters in IH space indicating high statistical significance (see also sections III and IV in *SI Text* and Table S2).

Fig. 5 visually illustrates the meaning of an antibody clique: The antibodies that belong to a specific clique show similar antigen reactivities for each specific individual although they differ between individuals. Therefore, for each individual we can assign a representative reactivity of the clique which is the average of the reactivities of the clique antigen reactivities for this individual, as is shown by the boxes in Fig. 5: for each clique, the representative antigen reactivity varies from individual to individual thus repre-
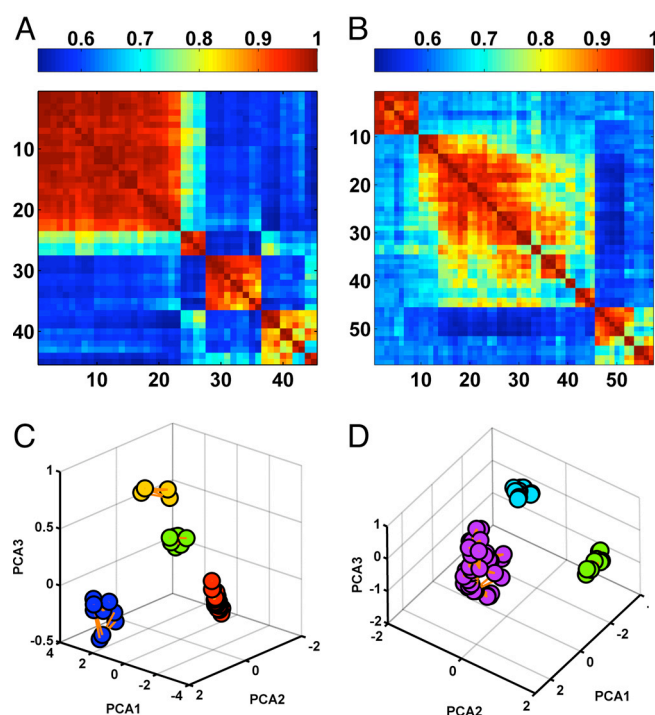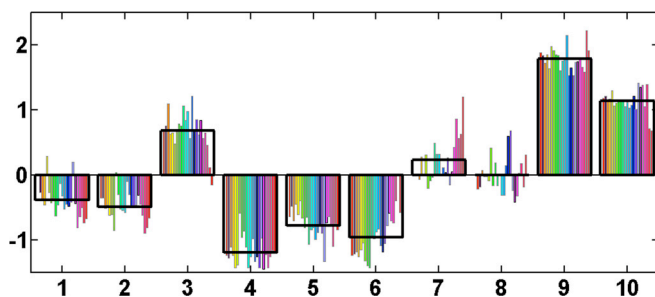

Fig. 4. The antigen cliques. (*A*) The normalized antigen correlation matrix of the IgM isotype of 45 antigens that compose the strong cliques in the maternal dataset. (*B*) Similar to *A* but for the IgG data. Note that in this case 57 antigens were selected and the IgM and IgG antigen clique members do not have to be similar. (*C* and *D*) The corresponding antigen networks. Note that 4 and 3 subgroups were identified for the IgM and IgG isotypes, respectively. In the presentation of the antigen networks, nodes (antigens) with high correlations (>0.85) are linked with color coded lines according to the correlation levels.

senting the response profile of the clique. Similarly, a specific individual manifests a different representative reactivity for the different cliques, which might reflect his or her individual immune state organization. We emphasize that for each individual, a different set of polyclonal antibodies can bind to the same given antigen. Hence each individual can express an individualized repertoire of autoantibodies that can correspond to a specific clique. For the effect of the additional females data on the identified cliques, see section VII in *SI Text* and Figs. S9 and S10.

Development of the Immune Modular Organization from Birth to Adulthood

Searching for antigen cliques in cord IgM reactivities, we found that about half of the autoantibodies belong to one large

Fig. 5. The reactivity profile of the antigens of a specific clique. The normalized reactivity values (relative to the average and divided by the STD) for the 10 mothers (marked 1–10) of the 22 antigens that belong to a specific IgM clique (the green cluster in Fig. 4*C*). For clarity, each antigen is colored with a different color. The black boxes mark the clique profile or representative clique reactivities—the averages of the reactivities of the clique's antigens for each individual.
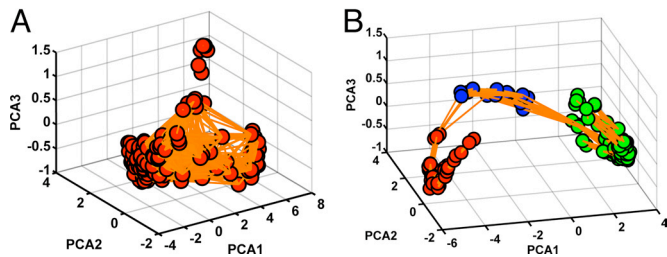
Fig. 6. IH: The cord antigen networks in the PCA space. (*A*) The antigen network for the large cluster of 150 antigens identified for the cords' IgM isotype. (*B*) The antigen network of the 132 antigens that form subgroups for the IgG isotype. For both networks, nodes (antigens) with high correlations (>0.85) are linked with color coded lines according to the correlation levels.

subgroup. Using the IH presentation, we found that this subgroup does not manifest any noticeable organization (Fig. 6*A*). This observation indicates that the universal congenital IgM immune network is not organized into antigen reactivity cliques as is the case for the mothers. Thus, the modular organization into autoantibody cliques found in the mothers had to have been formed during the development of the immune system after birth.

Calculating the IH manifold for the cords' IgG autoantibody correlation matrix (Fig. 6*B*), we identified three intermingled subgroups that were closely located with cross correlations, as if the organization of IgG detected in the mothers' repertoires was not fully present in the cord IgG repertoires. These findings indicate that whereas the IgG autoantibodies are transferred from the mother to the fetus, the global IgG profile at birth is less developed than that of the mothers.

In Fig. 7 we present additional comparisons between the mature and congenital IgG autoantibody organization (for additional angles of view, see the movies of rotating manifolds, section VIII in *SI Text* and Movies S5 and S6). Whereas a weak modular organization is detected for the congenital IgG, there is no actual separation between the different clusters. We also see that the cords' immune organization reflects that of the mothers—the different antigen reactivities maintain their relative positions in PCA space. These findings support the idea that the formation of structural organization (reflected by the formation of autoantibody

cliques) is associated with the normal evolution of the immune system. These findings also illustrate that the transfer of maternal autoantibodies to the fetus is not sufficient to transfer the maternal modular organization of the repertoire, implying that this organization evolves through immune learning.

Discussion

We compared maternal IgM and IgG autoantibody repertoires with those of their newborns using system level analysis of antibody reactivities measured by antigen microarrays. For the IgM autoantibodies, which are not transferred from mothers to fetus, we found that all of the newborns shared a universal immune profile, in agreement with the concept of the immunological homunculus—the immune system's internal image of key body molecules (2, 3). The mothers' IgM and IgG autoantibody repertoires are highly diverse, implying that healthy development of the autoimmune state from birth to adulthood arises from immunological learning according to ones personal life experience. In other words, the immunological homunculus is not static but responds with personal immune experience. At present, we do not know if healthy maturation of the autoantibody repertoire is induced by immune experience with foreign antigens cross-reactive with self, by autoimmune contact with sets of self antigens during normal body maintenance (2, 3), or by a combination of these processes. It would be important to learn whether particular types of autoimmune repertoire organization herald the later development of clinical autoimmune disease (9, 10).

Analyzing the maternal IgM and IgG autoantibody correlations, we found that immune state diversity goes hand in hand with the development of modular organization reflected by the formation of antigen reactivity cliques or functional immune groups (FIGs). To test the statistical validity of the results, we compared the FIGs identified for the 10 mothers with those identified by analyzing the dataset including also the additional 5 females and found high robustness (see *SI Text*). Our findings of system-level modular organization of the autoantibody repertoire are reminiscent of the idiotypic network idea by Jerne (31–33); idiotypic networks too might be deciphered using defined antibody microarrays and the recently developed special system-level analysis method adopted here. We also note that the modular organization we found represents a system level feature of the immune system that was not predicted by the clonal
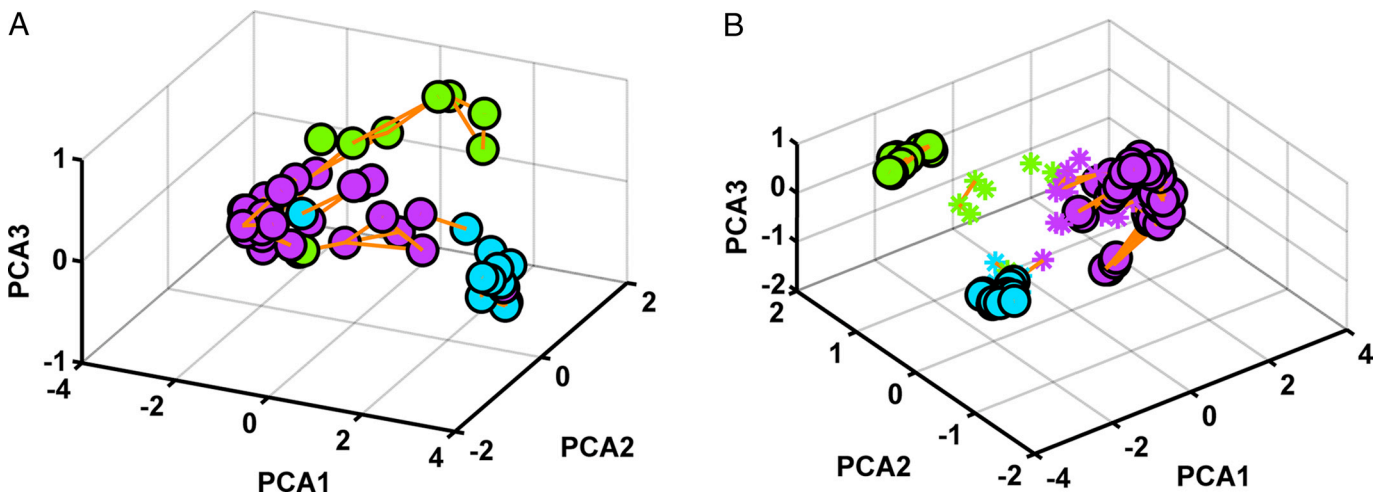
Fig. 7. Comparing congenital and mature immune organization. (*A*) The antigens of the maternal IgG cliques, presented according to their correlations in the cords dataset (compare with Fig. 4*B*, where the same antigens are presented according to their correlations in the maternal dataset). The color code of the cliques is the same as in Fig. 4*B* and the links between the antigens are for correlations (computed for the cords dataset) >0.85. (*B*) Combined presentation. The same antigens are presented according to both the maternal correlations (circles) and the cords correlations (asterisks) in the maternal PCA space. Nodes with correlations >0.85 are linked. The organization of the maternal antigen network is weakly detected in the cords' antigen network.

selection theory (2). However, larger datasets are needed to verify the existence of autoantibody cliques and the complete identification of the antigen reactivities of each clique.

Modular organization and antigen reactivity cliques were also identified for IgG autoantibodies. The IgG isotype is transported across the placenta from mother to fetus, yet, compared with the mothers, we found only a rudimentary modular organization in cord blood. Because the IgG autoantibody correlations for each mother-cord pair are high, the revealed differences call for closer inspection, which is presented in the *SI Text*.

The antibodies in a given clique, or FIG, are characterized by similar antigen reactivities for each mother, and the values of a clique's representative reactivity are different for the different mothers. These features can be used for identification of a personal autoimmune state of organizational motifs and hence might be used for various diagnostic purposes. For example, large variations between the reactivities of autoantibodies to the antigens that belong to the same clique or if one or more antigens have reactivities that substantially differ from the representative reactivity of the clique may indicate a disease (9, 10) or susceptibility for a disease (11, 12). For a specific mother, each autoantibody clique has its own representative antigen reactivity that differs from the representative antigen reactivity of the other cliques for the same mother; in other words, each mother expresses her own profile of representative clique reactivities. This individual profile reflects a personal immune state organization and hence it might be used for comparison between the immune states of different persons (12, 46).

In addition to the FIGs presented in this article, weaker cliques were also identified (see section IV in *SI Text* and Fig. S5). The existence of inner-layer structures and the ability to unmask the dominant motifs may become of special importance in the context of healthy versus pathologic structures. Whereas the leading motifs may not be changed in the case of an active (or past) disease, more sensitive changes might still be detected. Indeed, the approach presented here might be applied to study the immune system under different conditions and in various aspects of immunity: breast-feeding, childhood and growth, vaccinations, various disease states, autoimmune diseases, the influence of medical treatment (e.g., chemotherapy, Immunosuppressive drugs) and many more. In any case, the formation of immune modular organization expressed by autoantibody cliques reflects a functional system-level network organization of the immune system, an embodiment of the immunological homunculus.

To conclude, in this paper we introduce a powerful and efficient approach to the study of immune data. Whereas the available data are currently limited, the statistically significant results presented here will hopefully trigger additional experimental studies that will lead to a broader and deeper understanding of the immune system as a dynamically developing and complex network.

1. Janeway CA, Travers P (2005) *Immunobiology: The Immune System in Health and Disease* (Garland Science, New York).
2. Cohen IR (2000) *Tending Adam's Garden: Evolving the Cognitive Immune Self* (Elsevier Academic, London).
3. Cohen IR (1992) The cognitive paradigm and the immunological homunculus. *Immunol Today* 13:490–494.
4. Cohen IR (2000) Discrimination and dialogue in the immune system. *Semin Immunol* 12:215–219.
5. Perelson AS, Weisbuch G (1992) *Theoretical and Experimental Insights into Immunology* (Springer-Verlag, New York).
6. Perelson AS, Weisbuch G (1997) Immunology for physicists. *Rev Mod Phys*, 69(4):1219–1268.
7. Tauber A (1994) *The Immune Self: Theory or Metaphor?* (Cambridge Univ Press, Cambridge, UK).
8. Tauber A (2006) The biological notion of self and non-self. *The Stanford Encyclopedia of Philosophy*, Zalta EN, ed, http://plato.stanford.edu/archives/spr2006/entries/biology-self/.
9. Quintana FJ, Merbl Y, Sahar E, Domany E, Cohen IR (2006) Antigen-chip technology for accessing global information about the state of the body. *Lupus* 15:428–430.
10. Cohen IR (2007) Biomarkers, self-antigens and the immunological homunculus. *J Autoimmun* 29:246–249.
11. Quintana FJ, Cohen IR (2004) The natural autoantibody repertoire and autoimmune disease. *Biomed Pharmacother* 58:276–281.
12. Quintana FJ, et al. (2004) Functional immunomics: Microarray analysis of IgG autoantibody repertoires predicts the future response of mice to induced diabetes. *Proc Natl Acad Sci USA* 101(Suppl 2):14615–14621.
13. Robinson WH (2006) Antigen arrays for antibody profiling. *Curr Opin Chem Biol* 10:67–72.
14. Golub TR, et al. (1999) Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* 286:531–537.
15. Quinlan J (1992) *Programs for Machine Learning* (Morgan Kaufmann, San Mateo, CA).
16. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. *Parallel Distributed Processing*, eds Rumelhart DE, McClelland JL (MIT Press, Cambridge, MA), Vol 1, pages 318–362.
17. Furey TS, et al. (2000) Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16:906–914.
18. Hartigan JA (1975) *Clustering Algorithms* (Wiley, New York).
19. Everitt BS (1993) *Cluster Analysis* (Edward Arnold, London).
20. Mirkin B (1996) *Mathematical Classification and Clustering* (Kluwer Academic Publishers, Dordrecht, The Netherlands).
21. Hansen P, Jaumard B (1997) Cluster analysis and mathematical programming. *Math Program* 79:191–215.
22. Pilpel Y, Sudarsanam P, Church GM (2001) Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet* 29:153–159.
23. Getz G, Levine E, Domany E (2000) Coupled two-way clustering analysis of gene microarray data. *Proc Natl Acad Sci USA* 97:12079–12084.
24. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868.
25. Tamayo P, et al. (1999) Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA* 96:2907–2912.
26. Alon U, et al. (1999) Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc Natl Acad Sci USA* 96:6745–6750.
27. Hartuv E, et al. (2000) An algorithm for clustering cDNA fingerprints. *Genomics* 66:249–256.
28. Ben-Dor A, Shamir R, Yakhini Z (1999) Clustering Gene Expression Patterns. *J Comput Biol* 6:281–297.
29. Sharan R, Shamir R (2000) Center CLICK: A clustering algorithm with applications to gene expression analysis. *Proc Int Conf Intell Syst Mol Biol* 8:307–316.
30. Merbl Y, Zucker-Toledano M, Quintana FJ, Cohen IR (2007) Newborn humans manifest autoantibodies to defined self molecules detected by antigen microarray informatics. *J Clin Invest* 117:712–718.
31. Jerne NK (1967) Various basic problems of current immunology. *Landarzt* 43:1526–1530.
32. Jerne NK (1974) Towards a network theory of the immune system. *Ann Immunol (Paris)* 125C:373–389.
33. Jerne NK (1974) The immune system: A web of V-domains. *Harvey Lect* 70 Series:93–110.
34. Baruchi I, et al. (2006) Functional holography analysis: Simplifying the complexity of dynamical networks. *Chaos* 16:015112.
35. Madi A, et al. (2008) Genome holography: Deciphering function-form motifs from gene expression data. *PLoS ONE* 3:e2708.
36. Chou Y (1975) *Statistical Analysis* (Holt International, New York), section 17.9.
37. Mathworks (2006) MATLAB (Mathworks, Natick, MA).
38. Rodgers JL, Nicewander WA (1988) Thirteen ways to look at the correlation coefficient. *Am Stat* 42:59–66.
39. Jolliffe IT (1986) *Principal Component Analysis* (Springer-Verlag, New York).
40. Gidskehaug L, Anderssen E, Flatberg A, Alsberg BK (2007) A framework for significance analysis of gene expression data using dimension reduction methods. *BMC Bioinformatics* 8:346.
41. Boulesteix AL, Strimmer K (2007) Partial least squares: A versatile tool for the analysis of high-dimensional genomic data. *Brief Bioinform* 8:32–44.
42. Baruchi I, Ben-Jacob E (2004) Functional holography of recorded neuronal networks activity. *Neuroinformatics* 2:333–352.
43. De Boer RJ, Perelson AS (1993) How diverse should the immune system be? *Proc Biol Sci* 252:171–175.
44. Hanson LA, et al. (2003) Breast-feeding, infant formulas, and the immune system. *Ann NY Acad Sci* 987:199–206.
45. Bosko TJ, Todd DB, Sadus JR (2006) Analysis of the shape of dendrimers under shear. *J Chem Phys* 124:044910.
46. Cohen IR (2007) Real and artificial immune systems: Computing the state of the body. *Nat Rev Immunol* 7:569–574.