Systematic Classification of Primary Immunodeficiencies Based on Clinical, Pathological, and Laboratory Parameters

Crina Samarghitean, Csaba Ortutay and Mauno Vihinen

*J Immunol* 2009; 183:7569-7575; Prepublished online 16 November 2009;
doi: 10.4049/jimmunol.0901837
http://www.jimmunol.org/content/183/11/7569

Supplementary Material  http://www.jimmunol.org/content/suppl/2009/11/16/jimmunol.090183
7.DC1

Why *The JI*?

- **Rapid Reviews! 30 days** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Speedy Publication!** 4 weeks from acceptance to publication

*average

References  This article cites 71 articles, 9 of which you can access for free at:
http://www.jimmunol.org/content/183/11/7569.full#ref-list-1

Subscription  Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions  Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts  Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Systematic Classification of Primary Immunodeficiencies Based on Clinical, Pathological, and Laboratory Parameters

Crina Samarghitean,*† Csaba Ortutay,* and Mauno Vihinen2*†

The classification of diseases has several important applications ranging from diagnosis and choice of treatment to demographics. To date, classifications have been successfully created manually, often within international consortia. Some groups of diseases, such as primary immunodeficiencies (PIDs), are especially hard to nosologically cluster due, on one hand, to the presence of a wide variety of disorders and, in contrast, because of overlapping characteristics. More than 200 PIDs affecting components of the innate and adaptive immune systems have been described. Clinical, pathological, and laboratory characteristics were collected and used to group PIDs. A consensus of at least five independent methods provided a novel classification of 11 groups, which revealed previously unknown features and relationships of PIDs. Comparison of the classification to independent features, including the severity and therapy of the diseases, functional classification of proteins, and network vulnerability, indicated a strong statistical support. The method can be applied to any group of diseases.

Primary immunodeficiencies (PIDs) are a large and heterogeneous group of disorders that have been organized manually into different categories (1–8) sometimes without a consensus. PIDs are mainly rare hereditary disorders of the immune system that often have serious consequences (1, 2). These diseases represent a challenge in their diagnosis and treatment due to overlapping symptoms and similarities between diseases. Infections are the hallmarks of PIDs (1, 9–11). A diagnosis will often be considered when infections are frequent or severe, resistant to standard therapies, or caused by unusual (opportunistic) organisms. Other manifestations include autoimmune (12, 13) and cancer diseases (14), granulomatosis (15), hemophagocytic syndrome (16, 17), angioedema (18), autoinflammation (19–21), thrombotic microangiopathy, or predisposition to allergy.

Clinical descriptions have already been made for more than 200 PIDs (4, 7, 22), for which 167 genetic etiologies have been described. PIDs have historically been defined and classified according to the affected immune system that often have serious consequences (1, 2). These diseases represent a challenge in their diagnosis and treatment due to overlapping symptoms and similarities between diseases. Infections are the hallmarks of PIDs (1, 9–11). A diagnosis will often be considered when infections are frequent or severe, resistant to standard therapies, or caused by unusual (opportunistic) organisms. Other manifestations include autoimmune (12, 13) and cancer diseases (14), granulomatosis (15), hemophagocytic syndrome (16, 17), angioedema (18), autoinflammation (19–21), thrombotic microangiopathy, or predisposition to allergy.

Clinical descriptions have already been made for more than 200 PIDs (4, 7, 22), for which 167 genetic etiologies have been described. PIDs have historically been defined and classified according to the affected immune system that often have serious consequences (1, 2). These diseases represent a challenge in their diagnosis and treatment due to overlapping symptoms and similarities between diseases. Infections are the hallmarks of PIDs (1, 9–11). A diagnosis will often be considered when infections are frequent or severe, resistant to standard therapies, or caused by unusual (opportunistic) organisms. Other manifestations include autoimmune (12, 13) and cancer diseases (14), granulomatosis (15), hemophagocytic syndrome (16, 17), angioedema (18), autoinflammation (19–21), thrombotic microangiopathy, or predisposition to allergy.

Clinical descriptions have already been made for more than 200 PIDs (4, 7, 22), for which 167 genetic etiologies have been described. PIDs have historically been defined and classified according to the affected immune system that often have serious consequences (1, 2). These diseases represent a challenge in their diagnosis and treatment due to overlapping symptoms and similarities between diseases. Infections are the hallmarks of PIDs (1, 9–11). A diagnosis will often be considered when infections are frequent or severe, resistant to standard therapies, or caused by unusual (opportunistic) organisms. Other manifestations include autoimmune (12, 13) and cancer diseases (14), granulomatosis (15), hemophagocytic syndrome (16, 17), angioedema (18), autoinflammation (19–21), thrombotic microangiopathy, or predisposition to allergy.

Clinical descriptions have already been made for more than 200 PIDs (4, 7, 22), for which 167 genetic etiologies have been described. PIDs have historically been defined and classified according to the affected immune system that often have serious consequences (1, 2). These diseases represent a challenge in their diagnosis and treatment due to overlapping symptoms and similarities between diseases. Infections are the hallmarks of PIDs (1, 9–11). A diagnosis will often be considered when infections are frequent or severe, resistant to standard therapies, or caused by unusual (opportunistic) organisms. Other manifestations include autoimmune (12, 13) and cancer diseases (14), granulomatosis (15), hemophagocytic syndrome (16, 17), angioedema (18), autoinflammation (19–21), thrombotic microangiopathy, or predisposition to allergy.
Immunodeficiencies and selection of the parameters

Data for PIDs were collected from the ImmunoDeficiency Resource (IDR) (4), IDdiagnostics (37), IDbases (38), and literature (1, 2, 5–7, 39–43). Only detailed reports with statistical information including clinical symptoms and measured laboratory values characterizing PIDs were included. When diseases without specific symptoms were excluded, there were altogether 194 PIDs left.

For each disease, all signs, symptoms, and laboratory values mentioned in the literature were collected. The initial list contained 420 parameters. Parameters characterizing only one to four diseases were omitted or merged to others.

For example, IgM lymphoma, which is according to the literature associated only with hyper-IgM syndrome type 2, was merged to the more general term of other malignancies. Similarly, encephalitis, meningoencephalitis, menigitis, conjunctivitis, iritis, episcleritis, and brain abscess were grouped to a more general term of CNS infections. Cerebellar ataxia, pathognomonic for ataxia-telangiectasia, and ataxia-telangiectasia-like disease were grouped to neurological or CNS abnormalities together with other signs, such as peripheral neuropathy, speech delay, and convulsions. Finally, after iterative process, we had 87 informative parameters (supplemental Table S1). All of the parameters had an equal weight in the analysis.

Cluster and network community analysis

Cluster and network analyses were performed in the R statistical environment (44) using the igraph (45) and cluster program libraries. Three different variations of K-means clustering were used to analyze the dataset. The first was clustering Large Applications (clara) method computes a list representing the clustering of the data into k clusters. Partitioning Around Medoids (pam) partitions (clusters) the data into k clusters around medoids, which are representative objects of a dataset from which the distances to the other points in the cluster are computed. The Fuzzy Analysis Clustering (fanny) method computes a partition grouping of the data into k clusters. The number of clusters was chosen by maximizing the average width of the clusters. In fuzzy clustering, data elements can belong to more than one cluster, and thus, each disease has a set of membership levels, which indicate the strength of association to each cluster.

Three methods were applied to find highly interconnected parts of the network. Community structure via short random walks is a walktrap community analysis, which searches for densely connected subgraphs, i.e., communities (34). When moving from one node to a connected one, short random walks tend to stay in the same community. The second method uses community structure detection based on the leading eigenvector of the community matrix (35). The method looks for densely connected subgraphs by calculating the leading nonnegative eigenvector of the modularity matrix of the graph. The third method tries to find communities in graphs via a spin-glass model and simulated annealing (36).

Combination of clustering and network results

The data for PIDs are incomplete because in many diseases just a few, even a single, patient was known, and therefore, the most prominent signs were difficult to define. In rare diseases, some symptoms may occur frequently just by chance. Therefore, to obtain the most reliable and robust grouping and a consistent and robust view of the disease grouping patterns, a consensus classification based on the co-occurrence of the diseases in four, five, or six methods was generated. Using this approach, the results are expected to be independent from the biases of the individual methods.

Statistics

To evaluate which of the binary parameters significantly supported the identified disease clusters, we tested whether a parameter had significantly different distribution in the individual clusters compared with the entire dataset by calculating p values using the hypergeometric distribution. The threshold for significance used was a p value 0.05. A similar evaluation was performed for the binary clinical and functional properties of the diseases and respective proteins.

Correlation of disease clusters to clinical, genetic, functional, and network properties

The consensus network graph was used to visualize a number of properties for the diseases, and involved genes and proteins. Information about the prevalence of the PIDs was obtained from the IDR (4) and other sources (1, 2, 40, 46). When the prevalence was not known and only a few cases were reported in literature, it was assumed to be <1/107. Data on inheritance were retrieved from the IDR and literature (1, 2, 6, 40). For treatment modalities, the most common treatments were listed for each disease (1, 2, 39). Functional classifications of the proteins in the immunome (47) were obtained from the Immunome Knowledge Base (48).

Results

PID network

Novel PID grouping was obtained by applying altogether six methods for clustering and network analysis. Our approach revealed associations that were not previously obvious and led to the identification of distinct novel groups. Fig. 1A shows the results when at least five methods agreed on the clustering. The diseases are indicated by the affected genes, when known. If all six methods are required to agree, the only difference is that the 11 major groups are divided further to smaller subgroups. There is one giant cluster that contains the majority of the PIDs and some separate clusters and singleton PIDs. Some details of the classification may change in the future when more information becomes available, yet still the major features will remain. In the consensus graph for the PIDs (Fig. 1A), 1,285 pairs of diseases are grouped together by at least five methods of a possible 18,721, and of the 12,721 that are grouped by at least one method. The results are also available in an interactive web page at http://bioinf.uta.fi/PID_classification.

Disease clusters

The analytical approach combining six different and independent methods provides a highly robust classification when at least five of the methods were required to agree on the grouping details. The dendrogram reveals 11 well-defined disease clusters (DCs) of at least 4 diseases (Fig. 1A). First, we analyzed the nature of the PIDs in the clusters and then investigated the properties of the diseases and the corresponding genes and proteins in the clusters.

Most of the clusters are very homogeneous and contain related diseases. An overall view of the DCs shows that clusters III and VII contain (almost) exclusively SCIDs, whereas in group IX the diseases are related to the complement system and in DCs I and XI to phagocyte functions; in DC VIII are fever syndromes, and in DC X Fanconi anemias. All of the MHC II genes are in cluster III, and all of the known MHC I diseases are in cluster V. The classical complement pathway diseases are in clusters V and IX. Diseases in cluster II are related to DNA instability and DNA damage repair, except for G6PC3. DCs IV, V, and VI are the most heterogeneous. Cluster IV contains numerous receptor and signaling molecule-related diseases. Some of the proteins behind these disorders form transmembrane channels. DC V contains mainly Ab and complement deficiencies together with some SCIDs. The majority of the group VI diseases are related to phagocytosis and apoptosis. All of the groups have strong statistical support (supplemental Table S2).

These groups, which have been exclusively generated based on disease characteristics, indicate the power of the method. Other information for the PID genes, proteins, and their functions further support the classification (Fig. 1, B–E). Also, these results are statistically significant (supplemental Table S2).

The previous PID classifications have relied heavily on the cell types in which the disease-related genes are normally expressed. Thus, one of the major differences to these is that Ab deficiencies, combined PIDs, and diseases related to phagocytosis are widely scattered in our graph (Fig. 1). These diseases are very heterogeneous in their symptoms and signs, and affect numerous parts of the immune system. The highest concentration of SCIDs is in DCs II, IV, and VII, whereas Ab deficiencies mainly appear in clusters IV and V. Phagocyte diseases are exclusive to DCs I and XI, but also appear in DCs IV, V, and VI.
To test whether the etiologically rather homogeneous DCs shared any similarities, the distribution of characteristics describing independent clinical, functional, genetic, and network properties of the diseases and the corresponding genes and proteins were investigated. These features were not used for the original clustering of the PIDs.

The PIDs were divided into four groups according to their severity. Severity was not among the symptoms used in our classification because it is not routinely used in the clinical description of the diseases. SCID is a pediatric emergency situation because the condition is life threatening, whereas some of the other PIDs are just mild. In fact, the vast majority of PID cases have been thought to remain undiagnosed due to mild symptoms. The severity shows a very homogeneous pattern in some of the clusters (Fig. 1B). Diseases in DCs II, VII, and IX belong to two categories, whereas in almost all of the remaining clusters there are almost exclusively moderate and severe or severe and life-threatening diseases.

PIDs are treated in a number of ways that can be grouped to a small number of categories, including Ig treatment, the use of antibiotics, antifungals or antivirals, immunomodulators, and the reconstitution of the immune system by (haploidentical) bone marrow transplantation or hematopoietic stem cell transplantation. The data for the treatment of the diseases in Fig. 1C indicate that the majority of the clusters are very homogeneous in regard to treatments and there are clear differences between DCs. There are DCs in which all of the diseases are treated with the same battery of therapeutic modalities, and in almost all the clusters some of the therapeutic options can be used for all of the coclustered PIDs. Also, based on these results, the DCs reliably reflect the properties of diseases, and the therapy applied to diseases within DCs is usually similar.

The cellular functions have been determined for all the genes and proteins required in the immune system (entirety of immune system) (49). There are functional groups for, e.g., the surface receptors in clusters of differentiation classification, chemokines, and their receptors, humoral immunity proteins, and those involved in cellular immunity, Ag processing, and transcription factors. The distribution of the functional properties in the PID classification is shown in Fig. 1D.

The majority of the group I proteins are involved in inflammation and cellular immunity. Group II and VII proteins have functions as transcription factors involved in humoral and cellular immunity. Humoral immunity is most prevalent in DC IV, complement proteins in DC IX, and both humoral and complement functions in DC V. Inflammation is the function involved in DC VIII. Many of the proteins have several classifications because they are typically overlapping. Also, at this level, the grouped proteins and diseases share many common properties because the functions are very homogeneous within DCs and differ between them.

Vulnerability is a systems biology measure that indicates how crucial a certain node is for the network (30). Vulnerabilities of proteins in the immune protein interaction network were color coded compared with the average vulnerability in the entire network (Fig. 1E). There are some clusters in which the vulnerabilities are related, especially those in DCs III, VI, and VII. The most vulnerable diseases are widely scattered throughout the network. Only some of the SCID proteins that are related with the most severe PIDs are highly vulnerable. Also, previously, the vast majority of disease genes were shown to be peripheral in the network (27).

Inheritance pattern (supplemental Fig. S1) and prevalence (supplemental Fig. S2) did not show any correlation within DCs at all. This was expected because these characteristics are not likely to affect the etiology of diseases.

Discussion

We developed a novel approach for nosology and produced a systematic classification of PIDs based on parameters across several clinical, pathological, and physiological dimensions. Our approach combines existing clustering and network partition methods to classify these diseases. The new classification shares certain features with previous groupings, yet is different in a number of details. For example, the new classification indicates that cell-type expression, which has previously been one of the major classification criteria, cannot be very reliably used for classification of PIDs.

Clustering methods are widely used in many fields, such as in microarray data analysis. In medicine, cluster analysis has been used in the nosological splitting of different phenotypes, for example in Marshall and Stickler syndromes (50, 51), and more recently to classify patients with chronic pain (52) and to develop a new taxonomy for airway diseases (53, 54). Because the disease data were not complete due to many of the PIDs being extremely rare, we combined both network and clustering methods and used their consensus to obtain a robust grouping for PIDs. Considering the consensus of four, five, or six independent methods makes the results robust and independent from the individual methods.

Comparison with previous classifications

Previous in silico disease classification methods have been based on shared genes in disorders (27), protein interactions (55), protein complexes (28), or tissue-specific gene expression (29). For example, Human Phenotype Ontology terms describe clinical features and can be used for disease grouping (56). However, because only a few PIDs were included in these classification studies, we unfortunately cannot compare our results with them.

The new PID classification differs in details from those published earlier. The most comprehensive classification with the largest number of PIDs is from the European Society for Immunodeficiencies (ESID) registry (7) based on the International Union of Immunological Societies classification of 150 diseases (6). This scheme contains seven defined disease groups. The IDR uses and expands the classification (from Ref. 2) in 11 classes. The American Academy of Allergy, Asthma, and Immunology, the American College of Allergy, Asthma, and Immunology, and the Joint Council of Allergy, Asthma, and Immunology have classified 97 PIDs in 5 groups (39). The International Classification of Diseases (ICD10) contains 100 PIDs in 10 categories (8). These classifications have been useful, although they have disagreed on a number of disorders. The task of classifying the widely variable PIDs is hardly any more possible to do manually due to the very high dimensionality of the data. A systematic, mathematical classification can be used as an alternative or complement approach to the existing methods.

The earlier classifications were color coded and visualized in supplemental Fig. 3. The previous studies contained only subsections of the PIDs. The color codes were also chosen so that inter-classification comparisons are possible because related diseases in the different groupings have the same colors. The more homogenous the color is within a cluster, the more similar the classifications are. The IDR and ESID classifications agree very well in DCs I, II, IV, V, VII, VIII, IX, X, and XI. ICD10 agrees well in DCs III, V, VII, IX, X, and XI. There are, however, only 100 PIDs in the ICD system. The American Academy of Allergy, Asthma, and Immunology grouping behaves similarly to ICD, except for DC X, diseases that are not included at all. In conclusion, the novel
FIGURE 1. Consensus for the six methods used to group the PIDs. A, Relationship of the diseases. The white rectangles indicate the grouping of diseases by five (those attached to the first black bullet from the PIDs root) and six (the second bullet) methods. Diseases are indicated by the systematic names of affected genes, when known. Otherwise, the following names were used when no genes are identified in relation to the disease: AR-HIES, autosomal recessive hyperimmunoglobulin E recurrent infection syndrome; CMC, chronic mucocutaneous candidiasis; CVID, common variable immunodeficiency; FHL1, familial hemophagocytic lymphohistiocytosis type 1; HIGM4, hyper-IgM syndrome type 4; SADNI, specific Ab deficiency with normal Ig concentrations; THI, transient hypogammaglobulinemia of infancy; thymoma, thymoma with immunodeficiency (Good’s syndrome); and XLA/GHD, X-linked hypogammaglobulinemia with growth hormone deficiency. When one gene is involved in more than one disease, the following abbreviations were used:
classification agrees with numerous features in the earlier groupings. All of the previous classifications have a smaller number of groups, which are typically divided in the new nosology. There are also a number of diseases that are not consistent with the novel classification, or with the other previous groupings. The systematic mathematical approach is capable of solving these cases and places the diseases in groups using solely the given parameters.

**Parameters characterizing primary immunodeficiencies**

In PIDs, the signs and symptoms of infections may be repetitive, severe, or refractory to therapy and caused by organisms with low virulence (39). We grouped the parameters for infections according to microbial taxonomy and site of infection.

Autoimmune diseases and malignancies are complications of many immunodeficiencies. Based on the frequency of associations with autoimmune diseases, PIDs can be grouped in three groups, as follows: systematic (>80% of the patients with the disease have autoimmune disease symptoms), strong (20–80%), and mild (<20% of the patients) and absent (12, 57). Some PID patients appear susceptible also to atopy and lupus-like syndromes (58).

Malignancies occur with great frequency in certain immunodeficiencies. The types of malignancies depend on the PID, the age of the patient, and any possible viral infection(s). B cell malignancies, especially non-Hodgkin’s lymphomas, are predominant. Other types of malignancies encountered in PIDs are T cell malignancies and leukemia (14).

Problems in lymphoproliferation (hepatomegaly, splenomegaly, lymphadenopathy) are typical for some PIDs. EBV infection is associated with many lymphoproliferation-linked immunodeficiencies (59–61).

Some PID patients have chronic respiratory problems, such as asthma, chronic obstructive disease, chronic inflammatory lung disease, emphysema/lung cysts, or interstitial pneumonia (62–65). Cardiovascular diseases such as congenital cardiac anomalies, cardiomyopathy, and hematologic abnormalities are found in certain PIDs (66, 67).

Some of the gastrointestinal diseases are associated with PIDs, including esophageal atresia, Crohn’s disease, chronic inflammatory bowel disease, ulcerative colitis, granulomatous colitis, celiac disease, malabsorption, Hirschsprung disease, and anal stenosis (15, 68–70). Hepatobiliary tract diseases in PID patients include storage liver disease, hepatic vascular occlusion disease, and sclerosing cholangitis (71, 72). Kidney diseases associated with PIDs include renal anomalies, renal dysfunction, amyloidosis, renal failure, IgA nephropathy, and glomerulonephritis (73).

Although physical findings are often absent, may be nonspecific, or very discreet, many PIDs have characteristic features. A common feature in PID patients is failure to thrive (child) or wasting (adult). Facial abnormalities, such as microcephaly or dysmorphism, are characteristic in some PIDs. Neurological abnormalities can include ataxia, peripheral neuropathy, speech delay, mental retardation, retinal lesions, photophobia, convulsions, or psychomotor retardation, whereas gastrointestinal abnormalities appear as severe gingivostomatitis, recurrent aphtha, periodontitis, delayed shedding of primary teeth, palatal weakness/cleft, gastric outlet obstruction, or diarrhea (1, 2, 39–41).

Ligamentous laxity/hyperextensive joints, limited extension of elbows, costochondral junction abnormality, rib abnormalities, metaphyseal chondrodysplasia/dysostosis, pectus carinatum, spondyloepiphysyeal dysplasia, hip degeneration, or short limb dwarfism are among the skeletal abnormalities found in PIDs (74). The skin is frequently affected in immunodeficiencies. Erythoderma, eczema/ atopic dermatitis, pyoderma, or granuloma is common. The presence of petechiae or bruises suggests a bleeding problem, as in phagocyte disorders (75–79).

The definitive diagnosis of PIDs depends on laboratory evidence, including assessments of humoral and cellular immunity and molecular analysis. The immunologic phenotype is based on laboratory tests of immune function, such as serum Ig levels, specific Ab titers, peripheral blood lymphocyte subpopulations, measures of T cell function, assays of phagocytes, and complement function or serum component level (1, 2, 37, 39–41).

Most of the parameters have binary values (yes, no), whereas, for example, the laboratory parameters are quantitative. The parameters were chosen so that they represent different important features of PIDs.

**The new PID classification**

The obtained 11 disease clusters are very robust due to them being the consensus of at least five methods. The p values show the significance of the observations. More detailed subgrouping is available by using the consensus of all the six methods. In Fig. 1A, diseases coclustered by all the six methods are within the boxes, whereas in the DCs at least five methods agree on the placement of PIDs within the graph. In addition to the actual classification, the PID parameters could offer guidelines for medical descriptions of PIDs. The classification allows a novel and fresh look at the relationships of PIDs, the genes behind them, and the encoded proteins. The network is far more complex than the previous mainly cell-type-based groupings might have led to imagine.

Based on the classification and the parameters, it might be possible to develop novel diagnostic schemes for PIDs.

The correlation with other independent information not used for the original classification implies that the classification reliably reflects numerous properties of the diseases and the genes and proteins mutated in them. Data for the disease parameters were not complete, especially for the ultra-rare PIDs, and thus the homogeneity of the groups could even increase in the future when more reliable statistical information for the symptoms and laboratory characteristics will be available.

Diseases affecting genes involved in the same pathway do often cocluster, and they share a similar etiology, for example, JAK3 and IL2RG, and proteins in IFN-mediated immunity, including STAT1 and IFNGR2; MHC II diseases in DC III are all parts of the same protein complex as well as the MHC I components in DC V; complement components are in DC V and IX, and components of membrane bound oxidase in DC XI; BLNK, BTK, and MyD88, which have been suggested to be downstream of CD19 signaling, are all in DC V.

The new PID classification resulting from our approach can have several applications. Because it was generated independently from the existing classifications using solely mathematical analysis of clinical
and laboratory parameters, it can be used in evaluation and development of other classifications. Disease groups defined by experts and found also by our independent approach have a strong indication of their existence. The new classification will be used in IDR (4) for organizing diseases and information about them. Another possible application is detailed demographics and mortality records.

The classification and the dataset also serve the development of diagnostic expert systems, which requires objective criteria for diagnosis. Expert systems are useful, especially in case of rare diseases like PIDs. The computer-based disease classification applied in this study can also identify the key symptoms and laboratory parameters that can help the experts to diagnose correctly these diseases.

The approach can be applied also to other groups of diseases.

Acknowledgments

We thank Jukka Lehtiniemi for help with the figures.

Disclosures

The authors have no financial conflict of interest.

References


Systematic classification of primary immunodeficiencies based on clinical, pathological and laboratory parameters

Crina Samarghitean, Csaba Ortutay, Mauno Vihinen

Supplemental legends:

Fig. S1. Inheritance of PIDs.

Fig. S2. Prevalence of PIDs.

Fig. S3. Comparison of PID classification schemes.

Table S1. Parameters used to characterize signs, symptoms and laboratory values in PIDs.

Table S2. Parameters statistically enriched in Disease Clusters.
Fig. S1. Inheritance of PIDs. The layout of the graph is the same as in Fig. 1A.
Fig. S2. Prevalence of PIDs. The layout of the graph is the same as in Fig. 1A.
Fig. S3. Comparison of PID classification schemes. The layout of the graph is the same as in Fig. 1A.
The novel PID classification was compared with four previous systems In each cells the four coloured columns represents the following disease classification systems: IDR, ESID, AAAAI, ICD-10 from left to right, respectively. The colour coding for the individual systems is as follows:

ImmunoDeficiency Resource (IDR) classification
0 Not available
1 Combined B and T cell immunodeficiencies
2 Deficiencies predominantly affecting antibody production
3 Defects in lymphocyte apoptosis
4 Other well-defined immunodeficiency syndromes
5 Defects of phagocyte function
6 Defects of innate immune system, receptors and signaling components
7 Periodic fever syndromes
8 Defects of the classical complement cascade proteins
9 Defects of the alternative complement pathway
10 Defects of complement regulatory proteins
11 DNA breakage associated syndromes and DNA epigenetic modification syndromes

The European Society for Immunodeficiencies (ESID) classification
0 Not available
1 Predominantly T cell deficiencies, T B Severe combined immunodeficiency (SCID)
2 Predominantly antibody disorder
3 Other well defined PIDs
4 Complement deficiencies
5 Phagocytic disorders
6 Autoimmune and immunedysregulation syndromes
7 Autoinflammatory syndromes
American Academy of Allergy Asthma and Immunology (AAAAI) classification
0 Not available
1 Combined immunodeficiency
2 Humoral immunodeficiency
3 Cellular immunodeficiency
4 Phagocytic cell disorders
5 Complement deficiencies

International Classification of Diseases 10 (ICD-10)
0 Not available
1 Combined immunodeficiencies (D81)
2 Immunodeficiency with predominantly antibody defects (D80)
3 Immunodeficiency associated with other major defects (D82)
4 Other immunodeficiencies (D84)
5 Functional disorders of polymorphonuclear neutrophils (D71)
6 Agranulocytosis (D70)
7 Albinism (E70.3)
8 Cerebellar ataxia with defective DNA repair (G11.3)
9 Constitutional aplastic anaemia (D61)
10 Crohn's disease [regional enteritis] (K50)
**Table S1.** Parameters used to characterize signs, symptoms and laboratory values in PIDs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values(^a)</th>
<th>Number of diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infections</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>locations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>central nervous system</td>
<td>yes/no</td>
<td>44</td>
</tr>
<tr>
<td>upper respiratory tract</td>
<td>yes/no</td>
<td>91</td>
</tr>
<tr>
<td>lower respiratory tract</td>
<td>yes/no</td>
<td>92</td>
</tr>
<tr>
<td>skin</td>
<td>yes/no</td>
<td>52</td>
</tr>
<tr>
<td>gastro-intestinal</td>
<td>yes/no</td>
<td>33</td>
</tr>
<tr>
<td>musculo-skeletal</td>
<td>yes/no</td>
<td>31</td>
</tr>
<tr>
<td>other</td>
<td>yes/no</td>
<td>55</td>
</tr>
<tr>
<td><strong>taxonomy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Susceptibility to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- bacteria, Gram (+) (<em>Streptococcus pneumonia, Haemophilus influenzae</em>)</td>
<td>0: no</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>1: intermediate</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>2: high</td>
<td>30</td>
</tr>
<tr>
<td>- bacteria, Gram (+) (<em>Staphylococcus aureus</em>)</td>
<td>0: low/no</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>1: intermediate</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2: high</td>
<td>29</td>
</tr>
<tr>
<td>- bacteria, Gram (-) (<em>Neisseria</em>)</td>
<td>yes/no</td>
<td>13</td>
</tr>
<tr>
<td>- mycobacteria</td>
<td>0: low/no</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>1: intermediate</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2: high</td>
<td>27</td>
</tr>
<tr>
<td>- viruses</td>
<td>0: no</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>1: intermediate</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2: high</td>
<td>34</td>
</tr>
<tr>
<td>Susceptibility to fungi (especially <em>P. carinii</em>)</td>
<td>0: no</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>1: intermediate</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2: high</td>
<td>33</td>
</tr>
<tr>
<td>- <em>Candida</em></td>
<td>yes/no</td>
<td>43</td>
</tr>
<tr>
<td>Parameter</td>
<td>Values*</td>
<td>Number of diseases</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------</td>
<td>--------------------</td>
</tr>
<tr>
<td>- <em>Aspergillus</em></td>
<td>yes/no</td>
<td>11</td>
</tr>
<tr>
<td>- <em>Cryptococcus</em></td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>- <em>Histoplasma</em></td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>Susceptibility to Protozoa</td>
<td>1: <em>Giardia</em></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2: <em>Giardia+Cryptosporidium</em></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3: <em>Cryptosporidium</em></td>
<td>1</td>
</tr>
<tr>
<td><strong>Immune system dysregulation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>autoimmunity</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frequency</td>
<td>0: AID&lt;20%</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>1: 20&lt;AID&lt;80</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2: AID&gt;80%</td>
<td>11</td>
</tr>
<tr>
<td>systemic lupus erythematosus (and SLE-like syndrome)</td>
<td>yes/no</td>
<td>26</td>
</tr>
<tr>
<td>vasculitis</td>
<td>yes/no</td>
<td>13</td>
</tr>
<tr>
<td>rheumatoid arthritis/juvenile rheumatoid arthritis /arthritis</td>
<td>yes/no</td>
<td>21</td>
</tr>
<tr>
<td>autoimmune haemolytic anaemia (AIHA)</td>
<td>yes/no</td>
<td>23</td>
</tr>
<tr>
<td>idiopathic thrombocytopenia (ITP)</td>
<td>yes/no</td>
<td>17</td>
</tr>
<tr>
<td>endocrine AID</td>
<td>yes/no</td>
<td>6</td>
</tr>
<tr>
<td><strong>malignancies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B cell malignancies</td>
<td>yes/no</td>
<td>24</td>
</tr>
<tr>
<td>T cell malignancies</td>
<td>yes/no</td>
<td>12</td>
</tr>
<tr>
<td>leukemia</td>
<td>yes/no</td>
<td>19</td>
</tr>
<tr>
<td>other</td>
<td>yes/no</td>
<td>17</td>
</tr>
<tr>
<td><strong>allergy/atopy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Associated diseases</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chronic respiratory problems</td>
<td>yes/no</td>
<td>25</td>
</tr>
<tr>
<td>cardio-vascular diseases</td>
<td>yes/no</td>
<td>11</td>
</tr>
<tr>
<td>gastro-intestinal diseases</td>
<td>yes/no</td>
<td>25</td>
</tr>
<tr>
<td>biliary tract and liver diseases</td>
<td>yes/no</td>
<td>10</td>
</tr>
<tr>
<td>Parameter</td>
<td>Valuesa</td>
<td>Number of diseases</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>kidney and urogenital diseases</td>
<td>yes/no</td>
<td>36</td>
</tr>
<tr>
<td>infertility</td>
<td>yes/no</td>
<td>17</td>
</tr>
<tr>
<td><strong>Signs and symptoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>failure to thrive (child) or wasting (adult)</td>
<td>yes/no</td>
<td>36</td>
</tr>
<tr>
<td>chronic diarrhea</td>
<td>yes/no</td>
<td>38</td>
</tr>
<tr>
<td>delayed cord separation &lt; 2 weeks</td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>fever</td>
<td>yes/no</td>
<td>25</td>
</tr>
<tr>
<td>growth retardation</td>
<td>yes/no</td>
<td>45</td>
</tr>
<tr>
<td>pleural abnormalities</td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>gastrointestinal abnormalities</td>
<td>yes/no</td>
<td>29</td>
</tr>
<tr>
<td>hepatomegaly</td>
<td>yes/no</td>
<td>25</td>
</tr>
<tr>
<td>splenomegaly</td>
<td>yes/no</td>
<td>31</td>
</tr>
<tr>
<td>lymphadenopathy</td>
<td>yes/no</td>
<td>25</td>
</tr>
<tr>
<td>bleeding tendency</td>
<td>yes/no</td>
<td>35</td>
</tr>
<tr>
<td>skeletal or connective tissue abnormalities</td>
<td>yes/no</td>
<td>31</td>
</tr>
<tr>
<td><strong>facial abnormalities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- microcephaly</td>
<td>yes/no</td>
<td>21</td>
</tr>
<tr>
<td>- dental abnormalities/periodontitis</td>
<td>yes/no</td>
<td>10</td>
</tr>
<tr>
<td>- other facial anomalies/dysmorphism</td>
<td>yes/no</td>
<td>16</td>
</tr>
<tr>
<td><strong>skin defects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- photosensitivity</td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>- cutaneous telangiectasia</td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>- pigmentation defects</td>
<td>yes/no</td>
<td>14</td>
</tr>
<tr>
<td>- rash</td>
<td>yes/no</td>
<td>10</td>
</tr>
<tr>
<td>- eczema/atopic dermatitis</td>
<td>yes/no</td>
<td>17</td>
</tr>
<tr>
<td>- erythroderma</td>
<td>yes/no</td>
<td>7</td>
</tr>
<tr>
<td>- granuloma</td>
<td>yes/no</td>
<td>7</td>
</tr>
<tr>
<td>- other skin defects</td>
<td>yes/no</td>
<td>40</td>
</tr>
<tr>
<td>neurological or CNS abnormalities</td>
<td>yes/no</td>
<td>44</td>
</tr>
<tr>
<td>- mental retardation</td>
<td>yes/no</td>
<td>5</td>
</tr>
<tr>
<td>Parameter</td>
<td>Valuesa</td>
<td>Number of diseases</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>-ocular abnormalities</td>
<td>yes/no</td>
<td>7</td>
</tr>
<tr>
<td><strong>Laboratory features</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low red blood cell count (anemia)</td>
<td>yes/no</td>
<td>55</td>
</tr>
<tr>
<td>white blood cell (leukocytes) count</td>
<td>-1: decreased</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>21</td>
</tr>
<tr>
<td>low neutrophil count (&lt;500/ml)</td>
<td>yes/no</td>
<td>56</td>
</tr>
<tr>
<td>lymphocyte cell count affected</td>
<td>yes/no</td>
<td>82</td>
</tr>
<tr>
<td>thrombocytopenia</td>
<td>yes/no</td>
<td>28</td>
</tr>
<tr>
<td>polymorphonuclear cells affected</td>
<td>yes/no</td>
<td>32</td>
</tr>
<tr>
<td>monocytes/macrophages affected</td>
<td>yes/no</td>
<td>29</td>
</tr>
<tr>
<td>circulating T cells affected</td>
<td>-1: decreased</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>7</td>
</tr>
<tr>
<td>circulating B cells affected</td>
<td>-1: decreased:</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>7</td>
</tr>
<tr>
<td>NK cells affected</td>
<td>-1: decreased</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>1</td>
</tr>
<tr>
<td>IgG</td>
<td>-1: decreased</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>2</td>
</tr>
<tr>
<td>IgA</td>
<td>-1: decreased</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>5</td>
</tr>
<tr>
<td>IgM</td>
<td>-1: decreased</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>10</td>
</tr>
<tr>
<td>Parameter</td>
<td>Values(^a)</td>
<td>Number of diseases</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>IgE</td>
<td>-1: decreased</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>12</td>
</tr>
<tr>
<td>antibody response to booster immunization impaired</td>
<td>yes/no</td>
<td>19</td>
</tr>
<tr>
<td>low lymphocyte proliferation to mitogens</td>
<td>yes/no</td>
<td>20</td>
</tr>
<tr>
<td>CH50 abnormal (low or absent)</td>
<td>yes/no</td>
<td>21</td>
</tr>
<tr>
<td>AH50 abnormal (low or absent)</td>
<td>yes/no</td>
<td>12</td>
</tr>
<tr>
<td>low C4 or C3 concentrations</td>
<td>yes/no</td>
<td>7</td>
</tr>
<tr>
<td>spontaneous activation of the complement pathway</td>
<td>yes/no</td>
<td>19</td>
</tr>
<tr>
<td>radiation sensitivity/ chromosomal instability</td>
<td>yes/no</td>
<td>22</td>
</tr>
<tr>
<td>missing lymph nodes</td>
<td>yes/no</td>
<td>9</td>
</tr>
<tr>
<td>absence of thymus</td>
<td>yes/no</td>
<td>11</td>
</tr>
<tr>
<td>apoptosis</td>
<td>-1: decreased</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>0: normal/no</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>1: increased</td>
<td>13</td>
</tr>
<tr>
<td>chemotaxis</td>
<td>yes/no</td>
<td>8</td>
</tr>
<tr>
<td>killing (faulty O2 production)</td>
<td>yes/no</td>
<td>7</td>
</tr>
</tbody>
</table>

\(^a\) no – either the parameter is not related or data is missing
**Table S2.** Parameters statistically enriched in Disease Clusters. p-values are shown in brackets.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Diseases and affected genes</th>
<th>Parameters significantly overrepresented in the disease clusters</th>
<th>Treatments and functional categories significant in the clusters</th>
</tr>
</thead>
</table>
| I       | UNC13D STX11 RAB27A PRF1 FHL1 CN_ELA2 | other sites of infections (2.17e-03)   
ever (1.24e-06)   
neurological or CNS abnormalities (9.75e-04)  
gastrointestinal abnormalities (2.79e-03)  
hepatomegaly (7.21e-05)  
splenomegaly (1.96e-04)  
lymphadenopathy (7.21e-05)  
bleeding tendency (3.36e-04)  
low neutrophil count (<500/ml) (1.11e-04)  
thrombocytopenia (1.23e-04)  
lymphocyte cell count affected (6.02e-04) | BMT or HSCT (1.01e-02)  
immunomodulators (2.93e-02)  
cellular immunity (4.4e-02)  
nflammation (9.49e-03) |
| II      | G6PC3 DGCR ATM BLM | other sites of infection (3.26e-02)  
infections in upper respiratory tract (9.93e-03)  
infections in lower respiratory tract (1.02e-02)  
endocrine-AID (4.34e-03)  
B cell malignancies (4.54e-02)  
leukemia (3.84e-02)  
cardio-vascular diseases (1.44e-02)  
infertility (3.17e-02)  
growth retardation (2.18e-02)  
other facial anomalies/dysmorphism (2.85e-02)  
microcephaly (2.98e-03)  
photosensitivity (2.95e-03)  
cutaneous telangiectasia (2.95e-03)  
neurological or CNS abnormalities (2.08e-02)  
low lymphocyte proliferation to mitogens (4.19e-02)  
radiation sensitivity/chromosomal instability (4.9e-02)  
absence of thymus (1.44e-02) | - |
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Diseases and affected genes</th>
<th>Parameters significantly overrepresented in the disease clusters</th>
<th>Treatments and functional categories significant in the clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>III</td>
<td>PRKDC NFkBIA ADA COR01A RFXAP RFXANK RFX5 DNMT3B CIITA</td>
<td>infections in upper respiratory tract (2.60e-05) infections in lower respiratory tract (2.80e-05) skin infections (3.19e-04) susceptibility to bacteria, Gram (+) (<em>Staphylococcus aureus</em>) (3.03e-05) susceptibility to <em>Candida</em> (5.05e-06) failure to thrive or wasting (2.22e-08) chronic diarrhea (5.15e-05) gastro-intestinal diseases (1.12e-03) growth retardation (1.60e-04) neurological or CNS abnormalities (1.41e-04) lymphocyte cell count affected (1.31e-05) low lymphocyte proliferation to mitogens (3.57e-02)</td>
<td>immunoglobulin treatment (2.02e-05) BMT or HSCT (1.61e-05) antibiotics (5.36e-04) antifungals (2.32e-05) antivirals (8.31e-05) transcription factors (1.8e-04) humoral immunity (2.11e-02)</td>
</tr>
<tr>
<td>IV</td>
<td>STAT5B TNFRSF13B TMC8 TMC6 ORAI1 MPO LCK IGKC IGHE CMC CD79B CD3G AIRE CD3E CD247</td>
<td>infections in central nervous system (4.2e-02) infections in upper respiratory tract (2.61e-03) infections in lower respiratory tract (1.89e-02) susceptibility to <em>Candida</em> (1.9e-04) other malignancies (1.95e-02) low neutrophil count (&lt;500/ml) (1.85e-02)</td>
<td>immunoglobulin treatment (1.29e-02) antibiotics (1.45e-02) antifungals (8.61e-04) antivirals (6.96e-05)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Diseases and affected genes</td>
<td>Parameters significantly overrepresented in the disease clusters</td>
<td>Treatments and functional categories significant in the clusters</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>V</td>
<td>IRAK4 UNG TNFRSF13C Thymoma THI TAPBP TAP2 TAP1 ICOS HAX1 FCGR3A CXCR4 CD40 AICDA CSF3R CD19 BLNK IL12RB1 MYD88 IGLL1 IGHA2 IGHA1 IGAD1 CVID BTK SADNI IGHG3 IGHG2 IGHG1 LIG1 IGHG4 HIOM4 CFI C4B C4A C3 C2 C1S C1R C1QC C1QB C1QA CFD CD8A C4BPB C4BPA IGHM</td>
<td>infections in central nervous system (6.37e-03) infections in upper respiratory tract (2.4e-26) infections in lower respiratory tract (4.01e-26) gastro-intestinal infections (3e-02) susceptibility to mycobacteria (4.59e-03) susceptibility to Candida (9.07e-03) failure to thrive or wasting (1.20e-04) chronic diarrhea (5.26e-03) systemic lupus erythematosus (1.76e-03) rheumatoid arthritis (3.3e-06) idiopathic thrombocytopenia (3.90e-02) leukemia (6.87e-03) chronic respiratory problems (3.70e-02) cardio-vascular diseases (1.61e-02) infertility (1.13e-02) fever (1.26e-02) growth retardation (1.89e-04) other facial anomalies/dysmorphism (1.46e-02) microcephaly (4.19e-03) skin pigmentation defects (2.42e-02) neurological or CNS abnormalities (1.56e-03) gastrointestinal abnormalities (2.43e-02) hepatomegaly (1.58e-03) splenomegaly (3.75e-03) bleeding tendency (2.91e-02) skeletal or connective tissue abnormalities (3.81e-04) thrombocytopenia (7.72e-04) polymorphonuclear cells affected (3.02e-04) monocytes/macrophages affected (4.59e-03) CH50 abnormal (low or absent) (9.58e-04) low C4 or C3 concentrations (4.5e-02) spontaneous activation of the complement pathway (1.76e-03) radiation sensitivity/chromosomal instability (2.29e-02)</td>
<td>immunoglobulin treatment (2.75e-05) BMT or HSCT (1.64e-03) antibiotics (5.64e-16) antifungals (4.15e-02) immunomodulators (1.73e-11) complement system (8.87e-04) transcription factors (1.16e-02) humoral immunity (5.02e-05) cytokines (not chemokins) and receptors (4.69e-02) antigen processing and presenting (4.43e-02)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Diseases and affected genes</td>
<td>Parameters significantly overrepresented in the disease clusters</td>
<td>Treatments and functional categories significant in the clusters</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>VI</td>
<td>NRAS FPR1 CD59 XLA GHD G6PC FCGR1A CTSC SMARCAL1 ROBLD3 RNF168 NHEJ1 LIG4 TCN2 TCIRG1 TAZ STAT1 SBDS PSTPIP1 MYO5A MRE11A MLPH IL12B IFNGR2 G6PD ACTB DKC1 CEBPE AP3B1 CASP8 CASP10 CAPS XLN XIAP GF11 FASL CD55 LPIN2</td>
<td>infections in central nervous system (2.53e-04) infections in upper respiratory tract (4.12e-06) infections in lower respiratory tract (1.68e-07) skin infections (4.47e-02) gastro-intestinal infections (4.75e-02) susceptibility to bacteria, Gram (+) (<em>Staphylococcus aureus</em>) (1.14e-02) susceptibility to <em>Candida</em> (3.02e-03) chronic diarrhea (2.63e-02) systemic lupus erythematosus (3.66e-02) other malignancies (3.95e-02) growth retardation (3.83e-02) other facial anomalies/dysmorphism (3.1e-02) other skin defects (4.18e-02) eczema/atopic dermatitis (3.26e-02) splenomegaly (2.88e-02) low neutrophil count (&lt;500/ml) (1.17e-02) thrombocytopenia (2.73e-02) lymphocyte cell count affected (2.46e-02) monocytes/macrophages affected (2.44e-02) chemotaxis (3.29e-02) low lymphocyte proliferation to mitogens (1.84e-02) spontaneous activation of the complement pathway (2.22e-02)</td>
<td>BMT or HSCT (2.94e-02) antibiotics (3.57e-02) antifungals (3.84e-03) antivirals (1.35e-02) immunomodulators (2.09e-02) chemokins and receptors (3.88e-02) complement system (2.80e-02) humoral immunity (3.84e-03)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Diseases and affected genes</td>
<td>Parameters significantly overrepresented in the disease clusters</td>
<td>Treatments and functional categories significant in the clusters</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>VII</td>
<td>ZAP70 RAG2 RAG1 PTPRC JAK3 IL7R IL2RG IL2RA IKBKG FOXL1 CD3D</td>
<td>infections in upper respiratory tract (6e-05) infections in lower respiratory tract (6.48e-05) susceptibility to bacteria, Gram (+) (<em>Staphylococcus aureus</em>) (2.97e-05) susceptibility to mycobacteria (3.93e-11) susceptibility to <em>Candida</em> (4.03e-06) failure to thrive or wasting (3.21e-10) chronic diarrhea (4.44e-08) eczema/atopic dermatitis (6.17e-03) lymphocyte cell count affected (9.66e-07) low lymphocyte proliferation to mitogens (1.07e-02) absence of thymus (1.18e-03)</td>
<td>immunoglobulin treatment (1.67e-06) BMT or HSCT (3.58e-05) antibiotics (9.63e-05) antifungals (2.10e-08) antivirals (1.39e-07) humoral immunity (1.26e-02) cellular immunity (2.48e-04)</td>
</tr>
<tr>
<td>VIII</td>
<td>TNFRSF1A NOMID NOD2 MWS MVK MEFV FCAS</td>
<td>infections in central nervous system (2.07e-02) musculo-skeletal infections (4.99e-07) vasculitis (2.94e-04) kidney and urogenital diseases (1.15e-02) fever (1.10e-07) rash (4.56e-11) neurological or CNS abnormalities (1.95e-04) ocular abnormalities (5.74e-10) pleural abnormalities (8.5e-09) gastrointestinal abnormalities (1.86e-05) splenomegaly (7.15e-03) skeletal or connective tissue abnormalities (4.99e-07) polymorphonuclear cells affected (6.19e-07) monocytes/macrophages affected (2.26e-05)</td>
<td>antibiotics (1.65e-02) inflammation (1.43e-06)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Diseases and affected genes</td>
<td>Parameters significantly overrepresented in the disease clusters</td>
<td>Treatments and functional categories significant in the clusters</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>IX</td>
<td>CFP CFB C9 C8G C8B C8A C6 C5</td>
<td>infections in central nervous system (7.82e-07) infections in upper respiratory tract (4.4e-02) infections in lower respiratory tract (4.27e-02) susceptibility to bacteria, Gram (-) (Neisseria) (1.77e-11) systemic lupus erythematosus (4.72e-02) CH50 abnormal (low or absent) (1.02e-05) AH50 abnormal (low or absent) (7.06e-12) spontaneous activation of the complement pathway (5.52e-06)</td>
<td>immunoglobulin treatment (4.85e-02) antibiotics (1.34e-02) complement system (2.33e-08)</td>
</tr>
<tr>
<td>X</td>
<td>FANCM FANCL FANCI FANCG FANCF FANCE FANCD2 FANCC FANCB FANCA BRIP1 BRCA2</td>
<td>other sites of infections (4.63e-02) infections in upper respiratory tract (8.84e-03) infections in lower respiratory tract (8.47e-03) susceptibility to bacteria, Gram (+) (Staphylococcus aureus) (4.41e-02) leukemia (3.79e-15) kidney and urogenital diseases (3.66e-11) infertility (5.23e-16) growth retardation (6.75e-10) microcephaly (1.97e-14) skin pigmentation defects (9.18e-18) bleeding tendency (2.58e-11) skeletal or connective tissue abnormalities (5.41e-12) low neutrophil count (&lt;500/ml) (7.08e-09) thrombocytopenia (1.38e-12) lymphocyte cell count affected (1.31e-02) polymorphonuclear cells affected (8.2e-12) radiation sensitivity/chromosomal instability (4.1e-14)</td>
<td>immunoglobulin treatment (1.03e-02) BMT or HSCT (3.45e-07) antibiotics (4.06e-05) antifungals (4.93e-02) antivirals (2.77e-02) humoral immunity (4.93e-02)</td>
</tr>
<tr>
<td>Cluster</td>
<td>Diseases and affected genes</td>
<td>Parameters significantly overrepresented in the disease clusters</td>
<td>Treatments and functional categories significant in the clusters</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>XI</td>
<td>NCF2 NCF1 CYBB CYBA</td>
<td>other sites of infections (2.18e-03)</td>
<td>BMT or HSCT (8.08e-03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>infections in central nervous system (1.04e-03)</td>
<td>antibiotics (3.65e-02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>infections in upper respiratory tract (9.93e-03)</td>
<td>antifungals (2.02e-03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>infections in lower respiratory tract (1.02e-02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>skin infections (1.82e-03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gastro-intestinal infections (3.8e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>musculo-skeletal infections (3.03e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>susceptibility to bacteria, Gram (+) (Staphylococcus aureus)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.31e-03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>susceptibility to mycobacteria (2.68e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>susceptibility to Aspergillus (4.62e-06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>failure to thrive or wasting (5.19e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>chronic diarrhea (6.28e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>systemic lupus erythematosus (1.57e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>chronic respiratory problems (1.36e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gastro-intestinal diseases (1.36e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>kidney and urogenital diseases (5.19e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>growth retardation (1.21e-03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>eczema/atopic dermatitis (2.97e-05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>granuloma (5.3e-07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gastrointestinal abnormalities (2.37e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>hepatomegaly (1.36e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>splenomegaly (3.03e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>lymphadenopathy (1.36e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>low neutrophil count (~500/ml) (2.45e-03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>polymorphonuclear cells affected (3.4e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>monocytes/macrophages affected (2.68e-04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>killing (faulty O2 production) (5.3e-07)</td>
<td></td>
</tr>
</tbody>
</table>