Check for updates

DATA NOTE

REVISED A curated transcriptome dataset collection to investigate inborn errors of immunity [version 2; peer review: 2 approved]

Salim Bougarn, Sabri Boughorbel, Damien Chaussabel, Nico Marr

Systems Biology and Immunology, Sidra Medicine, Doha, Qatar

v2 **First published:** 15 Feb 2019, **8**:188 (https://doi.org/10.12688/f1000research.18048.1)

Latest published: 30 Aug 2019, **8**:188 (https://doi.org/10.12688/f1000research.18048.2)

Abstract

Primary immunodeficiencies (PIDs) are a heterogeneous group of inherited disorders, frequently caused by loss-of-function and less commonly by gain-of-function mutations, which can result in susceptibility to a broad or a very narrow range of infections but also in inflammatory, allergic or malignant diseases. Owing to the wide range in clinical manifestations and variability in penetrance and expressivity, there is an urgent need to better understand the underlying molecular, cellular and immunological phenotypes in PID patients in order to improve clinical diagnosis and management. Here we have compiled a manually curated collection of public transcriptome datasets mainly obtained from human whole blood, peripheral blood mononuclear cells (PBMCs) or fibroblasts of patients with PIDs and of control subjects for subsequent meta-analysis, query and interpretation. A total of eighteen (18) datasets derived from studies of PID patients were identified and retrieved from the NCBI Gene Expression Omnibus (GEO) database and loaded in GXB, a custom web application designed for interactive query and visualization of integrated large-scale data. The dataset collection includes samples from well characterized PID patients that were stimulated *ex vivo* under a variety of conditions to assess the molecular consequences of the underlying, naturally occurring gene defects on a genome-wide scale. Multiple sample groupings and rank lists were generated to facilitate comparisons of the transcriptional responses between different PID patients and control subjects. The GXB tool enables browsing of a single transcript across studies, thereby providing new perspectives on the role of a given molecule across biological systems and PID patients. This dataset collection is available at http://pid.gxbsidra.org/dm3/geneBrowser/list.

Keywords

Transcriptomics, microarray, primary immunodeficiency disorders, inborn errors of immunity.

This article is included in the Sidra Medicine gateway.

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
REVISED **version 2** published 30 Aug 2019	✓ report	✓ report
version 1 published 15 Feb 2019	? report	? report

1 **Bertrand De Meulder**, Université de Lyon, Lyon, France
Association EISBM, Vourles, France

2 **John B. Ziegler**, University of New South Wales, Sydney, Australia
Sydney Children's Hospital, Sydney, Australia

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Data: Use and Reuse collection.

Corresponding author: Salim Bougarn (sbougarn@sidra.org)

Author roles: Bougarn S: Conceptualization, Data Curation, Formal Analysis, Investigation, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Boughorbel S**: Data Curation, Resources, Software, Writing – Review & Editing; **Chaussabel D**: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing; **Marr N**: Conceptualization, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: All the authors listed on this publication received support from the Qatar Foundation. Support for this project was provided by the Qatar National Research Fund award NPRP10-0205-170348.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Bougarn S *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Bougarn S, Boughorbel S, Chaussabel D and Marr N. **A curated transcriptome dataset collection to investigate inborn errors of immunity [version 2; peer review: 2 approved]** F1000Research 2019, **8**:188 (https://doi.org/10.12688/f1000research.18048.2)

First published: 15 Feb 2019, **8**:188 (https://doi.org/10.12688/f1000research.18048.1)

REVISED **Amendments from Version 1**

In this new version, I took in consideration very helpful reviewers' comments and did some minor changes. The tutorials link on how to use the tools was updated. also, the GSE29536 datasets was removed from the instance and the datanote (Table 1). Finally, the Figure 1 was modified.

Any further responses from the reviewers can be found at the end of the article

Introduction

Primary immunodeficiencies (PIDs) are a heterogeneous group of inherited disorders, most often caused by loss-of-function mutations and less commonly by gain-of-function mutations, affecting components of the innate and/or adaptive immune system[1–3]. These inborn errors of immunity can result in profoundly increased susceptibility to a broad or a very narrow range of infections but also autoimmune disorders, allergies and malignancies[1,4–6]. The spectrum of clinical manifestations of PIDs is very broad and largely dependent upon the affected gene(s) and the degree to which normal gene function is lost or altered. In addition, a variety of other factors such as germline or somatic mosaicism, modifier genes and environmental factors can play an important role in the clinical penetrance and expressivity of a given disease phenotype[3,7,8]. To date, mutations in more than 300 genes have been identified to cause PIDs, which are classified into major groups reflecting the diverse immunological phenotypes[5,9]. Nonetheless, PIDs often go unrecognized or are not properly diagnosed[10]. However, with the recent developments and rapidly declining costs of next-generation sequencing technologies and other high-throughput methods, it is expected that many more as yet unknown disease-related genetic variants will be discovered in the near future[7].

A considerable challenge for identifying causal genetic variants—which is critical for the diagnosis and clinical management of PID patients—lies in the vast heterogeneity of the underlying immunological phenotypes and clinical manifestations on the one hand and in the degree of human genetic variation between individuals on the other hand. Despite considerable advances in recent years, specific gene functions in humans, their roles and regulation in biological processes, and essentiality of the redundancy in a function of a particular gene for protective immunity of the human host remains poorly understood[6]. In many cases, the use of forward and reverse genetics in mice or other model organisms has provided insufficient insights into the pathophysiology of PIDs, this due to interspecies differences and the fact that inbreeding has let to various deficiencies in laboratory animals, rendering them susceptible to a broad range of infections that often poorly recapitulates the clinical phenotypes in humans[11]. On the other hand, studying naturally occurring genetic defects in humans is much more challenging given the difficulty of obtaining biological samples, ethical implications and potential risks that go along with it. An additional challenge is the low frequency of most null alleles. Although PIDs are not necessarily rare when considered collectively, the small number of individuals that suffer from a specific deficiency usually does not permit classic case-control or family-based genetic association studies. Indeed, a considerable proportion of monogenic etiologies of PIDs were initially reported in single patients[12]. The ability to identify single-gene inborn errors in PID patients requires validation of the disease-causing variant by in-depth mechanistic studies demonstrating the structural and functional consequences of the mutations using blood or other accessible biological samples such as fibroblasts from skin biopsies[12]. In this context, several transcriptomics studies have been conducted using whole blood, PBMCs and fibroblasts of well-characterized PID patients, to assess the underlying immunological phenotypes at the molecular and cellular levels in more detail (Table 1) and in many cases, to further validate the causal relationship between the underlying genotypes and clinical phenotypes. Notable are several seminal studies of PID patients with susceptibility to a very narrow range of pathogens, such as patients with MYD88 or IRAK4 deficiency who are primarily susceptible to pyrogenic bacterial infections[13,14], patients with TBK1, TRIF or TLR3 deficiency[15–17] which underlies herpes simplex encephalitis of childhood, or a recent study of a child with IRF7 deficiency who was primarily susceptible to severe influenza but otherwise immunocompetent with regard to other common infectious diseases[18]. Such studies have highlighted that often, the underlying gene defect may only affect a narrow repertoire of transcriptional responses while the affected individual's cells remain highly responsive to specific stimulation through alternate receptors, pathways and signaling networks and in particular to *ex vivo* stimulation with whole organisms, reflecting the high degree of human gene redundancy in host defenses[6].

Here, we compiled a curated collection of 18 transcriptome datasets, retrieved from the NCBI's Gene Expression Omnibus (GEO) database, to provide as resource for the investigation on inborn errors of immunity. The datasets were loaded into a custom interactive web application, the Gene Expression Browser (GXB), (http://pid.gxbsidra.org/dm3/geneBrowser/list), which allows seamless access to the data and interactive visualization of the transcriptional responses, along with demographic and clinical information[19]. The user can customize data plots by adding multiple layers of parameters (e.g. age, gender, sample type and type of genetic defects), select and modify the sample ordering and gene rank lists, and generate links (mini URL) that can be shared via e-mail or used in publications. The GXB tool enables browsing of a single transcript across multiple studies and datasets, providing new perspectives on the role of a given molecule across biological systems and PID patients. In summary, this dataset collection can aid clinicians and researchers to study and quickly visualize the functional consequences of a variety of well characterized, naturally occurring mutations on a genome-wide scale.

Methods

A total of 163 datasets were identified in GEO using the following search query: Homo sapiens AND ("primary immunodeficiency diseases" OR "primary immunodeficiencies diseases" OR PID OR "autosomal recessive" OR "autosomal dominant" OR "inherited deficiency") AND ("Expression profiling by array"). All GEO entries that were returned with this query were manually curated. This process involved reading all the descriptions available for the datasets, the study designs and corresponding research articles. Finally, a total of 18 datasets were retained because they contained samples which were obtained from well

Table 1. List of datasets constituting the collection.

Title	Platforms used	PID classification	Disease	Genetic defects	Cell type/ Tissues	Number of samples	Citation #	GEO ID
Complete TLR3 deficiency. GSE30951.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	MyD88, TLR3, UNC93B1 deficiencies	Mutations of *MyD88*, *TLR3*, and *UNC93B1*	Fibroblast/ PBMC	44	17	GSE30951
Genome-wide profiling of whole blood from patients with defects in Toll-like receptors (TLRs) and IL-1Rs (the TIR pathway) signaling.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	IRAK4, MyD88 deficiencies	Mutations of *IRAK4*, and *MyD88*	Whole blood	365	14	GSE25742
Identification of IL-21-induced STAT3 dependent genes in human B cells.	Affymetrix HuGene 1.0 ST v1	Diseases of immune dysregulation	STAT3 GOF mutations	Mutations in *STAT3*	B cells	14	20	GSE51587
Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells.	Illumina HumanWG-6 v3	Defects in intrinsic and innate immunity	UNC93B1 deficiency	Mutations of *UNC93B1*	iPS	18	21	GSE40593
In vitro response of fibroblasts isolated from patients with immunodeficiencies.	Illumina Human-6 v2	Defects in intrinsic and innate immunity Combined immunodeficiencies with associated or syndromic features	UNC93B1, MyD88, IRAK4, STAT1, NEMO deficiencies	Mutations of *UNC93B1*, *MyD88*, *IRAK4*, *STAT1*, and *NEMO*	Fibroblast	72	13	GSE12124
In vitro response of fibroblasts isolated from patients with RBCK1 deficiency.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity Combined immunodeficiencies with associated or syndromic features	MyD88, HOIL1, NEMO deficiencies	Mutations of *MyD88*, *HOIL1/ RBCK1*, and *NEMO*	Fibroblast	46	22	GSE31064
In vitro responses of fibroblasts from patients with TBK1 deficiency after TLR3 dependent and independent stimuli.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	TLR3, TBK1, STAT1 deficiencies	Mutations of *TLR3*, *TBK1*, and *STAT1*	Fibroblast	58	15	GSE38652
In vitro responses of PBMC and fibroblasts from patients with TRIF deficiency after TRIF dependent and independent stimuli.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	TLR3, TRIF and MyD88 deficiencies	Mutations of *TLR3*, *TRIF*, and *MyD88*	Fibroblast, PBMC	27	16	GSE32390
Increased Wnt and Notch signaling: A clue to the renal disease in Schimke immuno-osseous dysplasia?.	Affymetrix HG-U133_Plus_2	Combined immunodeficiencies with associated or syndromic features	Schimke immuno-osseous dysplasia (SIOD)	Mutations in *SMARCAL1*	Kidney biopsy	5	23	GSE81156
Inherited human IRAK-1 deficiency selectively abolishes TLR signaling in fibroblasts.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	IRAK4, MyD88 deficiencies	Mutations of *IRAK1*, *IRAK4*, *MyD88*, and *MECP2*	Fibroblast	64	24	GSE92466

Title	Platforms used	PID classification	Disease	Genetic defects	Cell type/ Tissues	Number of samples	Citation #	GEO ID
Interferon Signature in the Blood in Inflammatory Common Variable Immune Deficiency [Test Set].	Illumina HumanHT-12 v4	Predominantly antibody deficiencies	Common Variable Immune Deficiency (CVID)	Unknown	Whole blood	53	25	GSE51404
Interferon Signature in the Blood in Inflammatory Common Variable Immune Deficiency [Training Set].	Illumina HumanHT-12 v3	Predominantly antibody deficiencies	Common Variable Immune Deficiency (CVID)	Unknown	Whole blood	83	25	GSE51405
IRAK-4- and MyD88-dependent pathways are essential for the removal of developing autoreactive B cells in humans.	Affymetrix HG-U133_Plus_2	Defects in intrinsic and innate immunity	UNC93B1, MyD88, and IRAK4 deficiencies	Mutations of *UNC93B1*, *MyD88*, and *IRAK4*	B cells	5	26	GSE13300
Response of IRF7-deficient peripheral blood mononuclear cells to pH1N1 influenza virus infection.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity	UNC93B1, and IRF7 deficiencies	Mutations of *UNC93B1*, and *IRF7*	PBMC	18	18	GSE66486
Screening for differentially expressed genes in patients with a novel immunodeficiency syndrome.	Affymetrix HG-U133A	Congenital defects of phagocyte number, functions, or both	p14/ LAMTOR2 deficiency	Mutations in *ROBLD3/ LAMTOR2*	B cells	4	27	GSE6322
Transcriptional analysis of whole blood in patients with auto-inflammatory disorders.	Illumina HumanHT-12 v3	Defects in intrinsic and innate immunity Autoinflammatory disorders	UNC93B1, MyD88, and IRAK4 deficiencies Defects affecting the inflammsome	Mutations of *UNC93B1*, *MyD88*, and *IRAK4, NLRP3, MVK*	Whole blood	51	22	GSE40561
Transcriptome analysis in peripheral blood mononuclear cells (PBMC) from HOIL-1-deficient patients upon TNF-a or IL-1b stimulation.	Illumina HumanHT-12 v4	Combined immunodeficiencies with associated or syndromic features	HOLI1 deficiency	Mutations of *HOLI1/RBCK1*	PBMC	45	22	GSE40838
Transcriptome analysis in primary fibroblasts from HOIL-1-deficient patients upon TNF-a or IL-1b stimulation.	Illumina HumanHT-12 v4	Defects in intrinsic and innate immunity Combined immunodeficiencies with associated or syndromic features	MyD88, HOIL1, NEMO deficiencies	Mutations of *MyD88, HOIL1/ RBCK1,* and *NEMO*	Fibroblast	54	22	GSE40560

F1000Research 2019, 8:188 Last updated: 17 SEP 2019

characterized patients with known PIDs (i.e. the genetic etiology had been identified) or the samples were obtained from patients which were considered to have common variable immunodeficiency (CVID). These include datasets that were generated from whole blood, PBMCs, fibroblasts, B cells, Induced Pluripotent Stem Cells (iPS) and kidney biopsy of individuals with defects in intrinsic and innate immunity, combined immunodeficiencies with associated or syndromic features, autoinflammatory disorders, congenital defects of phagocyte number, functions, or both, predominantly antibody deficiencies and diseases of immune dysregulation. However, In the future, this current instance might be updated by adding additional datasets. The selected datasets are listed in Table 1. A breakdown of the dataset collection by category in accordance to the most recently published update on PID classification from International Union of Immunological Societies Expert Committee[5,9] is shown in Figure 1.

The selected datasets were downloaded from GEO using the SOFT file format. Then, the datasets were uploaded onto our web tool, called the Gene Expression Browser (GXB), an interactive application hosted on the Amazon Web Services cloud[19]. Information about samples and study design were also uploaded. The available samples were assigned to groups based on the individuals and deficiencies studied and genes were ranked according to different group comparisons allowing the identification of transcripts that were differentially expressed between the patient's and control subject's cells cultured or stimulated *ex vivo* under the same conditions. Our dataset collection, uploaded in GXB, is available at http://pid.gxbsidra.org/dm3/geneBrowser/list. A web tutorial for the use of GXB can be accessed at: http://pid.gxbsidra.org/dm3/tutorials.gsp.

A detailed description of GXB has been recently published[19,28–30] and is reproduced here so that readers can use this article as a standalone resource. Briefly, datasets of interest can be quickly identified either by filtering on criteria from pre-defined sections on the left or by entering a query term in the search box at the top of the dataset navigation page. Clicking on one of the studies listed in the dataset navigation page opens a viewer designed to provide interactive browsing and graphic representations of large-scale data in an interpretable format. This interface is designed to present ranked gene lists and display expression results graphically in a context-rich environment. Selecting a gene from the rank ordered list on the left of the data-viewing interface will display its expression values graphically in the screen's central panel. Directly above the graphical display drop down menus give users the ability: a) To change how the gene list is ranked - this allows the user to change the method used to rank the genes, or to only include genes that are selected for specific biological interest; b) To change sample grouping (Group Set button) - in some datasets, a user can switch between groups based on cell type to groups based on disease type, for example; c) To sort individual samples within a group based on associated categorical or continuous variables (e.g. gender or age); d) To toggle between the bar chart view and a box plot view, with expression values represented as a single point for each sample. Samples are split into the same groups whether displayed as a bar chart or box plot; e) To provide a color legend for the sample groups; f) To select categorical information that is to be overlaid at the bottom of the graph - for example, the user can display gender or smoking status in this manner; g) To provide a color legend for the categorical information overlaid at the bottom of the graph; h) To download the graph as a portable network graphics (png) image. Measurements have no intrinsic utility in absence of contextual information. It is this contextual information that makes the results of a study or experiment interpretable. It is therefore important to capture, integrate and display information that will give users the ability to interpret data and gain new insights from it. We have organized this information under different tabs directly

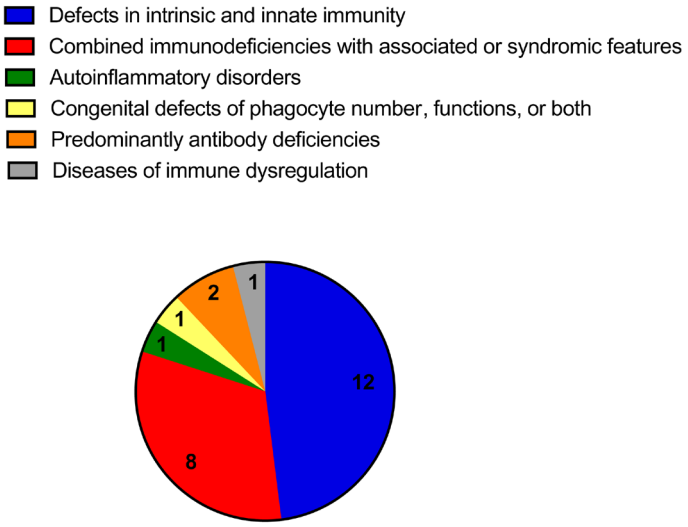

Figure 1. Break down of the dataset collection by category. The pie chart indicates the PIDs classification out for the 18 datasets.

above the graphical display. The tabs can be hidden to make more room for displaying the data plots, or revealed by clicking on the blue "show info panel" button on the top right corner of the display. Information about the gene selected from the list on the left side of the display is available under the "Gene" tab. Information about the study is available under the "Study" tab. Rolling the mouse cursor over a bar chart feature while displaying the "Sample" tab lists any clinical, demographic, or laboratory information available for the selected sample. Finally, the "Downloads" tab allows advanced users to retrieve the original dataset for analysis outside this tool. It also provides all available sample annotation data for use alongside the expression data in third party analysis software. Other functionalities are provided under the "Tools" drop-down menu located in the top right corner of the user interface. Some of the notable functionalities available through this menu include: a) Annotations, which provides access to all the ancillary information about the study, samples and dataset organized across different tabs; b) Cross-project view; which provides the ability for a given gene to browse through all available studies; c) Copy link, which generates a mini-URL encapsulating information about the display settings in use and that can be saved and shared with others (clicking on the envelope icon on the toolbar inserts the URL in an email message via the local email client); d) Chart options; which gives user the option to customize chart labels.

Data availability

All datasets included in our curated collection are available publicly via the NCBI GEO website: https://www.ncbi.nlm.nih.gov/gds/ and are referenced throughout the manuscript by their GEO accession numbers (e.g. GSE92466). Signal files and sample description files can also be downloaded from the GXB tool under the "downloads" tab.

Acknowledgments

We would like to thank all the investigators who decided to make their datasets publicly available by depositing them in GEO.

References

1. Seleman M, Hoyos-Bachiloglu R, Geha RS, *et al.*: **Uses of Next-Generation Sequencing Technologies for the Diagnosis of Primary Immunodeficiencies.** *Front Immunol.* 2017; **8**: 847.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Casanova JL: **Human genetic basis of interindividual variability in the course of infection.** *Proc Natl Acad Sci U S A.* 2015; **112**(51): E7118–27.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Casanova JL: **Severe infectious diseases of childhood as monogenic inborn errors of immunity.** *Proc Natl Acad Sci U S A.* 2015; **112**(51): E7128–37.
PubMed Abstract | Publisher Full Text | Free Full Text
4. Allenspach E, Torgerson TR: **Autoimmunity and Primary Immunodeficiency Disorders.** *J Clin Immunol.* 2016; **36**(Suppl 1): 57–67.
PubMed Abstract | Publisher Full Text
5. Picard C, Al-Herz W, Bousfiha A, *et al.*: **Primary Immunodeficiency Diseases: an Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015.** *J Clin Immunol.* 2015; **35**(8): 696–726.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Casanova JL, Abel L: **Human genetics of infectious diseases: Unique insights into immunological redundancy.** *Semin Immunol.* 2018; **36**: 1–12.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Stray-Pedersen A, Sorte HS, Samarakoon P, *et al.*: **Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders.** *J Allergy Clin Immunol.* 2017; **139**(1): 232–245.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Wolach O, Kuijpers T, Ben-Ari J, *et al.*: **Variable clinical expressivity of STAT3 mutation in hyperimmunoglobulin E syndrome: genetic and clinical studies of six patients.** *J Clin Immunol.* 2014; **34**(2): 163–70.
PubMed Abstract | Publisher Full Text
9. Bousfiha A, Jeddane L, Picard C, *et al.*: **The 2017 IUIS Phenotypic Classification for Primary Immunodeficiencies.** *J Clin Immunol.* 2018; **38**(1): 129–143.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Berger M, Geng B, Cameron DW, *et al.*: **Primary immune deficiency diseases as unrecognized causes of chronic respiratory disease.** *Respir Med.* 2017; **132**: 181–188.
PubMed Abstract | Publisher Full Text
11. Ciancanelli MJ, Abel L, Zhang SY, *et al.*: **Host genetics of severe influenza: from mouse Mx1 to human IRF7.** *Curr Opin Immunol.* 2016; **38**: 109–20.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Casanova JL, Conley ME, Seligman SJ, *et al.*: **Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies.** *J Exp Med.* 2014; **211**(11): 2137–49.
PubMed Abstract | Publisher Full Text | Free Full Text
13. von Bernuth H, Picard C, Jin Z, *et al.*: **Pyogenic bacterial infections in humans with MyD88 deficiency.** *Science.* 2008; **321**(5889): 691–696.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Alsina L, Israelsson E, Altman MC, *et al.*: **A narrow repertoire of transcriptional modules responsive to pyogenic bacteria is impaired in patients carrying loss-of-function mutations in *MYD88* or *IRAK4*.** *Nat Immunol.* 2014; **15**(12): 1134–42.
PubMed Abstract | Publisher Full Text | Free Full Text
15. Herman M, Ciancanelli M, Ou YH, *et al.*: **Heterozygous *TBK1* mutations impair TLR3 immunity and underlie herpes simplex encephalitis of childhood.** *J Exp Med.* 2012; **209**(9): 1567–1582.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Sancho-Shimizu V, Pérez de Diego R, Lorenzo L, *et al.*: **Herpes simplex encephalitis in children with autosomal recessive and dominant TRIF deficiency.** *J Clin Invest.* 2011; **121**(12): 4889–902.
PubMed Abstract | Publisher Full Text | Free Full Text
17. Guo YQ, Audry M, Ciancanelli M, *et al.*: **Herpes simplex virus encephalitis in a patient with complete TLR3 deficiency: TLR3 is otherwise redundant in protective immunity.** *J Exp Med.* 2011; **208**(10): 2083–2098.
PubMed Abstract | Publisher Full Text | Free Full Text
18. Ciancanelli MJ, Huang SX, Luthra P, *et al.*: **Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency.** *Science.* 2015; **348**(6233): 448–53.
PubMed Abstract | Publisher Full Text | Free Full Text
19. Speake C, Presnell S, Domico K, *et al.*: **An interactive web application for the dissemination of human systems immunology data.** *J Transl Med.* 2015; **13**: 196.
PubMed Abstract | Publisher Full Text | Free Full Text
20. Berglund LJ, Avery DT, Ma CS, *et al.*: **IL-21 signalling via STAT3 primes human naive B cells to respond to IL-2 to enhance their differentiation into plasmablasts.** *Blood.* 2013; **122**(24): 3940–50.
PubMed Abstract | Publisher Full Text | Free Full Text
21. Lafaille FG, Pessach IM, Zhang SY, *et al.*: **Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells.** *Nature.* 2012; **491**(7426): 769–73.
PubMed Abstract | Publisher Full Text | Free Full Text
22. Boisson B, Laplantine E, Prando C, *et al.*: **Immunodeficiency, autoinflammation and amylopectinosis in humans with inherited HOIL-1 and LUBAC deficiency.** *Nat Immunol.* 2012; **13**(12): 1178–86.
PubMed Abstract | Publisher Full Text | Free Full Text
23. Morimoto M, Myung C, Beirnes K, *et al.*: **Increased Wnt and Notch signaling: a clue to the renal disease in Schimke immuno-osseous dysplasia?** *Orphanet J Rare Dis.* 2016; **11**(1): 149.
PubMed Abstract | Publisher Full Text | Free Full Text
24. Della Mina E, Borghesi A, Zhou H, *et al.*: **Inherited human IRAK-1 deficiency selectively impairs TLR signaling in fibroblasts.** *Proc Natl Acad Sci U S A.* 2017; **114**(4): E514–E523.
PubMed Abstract | Publisher Full Text | Free Full Text
25. Park J, Munagala I, Xu H, *et al.*: **Interferon Signature in the Blood in Inflammatory Common Variable Immune Deficiency.** *PLoS One.* 2013; **8**(9): e74893.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Isnardi I, Ng YS, Srdanovic I, *et al.*: **IRAK-4- and MyD88-dependent pathways are essential for the removal of developing autoreactive B cells in humans.** *Immunity.* 2008; **29**(5): 746–57.
PubMed Abstract | Publisher Full Text | Free Full Text

27. Bohn G, Allroth A, Brandes G, *et al.*: **A novel human primary immunodeficiency syndrome caused by deficiency of the endosomal adaptor protein p14.** *Nat Med.* 2007; **13**(1): 38–45.
PubMed Abstract | Publisher Full Text

28. Rahman M, Boughorbel S, Presnell S, *et al.*: **A curated transcriptome dataset collection to investigate the functional programming of human hematopoietic cells in early life [version 1; referees: 2 approved].** *F1000Res.* 2016; **5**: 414.
PubMed Abstract | Publisher Full Text | Free Full Text

29. Marr AK, Boughorbel S, Presnell S, *et al.*: **A curated transcriptome dataset collection to investigate the development and differentiation of the human placenta and its associated pathologies [version 2; referees: 2 approved]**. *F1000Res.* 2016; **5**: 305.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Rinchai D, Boughorbel S, Presnell S, *et al.*: **A curated compendium of monocyte transcriptome datasets of relevance to human monocyte immunobiology research [version 2; referees: 2 approved].** *F1000Res.* 2016; **5**: 291.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 17 September 2019

https://doi.org/10.5256/f1000research.22465.r53217

© 2019 Ziegler J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **John B. Ziegler**
[1] School of Women's & Children's Health, University of New South Wales, Sydney, NSW, Australia
[2] Department of Immunology & infectious Diseases, Sydney Children's Hospital, Sydney, Australia

All is now in order with significant improvements in response to peer review.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical immunology including immunogenetics of PIDs.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 09 September 2019

https://doi.org/10.5256/f1000research.22465.r53218

© 2019 De Meulder B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Bertrand De Meulder**
[1] European Institute for Systems Biology & Medicine, Université de Lyon, Lyon, France
[2] Association EISBM, Vourles, France

I feel my comments have been taken into account and hence approve the indexing of this manuscript.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, transcriptomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 28 March 2019

https://doi.org/10.5256/f1000research.19737.r45485

© 2019 Ziegler J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

John B. Ziegler
[1] School of Women's & Children's Health, University of New South Wales, Sydney, NSW, Australia
[2] Department of Immunology & infectious Diseases, Sydney Children's Hospital, Sydney, Australia

This is a useful tool to examine publicly available gene transcript data.

The paper would be more useful if it included more detailed instructions for its use, perhaps including screenshots to illustrate the text.

The location of the PubMed link of each sample wasn't immediately obvious to this reader. I now realise it is accessed by clicking a part of the PubMed.gov icon.

The meaning of the term "ranking" becomes apparent with use but it would not be burdensome to more experienced readers to have it defined or explained.

The acronym for induced pluripotent stem cells (iPS) should be defined with first use.

The sample set "Whole Blood Transcriptional Modules generated on Illumina Hu-6 V2 Beadchips. GSE29536" is listed with the disease category "Immunodeficiencies" but the data presented appears to include none from immunodeficiency settings. I see transcripts from infections, SLE, diabetes, Still's disease. I can't get more detail of the study because the PubMed link is wrong. The link provided is https://www.ncbi.nlm.nih.gov/pubmed/24069364 which is actually the link for other listed samples (GSE51404 and GSE51405). I haven't verified all links in the table.

If improvements such as those above were addressed I believe the paper would be very useful to those conducting immunological research.

If improvements such as those above were addressed I believe the paper would be very useful to those conducting immunological research.

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of methods and materials provided to allow replication by others?
No

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical immunology including immunogenetics of PIDs.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 27 Aug 2019

Salim Bougarn, Sidra Medicine, Doha, Qatar

Dear Dr. Ziegler,
We are thankful to Dr. Ziegler for his positive feed back regarding our manuscript and for the careful revision. We respond to the specific comments and describe the introduced changes to the new version below in the text and in bold.

The paper would be more useful if it included more detailed instructions for its use, perhaps including screenshots to illustrate the text.

We have updated a link to the tutorial (http://pid.gxbsidra.org/dm3/tutorials.gsp) as well as a reference to a paper describing [Speake et. al) on how to use GXB with detailed instructions including screenshots.

The location of the PubMed link of each sample wasn't immediately obvious to this reader. I now realise it is accessed by clicking a part of the PubMed.gov icon.

The meaning of the term "ranking" becomes apparent with use but it would not be burdensome to more experienced readers to have it defined or explained.

The ranking was explained in material and methods section in the paper

The acronym for induced pluripotent stem cells (iPS) should be defined with first use.

We have added to the paper the definition of iPS.

The sample set "Whole Blood Transcriptional Modules generated on Illumina Hu-6 V2 Beadchips. GSE29536" is listed with the disease category "Immunodeficiencies" but the data presented appears to include none from immunodeficiency settings. I see transcripts from infections, SLE, diabetes, Still's disease. I can't get more detail of the study because the PubMed link is wrong. The link provided is https://www.ncbi.nlm.nih.gov/pubmed/24069364 which is actually the link for other listed samples (GSE51404 and GSE51405). I haven't verified all links in the table.

We have removed the GSE29536 from the instance and the datanote and modified the total number of datasets retrieved and loaded into GXB. But also, we modified the table 1 and figure 1.

We have checked and verified all PubMed links in other datasets.

If improvements such as those above were addressed I believe the paper would be very useful to those conducting immunological research.
The requested improvements have been incorporated in the paper.

Please note that other minor modification updates were done on the manuscript, such as the total number of datasets found in the GEO database using the criteria mentioned in the paper.

Competing Interests: No competing interests were disclosed.

Reviewer Report 21 March 2019

https://doi.org/10.5256/f1000research.19737.r45968

© 2019 De Meulder B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Bertrand De Meulder
[1] European Institute for Systems Biology & Medicine, Université de Lyon, Lyon, France
[2] Association EISBM, Vourles, France

This article presents a tool to gather and make preliminary analyses on a set of curated sequencing datasets related to primary immunodeficiencies.

Regarding replication, I could not reproduce the list of datasets with the criteria that are mentioned in the text. I would have to believe that the authors are indeed reporting their results, and that the manual curation of the dataset list was accurate.

The tool that is presented has a number of issues:

- The list of datasets is rather small, and the authors do not mention whether they plan on adding datasets to the list,
- The tutorials on how to use the tools are only available to registered users, and I could not find how to register
- Graphical issues are present when displaying the charts, both in-tool and in png images.

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?

Yes

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, transcriptomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 27 Aug 2019

Salim Bougarn, Sidra Medicine, Doha, Qatar

Dear Bertrand,

We are thankful to Dr. De Meulder for his positive feed back regarding our manuscript and for the careful revision. We respond to the specific comments and describe the introduced changes to the new version below in the text and in bold.

Regarding replication, I could not reproduce the list of datasets with the criteria that are mentioned in the text. I would have to believe that the authors are indeed reporting their results, and that the manual curation of the dataset list was accurate.
This is not related to GXB but on the pubmed search criteria. Also, please note that the GEO database is growing over time and for your attention the query was run in date 2017-2-1. I have updated in the paper the total number of datasets found in the GEO database using the criteria mentioned in the paper.

The tool that is presented has a number of issues:

- The list of datasets is rather small, and the authors do not mention whether they plan on adding datasets to the list,

The curation of the datasets requires substantial effort so we decide to freeze the addition of dataset for this version. We are planning to update the datasets in a future version of the paper. A sentence was added to the material and method in the paper.

- The tutorials on how to use the tools are only available to registered users, and I could not find how to register

The tutorial is available also for non registered user via the following link: http://pid.gxbsidra.org/dm3/tutorials.gsp. The link in the paper was modified.

- Graphical issues are present when displaying the charts, both in-tool and in png images.

It is not clear what graphical issue exactly you encounter. It is possible to hide info panel information to resize the graphs and get a better visualization.

Please note that the GSE29536 datasets was removed from the instance and the datanote

(Table 1). The figure 1 was slightly modified. Also, the total number of datasets retrieved, selected and loaded to GXB was modified.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research