

## A Non-targeted Proteomics Newborn Screening Platform for Genetic Disorders

Hirofumi Shibata<sup>1#</sup>, Daisuke Nakajima<sup>2#</sup>, Ryo Konno<sup>2</sup>, Atsuhiko Hijikata<sup>3</sup>, Motoko Higashiguchi<sup>1</sup>, Hiroshi Nihira<sup>1</sup>, Saeko Shimodera<sup>1</sup>, Takayuki Miyamoto<sup>1</sup>, Masahiko Nishitani-Isa<sup>1</sup>, Eitaro Hiejima<sup>1</sup>, Kazushi Izawa<sup>1</sup>, Junko Takita<sup>1</sup>, Toshio Heike<sup>1</sup>, Ken Okamura<sup>4</sup>, Hidenori Ohnishi<sup>5</sup>, Masataka Ishimura<sup>6</sup>, Satoshi Okada<sup>7</sup>, Motoi Yamashita<sup>8</sup>, Tomohiro Morio<sup>8</sup>, Hirokazu Kanegane<sup>9</sup>, Kohsuke Imai<sup>10</sup>, Yasuko Nakamura<sup>10</sup>, Shigeaki Nonoyama<sup>10</sup>, Toru Uchiyama<sup>11</sup>, Masafumi Onodera<sup>12</sup>, Ryuta Nishikomori<sup>13</sup>, Osamu Ohara<sup>2</sup>, Yusuke Kawashima<sup>2\*</sup>, and Takahiro Yasumi<sup>1\*</sup>

<sup>1</sup>Department of Pediatrics, Kyoto University Graduate School of Medicine, Kyoto, Japan

<sup>2</sup>Department of Applied Genomics, Kazusa DNA Research Institute, Kisarazu, Japan

<sup>3</sup>School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, Tokyo, Japan

<sup>4</sup>Department of Dermatology, Faculty of Medicine, Yamagata University, Yamagata, Japan

<sup>5</sup>Department of Pediatrics, Gifu University Graduate School of Medicine, Gifu, Japan

<sup>6</sup>Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

<sup>7</sup>Department of Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Hiroshima, Japan

<sup>8</sup>Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University (TMDU), Tokyo, Japan

<sup>9</sup>Department of Child Health and Development, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University (TMDU), Tokyo, Japan

<sup>10</sup>Department of Pediatrics, National Defense Medical College, Tokorozawa, Japan

<sup>11</sup>Department of Human Genetics, National Center for Child Health and Development, Tokyo, Japan

<sup>12</sup>Gene & Cell Therapy Promotion Center, National Center for Child Health and Development, Tokyo, Japan

<sup>13</sup>Department of Pediatrics and Child Health, Kurume University School of Medicine, Kurume, Japan

<sup>#</sup>These authors contributed equally to this work.

<sup>\*</sup>Corresponding authors: Takahiro Yasumi and Yusuke Kawashima

Takahiro Yasumi (ORCID ID: 0000-0003-3257-5071)

Department of Pediatrics, Kyoto University Graduate School of Medicine

54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto 606-8507, JAPAN

Phone: (+81) 75-751-3291, Fax: (+81) 75-752-2361, E-mail: yasumi@kuhp.kyoto-u.ac.jp

Yusuke Kawashima (ORCID ID: 0000-0002-9779-8199)

Department of Applied Genomics, Kazusa DNA Research Institute

2-5-23 Kazusa-Kamatari, Kisarazu, 292-0818, JAPAN

Phone: (+81) 438-52-3580, Fax: (+81) 438-52-3501, E-mail: ykawashi@kazusa.or.jp

The authors have declared that no conflict of interest exists.

### Summary:

Non-targeted proteomics analysis of newborn dried blood spots can evaluate protein levels encoded by >1000 genes in the Online Mendelian Inheritance in Man database, identify patients with inborn errors of immunity, and provide a novel newborn disease screening platform.

## **Abstract**

Newborn screening using dried blood spot (DBS) samples has made a substantial contribution to public healthcare by detecting patients with genetic disorders as neonates. Targeted measurements of nucleic acids and metabolites have played major roles in newborn screening to date, while the feasibility of new non-targeted approaches, including genome-wide DNA sequencing, has been explored. Here, we investigated the applicability of non-targeted quantitative proteomics analysis to newborn screening for genetic diseases. DBS protein profiling allowed monitoring of levels of proteins encoded by 2912 genes, including 1106 listed in the Online Mendelian Inheritance in Man database, in healthy newborn samples, and was useful in identifying patients with inborn errors of immunity by detecting reduced levels of disease causative proteins and cell-phenotypical alterations. Our results indicate that application of non-targeted quantitative protein profiling of DBS samples can forge a new path in screening for genetic disorders.

## Introduction

Dried blood spot (DBS) samples, comprising spotted blood dried in filter paper, are widely used for pre-symptomatic disease screening because of their simplicity in terms of sampling, shipping, and storage.

Since Dr. Guthrie introduced the first screening test for phenylketonuria (GUTHRIE and SUSI, 1963), newborn screening (NBS) programs using DBS samples have made considerable contributions to public healthcare. NBS can save neonates with serious inborn disorders from developing devastating symptoms, by introducing appropriate therapeutic intervention before disease onset, which can improve their long-term prognosis. Notably, the emergence of tandem mass spectrometry has enabled use of a single platform to simultaneously screen for a wide range of inborn errors of metabolism (IEMs), playing a critical role in the spread of NBS (Wilcken et al., 2003), and serving as a successful example of the power of non-targeted measurement platforms for pre-symptomatic screening for inherited diseases, followed by confirmatory diagnosis by genetic analyses.

Although NBS for IEMs can be achieved by assessing metabolic function through monitoring activity of metabolites or metabolic enzymes, direct measurement of molecular function is frequently too complex and infeasible in the context of NBS for other inherited conditions. As recent advances in DNA sequencing technology have dramatically improved cost-performance and throughput, the use of genome-wide sequencing has been proposed as a new alternative modality for NBS (Berg et al., 2017); however, there remain many technical, clinical, ethical, and social issues to be addressed before implementation of genome

sequencing for NBS (Spiekerkoetter et al., 2023; Ulph and Bennett, 2023). In addition, concerns regarding the use of genome sequencing persist, due to an inability to predict gene product functionality from structural genetic information with high accuracy; in the absence of accumulated in-depth genotype-phenotype data, gene structure information cannot be linked with clinical phenotypes. Indeed, Adhikari et al. reported that exome sequencing was not sufficiently sensitive or specific as a primary screen for use in NBS for IEM, while it is useful as a second-tier test for some infants who test positive in first-tier screening, assisted by tandem mass spectrometry (Adhikari et al., 2020).

As the majority of inherited diseases occur due to loss or gain of protein function, we hypothesized that measurement of the quality and/or quantity of proteins encoded by causative genes could represent a new modality for NBS with advantages over structural gene analyses. We recently reported that proteome analysis of peripheral blood mononuclear cells can complement genetic analysis in identifying the causes of inborn errors of immunity (IEIs) in undiagnosed patients (Sakura et al., 2023). Here, we explore whether proteome analysis of DBS samples can contribute to the diagnosis/screening of IEIs and other genetic disorders. Targeted quantitative measurement of proteins of interest in DBS have been reported as test trials (Collins et al., 2020; Dezfouli et al., 2020); however, until recently, cost and throughput issues have not been surmounted. Recent advances in proteome technology have allowed us to address these issues; for example, advanced liquid chromatography-assisted mass spectrometry (LC-MS/MS) in data-independent acquisition (DIA) mode facilitates highly sensitive quantification of protein levels (Kawashima et al., 2022; Kawashima

et al., 2019). Furthermore, a non-targeted method for in-depth DBS protein profiling has also been developed (Nakajima et al., 2020). These technical advances have allowed us to overcome the obstacles impeding progress in proteome analysis-based NBS.

In this context, the aims of the present study were to evaluate whether non-targeted proteomics analysis of newborn DBS by LC-MS/MS could be used to identify and quantify disease-causing or -related proteins, thereby clarifying the molecular phenotypes underlying genetic disorders at a pre-symptomatic stage. For this purpose, DBS samples from 40 healthy newborns were first subjected to non-targeted proteomics analysis, to assess their consistency and reliability for use in protein profiling by comparison with similar data generated from healthy adult samples. Subsequently, DBS samples from patients with IEIs, as a representative disease spectrum primarily caused by defects in genes expressed in blood, were analyzed to determine whether patients with these conditions can be identified by quantitative protein profiling. Our results indicate that non-targeted DBS proteomics can sensitively detect changes in molecular and cellular phenotypes underlying many IEIs, indicating that this approach may be useful for newborn diagnosis and/or screening for genetic diseases.

## **Results**

### **Comparison of protein profiles from healthy newborn and healthy adult samples**

LC-MS/MS and non-targeted deep proteome analyses of DBS samples from 40 healthy newborns enabled

the identification and quantification of 3587 proteins within the identification threshold (Table S1, <https://doi.org/10.5281/zenodo.10539971>). These data were filtered to include only proteins with  $\geq 2$  identified peptides, resulting in a total of 2912 proteins that could be quantified within measurable ranges in all newborn samples. Of these, 1106 proteins encoded by genes listed in the Online Mendelian Inheritance in Man (OMIM) database were subjected to analysis after standardization based on  $\beta$ -actin levels in the same sample (Table S2, <https://doi.org/10.5281/zenodo.10539971>). We also analyzed samples from eight healthy adults. Correlation analysis indicated that samples in the newborn and adult DBS groups showed high within-group coefficient values (Pearson  $r = 0.849$ – $0.973$  and  $0.898$ – $0.962$ , respectively). Next, we conducted multiple evaluation of Pearson's correlation coefficient values between the two sample types, along with principal component analysis, and the results showed that newborn samples clustered separately from adult samples (Fig. 1, A and B). Detailed analysis of protein levels to search for differentially expressed proteins (DEPs) identified 954 upregulated and 108 downregulated proteins in newborn, relative to adult, samples (Fig. 2 A). Upregulated proteins included components of fetal and embryonic hemoglobin (*HBG1/HBG2* and *HBE1/HBZ*, respectively) and alfa fetoprotein (*AFP*). Downregulated proteins included antithrombin III (*SERPINC1*), heparin cofactor 2 (*SERPIND1*), protein C (*PROC*), haptoglobin (*HP*), Hb subunit delta (*HBD*), immunoglobulin heavy constant mu (*IGHM*), and joining chain of multimeric IgA and IgM (*JCHAIN*) (Fig. 2, B and C, and Fig. S1). In Gene Ontology (GO) analysis of DEPs, translational initiation and regulation of complement activation pathways were ranked as highly enriched for proteins

upregulated and downregulated in newborn relative to adult samples, respectively (Fig. 2 D and E). These results are generally consistent with previous reports (Kanakoudi et al., 1995; Reverdiau-Moalic et al., 1996) and collectively indicate the reliability of the current method for non-targeted quantitative evaluation of proteins in DBS samples.

### **Proteins causative of various genetic disorders can be quantified in neonatal DBS samples**

Next, we considered which diseases are feasible potential targets for diagnosis/screening by DIA-LC-MS/MS using DBS samples. Diseases with autosomal recessive or X-linked inheritance were identified as primary targets; as diseases with autosomal dominant inheritance are caused by anomaly of a single allele, while the other allele remains intact, which would be expected to minimally affect protein levels, they were not considered ideal for screening using this approach. Of the 1110 identified proteins listed in OMIM, 781 were related to diseases with autosomal recessive or X-linked inheritance, of which, 432, 71, and 139 proteins were categorized as metabolic/neurological, hematological, and immunological diseases, respectively (Table S2 and Fig. S2 A). Of the 139 proteins related to immunological diseases, 124 were included as causative molecules in the latest list of IEIs (Table S2) and most of these were distributed similarly in healthy newborn and adult samples (Fig. S2 B). To explore the possibility of direct detection of changes in protein levels caused by genetic mutations and their use for diagnosis/screening of genetic disorders, we next focused our efforts on characterization of DBS protein levels in samples from patients

with IEIs (Table 1), as most IEI causative proteins are abundantly expressed in peripheral blood.

### **Primary effect of IEI gene mutations on levels of their encoded proteins**

First, we evaluated whether levels of IEI disease-causing proteins were exclusively reduced in patient samples and suitable for use in disease diagnosis. All 10 samples from patients with familial hemophagocytic lymphohistiocytosis (FHL) type 3, including five obtained before and five obtained after the onset of disease, showed significant reductions in Munc13-4 levels compared to control samples (Fig. 3 A and Fig. S3). This finding is consistent with reports showing that FHL type 3 can be reliably diagnosed by assessing intra-platelet Munc13-4 levels (Murata et al., 2011; Shibata et al., 2018). A similar result was observed in a patient with FHL type 5, in that Munc18-2 levels were markedly lower in the patient than in control samples (Fig. 3 B). By contrast, perforin levels did not differ significantly between patients with FHL type 2 and controls (Fig. 3 D), although flow cytometry showed that perforin levels were reduced in natural killer cells from patients (Fig. S4). Analysis of protein profiles in patients with other IEIs showed that levels of p91-phox, Wiskott-Aldrich syndrome protein (WASP), and Bruton's tyrosine kinase (BTK) in samples from patients with p91-phox deficiency, Wiskott-Aldrich syndrome (WAS), and X-linked agammaglobulinemia (XLA), respectively, were markedly lower than those in control samples (Fig. 3, E, H, and I). A sample from a patient with p22-phox deficiency also showed markedly reduced levels of its causative protein (Fig. 3 F). Interestingly, p22-phox levels paralleled those of p91-phox in this patient (Fig. 3, E and F), likely reflecting



the tight interaction between these molecules (Parkos et al., 1989; Porter et al., 1994). A similar effect was observed in an FHL type 5 sample, in that Syntaxin-11 levels were reduced in parallel with those of Munc18-2 (Fig. 3, B and C). Levels of the causative protein were also lower in sample from a patient with p47-phox deficiency than those in healthy controls; however, they were only slightly below the normal range (Fig. 3 G). Samples from patients with Hermansky-Pudlak syndrome (HPS) type 2 also had reduced p47-phox protein levels (Fig. 3 G). By contrast, levels of lysosomal trafficking regulator (LYST), adaptor related protein complex 3 subunit beta 1 (AP3B1), and adenosine deaminase (ADA) in patients with Chédiak-Higashi syndrome (CHS), HPS type 2, and ADA deficiency, respectively, did not differ significantly from those in control samples (Fig. 3, J–L). Levels of common  $\gamma$  chain (IL2RG) were below the limit of detection in every sample. Absolute expression levels of the disease-causing protein in every patient, together with those in healthy controls, are summarized in Table 1.

We next determined whether the effect of a gene mutation on its protein level can be reliably predicted by *in silico* analysis. Table 2 shows the results of mutation effect predictions, together with the observed protein levels. As missense mutations often affect protein abundance by destabilizing protein structures (Yue et al., 2005), we performed computational prediction of the impact of missense mutations on protein stability using FoldX, a state-of-the art method for stability estimation based on the three-dimensional protein structures (Delgado et al., 2019). Of the 14 missense mutations identified in patients, 10 led to reduced protein abundance. Of those mutations leading to protein loss, 8 (80%) were

successfully predicted to be destabilized, based on a total energy change threshold of  $> 1.58$  kcal/mol for affecting protein stability upon a missense mutation (Gerasimavicius et al., 2020); however, two mutations, namely p.Pro359Thr in WASP and p.Arg525Gln in BTK, were falsely predicted to result in protein loss. Further, three mutations, namely p.Gly220Arg in perforin, p.Met63Lys in AP3B1, and p.Arg211His in ADA, were predicted to destabilize the proteins, while proteins harboring the mutations showed no loss of abundance in our experiments. These results indicate that use of an *in silico* approach is not sufficiently reliable, in terms of sensitivity and specificity, to evaluate the impact of missense mutations on protein abundance.

### **Secondary effects of IEI gene mutations on DBS protein profiles**

Next, we evaluated the secondary effects of IEI mutations and searched for proteins that could serve as alternative markers for diagnosis/screening of IEIs, particularly those that were unidentifiable by direct evaluation of the causative gene products. For accurate evaluation of DEPs between patients and healthy controls, age-related physiological alterations in protein levels must be taken into account; however, because of their extreme rarity, collection of newborn DBS samples for every IEI is almost impossible. Furthermore, secondary protein profile modification must be considered, since many patient samples were collected after disease onset. Therefore, we searched for possible disease markers by evaluating changes in expression beyond the level of age-dependent variation and secondary modification after disease onset. To this end, we

first compared the protein profiles of DBS samples from patients with specific diseases with those from healthy newborns and with those from adults. We then searched for biological pathways that were commonly altered in patient samples relative to both control groups, followed by identification of specific molecules in these pathways that were also in the list of DEPs.

The neutrophil degranulation pathway was commonly altered in patients with HPS type 2 relative to healthy newborns and adults, as well as in patients with CHS compared to the two control groups (Fig. 4, A and B). Among common DEPs shared in this pathway, levels of p22-phox (*CYBA*), cathepsin G (*CTSG*), and myeloperoxidase (*MPO*), were reduced in samples from patients with HPS type 2, whereas those of cathepsin G and ELANE (*ELANE*) were reduced in patients with CHS relative to controls (Fig. 4 C). These findings likely reflect the phenotypes of each disease; that is, neutropenia in HPS type 2 and cell granule abnormalities in CHS. By combining ADA and common  $\gamma$  chain deficiencies together as SCID, immune system and neutrophil degranulation pathways were commonly altered in patients with SCID compared to healthy newborns and adults (Fig. 4 D). Among common DEPs in these pathways, levels of CD3E were exclusively lower in samples from patients with SCID than in healthy controls (Fig. 4 E), likely reflecting the phenotypic alteration in blood cell numbers in this disease (T cell lymphopenia). No protein was found to be common to pathways that were altered in patients with FHL type 2 relative to healthy newborns and adults.

Interestingly, the hemostasis pathway was commonly altered in patients with WAS relative to healthy newborns and adults (Fig. 4 G). Among the common DEPs in this pathway, CD41 (*ITGA2B*) and

Rab27b (*RAB27B*) were selectively reduced in patients with WAS, likely reflecting thrombocytopenia (Fig. 4 H). This explains the reduction of Munc13-4 levels in WAS samples, as Munc13-4 is abundant in platelets (Fig. 4 A) (Murata et al., 2011). We also found that levels of proteins related to T lymphocytes (CD2, CD3D, CD3E, CD5, CD247), platelets (CD41, CD42b), and neutrophils (CD11b, CD18, CD33, CD66b) were reduced in patients with T cell lymphopenia, thrombocytopenia, and neutropenia, respectively, and could be clustered as distinct groups (Fig. 5), suggesting that assays of these proteins could be used to identify patients with these cell-phenotypical alterations.

## Discussion

Screening for genetic diseases enables early identification of patients who need life-saving treatment before disease onset. The Wilson and Jungner recommendations for population-based screening have guided decisions regarding candidate disease inclusion in NBS programs for the past 50 years; however, the changing landscape of clinical practice and newer technologies, including genome-based methodologies, as well as developing perspectives on the ethical, legal, and social implications of population-based screening practices, necessitate review of these historic criteria (King et al., 2021). The analysis of multiple genes in newborns raises ethical considerations, and difficulties arise in interpreting variants with yet undetermined significance. The primary purpose of NBS is to convincingly detect significant signs of suspected genetic disease at the pre-symptomatic stage, rather than to conclusively diagnose newborns. Thus, it is reasonable

that IEMs have been screened by measuring metabolite levels or activities of metabolic enzymes, as a first-tier test, as abnormal levels of the metabolite or enzymatic activity of interest are directly relevant to IEM pathogenesis; however, suitable markers for other inherited diseases, which can be used to detect functional anomalies at molecular and/or cellular levels before macroscopic disease signs manifest, remain elusive. In this context, we examined the applicability of quantitative protein profiling of DBS, supported by recent progress in DIA-LC-MS/MS technologies, including the development of a simple method for pre-treatment of DBS for non-targeted proteome measurements (Nakajima et al., 2020).

Using the newly-developed DIA-LC-MS/MS method, we successfully identified and quantified 2912 proteins in all healthy newborn samples tested. Evaluation of these proteins in newborn DBS samples revealed that their profiles were strongly correlated among newborn samples, which clustered distinctly from adult samples (Fig. 1, A and B). In agreement with previous reports, newborn DBS contained more fetus-specific proteins (AFP, fetal Hb-related components), as well as proteins related to translational initiation pathways and cell-cell adhesion, and fewer IgA/IgM-related proteins and proteins associated with coagulation and complement pathways (Fig. 2, A–D). Of consistently quantified proteins, 1110 were listed in OMIM and 781 were related to diseases with autosomal recessive or X-linked inheritance (Table S2). These included candidate diseases for NBS, such as Menkes disease, Wilson disease, congenital platelet abnormalities, and coagulation factor deficiencies, as well as various IEMs that have potential for screening by

detection of reduced levels of disease-responsible proteins.

IEIs are caused by monogenic germline variants that result in loss or gain of function of the encoded proteins, and patients with these conditions present with increased susceptibility to a spectrum of infections, as well as various autoimmune, autoinflammatory, allergic, and/or malignant phenotypes (Bousfiha et al., 2022). IEIs are associated with significant morbidity and mortality if left untreated (King and Hammarström, 2018). Patients with severe IEIs require hematopoietic cell transplantation or other curative approaches, with earlier intervention resulting in better outcomes (Bakhtiar et al., 2021; Bergsten et al., 2020; Burroughs et al., 2020; Chiesa et al., 2020; Lucchini et al., 2018). Measuring T cell receptor excision circles (TRECs) and kappa-deleting element recombination circles (KRECs) in DBS samples can identify patients with T and B cell lymphopenias, respectively, and facilitates early treatment intervention and reduction of overall medical costs (Borte et al., 2012; Chan and Puck, 2005; Elsink et al., 2020; King et al., 2021; Nakagawa et al., 2011; Sheller et al., 2020). These results indicate that feasible NBS-based methods to detect other severe IEIs could improve patient outcomes. Furthermore, live vaccines administered in early infancy often cause opportunistic infections in certain groups of infants with IEIs, such as those with SCID and chronic granulomatous disease (Conti et al., 2016), which could be prevented by NBS.

Our data demonstrate that non-targeted DBS proteomics analysis can detect reduced levels of disease-responsible proteins in samples from patients with FHL type 3, FHL type 5, p91-phox, and p22-phox deficiencies, WAS, and XLA, relative to healthy newborn control samples (Fig. 3, A, B, E, F, H, and I).

Moreover, the levels of p22-phox and Syntaxin-11 were reduced in parallel with those of p91-phox and Munc18-2, respectively (Fig. 3, B–C, E–F), in agreement with previous reports (Parkos et al., 1989; zur Stadt et al., 2009). Munc13-4 protein was specifically decreased in all pre-symptomatic newborn samples from patients with FHL type 3, which had been used for screening for metabolic disorders and stored for months to years, in the same way as detected in post-symptomatic samples (Fig. 3 A). These findings demonstrate the robustness of DBS proteomics for NBS of IEIs and also suggest its potential usefulness for diagnosing patients living in regions where other diagnostic modalities are unavailable.

The levels of perforin, *LYST*, and ADA proteins were not significantly reduced in patients with FHL type 2, CHS, and ADA deficiency, respectively (Fig. 3, D, J and L). Many *PRF1* mutations result in inappropriate folding or polymerization of perforin molecules, with residual protein expression on western blotting analysis (Voskoboinik et al., 2005). Further, many *LYST* mutations result in residual expression of truncated protein (Certain et al., 2000), which cannot be distinguished from full-length *LYST* on LC-MS/MS, unless a specific peptide near the C-terminus is evaluated. Furthermore, measurement of ADA by the current method may not be reliable because ADA is predominantly cytosolic and hydrophilic. Interestingly, *AP3B1* levels varied greatly among patients with HPS type 2, with the highest level in a patient with a splicing site mutation near the 3' end of one allele (Patient #13 in Table 1) and the lowest level in a patient who had experienced an episode of hemophagocytic lymphohistiocytosis (Patient #12 in Table 1) (Fig. 3 K), suggesting that HPS type 2 disease severity may correlate with residual *AP3B1* expression.

As our method was non-targeted, we could also detect disease markers that reflect disease-associated phenotypes. Levels of p22-phox, cathepsin G, and myeloperoxidase were lower in DBS from patients with HPS type 2, and those of cathepsin G and ELANE were lower in patients with CHS, compared with controls (Fig. 4 C). These findings reflect the clinical phenotypes of the conditions; neutropenia in patients with HPS type 2 (Jung et al., 2006), and abnormal azurophil granules in patients with CHS (Burnett et al., 1995). As p22-phox, ELANE, and cathepsin G are primarily expressed in granulocytes, analyses of these proteins may also be useful in identifying patients with other diseases that are accompanied by neutropenia. CD3E levels were found to be reduced in patients with SCID (Fig. 4 E), and those of Rab27B and platelet-related proteins were low in patients with WAS (Fig. 4 H), likely reflecting thrombocytopenia and T cell lymphopenia, respectively, in these two IEIs. These results demonstrate that DBS proteomics can detect major changes in the number and/or characteristics of specific cell populations that can currently be analyzed only in specialized laboratories (Fig. 5). Notably, our method could identify patients with T cell lymphopenia and BTK deficiency and may complement TRECs/KRECs analyses, which have high sensitivity but relatively low specificity. DBS proteomics could be performed in parallel with TRECs/KRECs screening, using an additional punchout from the same sample, thus providing phenotypic specificity.

Another important aspect of the current study is that proteome analysis can provide information more directly relevant to patient conditions than that generated by genetic analysis. As shown in Table 2, the



effects of genetic variants on levels of their encoded proteins are often unpredictable using available *in silico* analysis methods, and interpretation is difficult unless pathogenicity has been reported previously. Non-targeted DBS proteomics can provide in-depth molecular phenotype information and will aid in linking gene structural information with clinical phenotypes.

We acknowledge that this study has several limitations. First, most IEI samples analyzed here were obtained after disease onset, with only six pre-symptomatic neonatal samples included; five from patients with FHL type 3 and one from a patient with CHS. Therefore, further analysis of newborn IEI samples is required to prove the usefulness of the current method for NBS; however, most IEIs analyzed in this study are characterized by significant reductions in expression of disease-related proteins, or in specific cell populations, during the newborn period, which enables the diagnosis of siblings of index patients during their neonatal period before disease onset. In addition, healthy newborns experience physiological leukocytosis, resulting in high expression of leukocyte-related proteins during the neonatal period, and the proteins analyzed in the current study were consistently expressed in all healthy newborn samples. Thus, we do not anticipate that detecting reduced levels of disease marker proteins in newborn samples would be difficult. Second, not all patients display reduced marker protein levels, but protein expression levels often correlate with disease severity (Almarza Novoa et al., 2018; Kuhns et al., 2010; Zhu et al., 1997) and screening for patients with reduced protein levels would be of significant clinical impact. Third, we conducted DIA proteomic analysis with LC-MS/MS to evaluate a wide variety of non-target proteins, which is more time

consuming than targeted analysis, such as monitoring of selected or parallel reactions, and could lead to challenges in providing prompt results for NBS; however, rapid advances in MS instrumentation and DIA-LC-MS/MS technology have led to the development of deep non-targeted proteome analysis with short measurement time (Ishikawa et al., 2022; Guzman et al., 2023). Therefore, many patients with IEIs characterized by significant reductions in disease marker proteins could be amenable to screening using non-targeted DIA-LC-MS/MS methods in the near future. Forth, we did not evaluate the unit cost of the test in the current study; the low incidence of severe genetic diseases means that the unit cost of a test is key to its potential for application. The method described in this study can be applied to a wide variety of congenital disorders, including metabolic/neurological and hematological diseases, associated with reduced levels of causative proteins in hydrophobic fractions of peripheral blood cells. Further, this method can readily be performed in parallel with other screening methods that use the same DBS samples, including TRECs/KRECs analysis and proteome analysis with pre-enrichment of target proteins using specific antibodies. Furthermore, advances in proteomics technology may expand the range of detectable peptides from disease-causing gene products, to allow detection of single amino acid alterations in abundant proteins (Salz et al., 2021). Together, such approaches will increase the number of disorders that can be screened simultaneously and reduce the per disease cost. In addition, DBS samples can be used for genome-wide genetic analysis, to confirm diagnoses and potentially lead to early intervention. Although additional improvements are needed for population-based applications, methodological refinements and innovations

will enable the use of this approach to screen newborns for various diseases and strengthen the application of genome-wide sequencing-based screening in the future (King et al., 2021; Strand et al., 2020).

In conclusion, our findings demonstrate that levels of multiple proteins associated with genetic disorders in DBS samples can be semi-quantified, with no need for pre-enrichment using specific antibodies. Proteomic analysis of DBS samples can serve as a “single-platform-to-assess-many-diseases” NBS approach for IEIs and various other genetic disorders, enabling the identification of patients before disease onset, allowing earlier intervention in the absence of severe complications, and improving patient prognosis.

## **Materials and methods**

### **Patients and samples**

DBS samples on Guthrie cards from 40 healthy newborns used to screen for metabolic disorders were collected, as were samples from 8 healthy adults, and 45 samples from 43 patients with 12 IEI diseases (Table 1). Of the IEI samples, 38 were obtained after disease onset and stored at their respective institutions for research purposes, and the remaining FHL type 3 (n = 5) and CHS (n = 1) samples were obtained pre-symptomatically for use in public health screening for metabolic diseases. The study protocol, which conformed to the principles of the Declaration of Helsinki, was approved by the ethics committee of Kyoto University Hospital. Written informed consent was obtained from all participants or their legal guardians.

### **Preparation and digestion of the insoluble protein fraction**

Soluble fractions were reduced and proteins in DBS digested as previously described (Nakajima et al., 2020), with slight modification. Briefly, a disc (3.2 mm diameter) punched from a DBS was suspended in 1 mL of Na<sub>2</sub>CO<sub>3</sub> solution and vigorously shaken in the presence of a 5 mm zirconia bead (Tomy Seiko, Tokyo, Japan) using a Tissue Lyser (Qiagen, Hilden, Germany) oscillating at 25 Hz for 5 min at room temperature. The remaining disc was removed by centrifugation (3,000 × g, 3 min, 4°C) and the supernatant (approximately 850 µL) was centrifuged (17,400 g, 15 min, 4°C). The precipitate was resuspended in 200 µL 100 mM sodium carbonate aqueous solution by vigorous vortexing at room temperature. After centrifugation (17,400 g, 5 min, 4°C), the supernatant was completely removed and the precipitate dissolved in 100 µL 12 mM sodium deoxycholate, 12 mM N-lauroylsarcosine, and 100 mM Tris-Cl pH 8.0 by vigorous vortexing for 1 min, followed by sonication using Bioruptor II (CosmoBio, Tokyo, Japan) on high power mode for 10 min in 30 s on/30 s off cycles. A 500 ng aliquot of trypsin/lys-C mix (Promega, Madison, WI) was added to each protein extract and samples were digested overnight at 37°C. After precipitation, detergent was removed by acidification with 30 µL 5% trifluoroacetic acid, followed by sonication with a Bioruptor II in high power mode for 5 min in 30 s on/30 s off cycles. The mixture was shaken for 5 min and centrifuged (15,000 g, 5 min). The supernatant was desalted using GL-Tip SDB (GL Sciences Inc., Tokyo, Japan), according to the manufacturer's instructions, and dried using a centrifugal evaporator (miVac Duo concentrator, Genevac, Ipswich, UK). Dried peptides were redissolved in 3% ACN, 0.1% formic acid.

## **LC-MS/MS**

Digested peptides were directly injected onto a 75  $\mu\text{m} \times 20$  cm PicoFrit emitter (New Objective, Woburn, MA) packed in-house with C18 core-shell particles (CAPCELL CORE MP 2.7  $\mu\text{m}$ , 160  $\text{\AA}$  material; Osaka Soda Co., Ltd., Osaka, Japan) at 50°C and peptides separated using an UltiMate 3000 RSLCnano LC system (Thermo Fisher Scientific, Waltham, MA), with an 80 min gradient at 100 nL/min. Peptides eluted from the column were analyzed on an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific) for overlapping window DIA-MS (Amodei et al., 2019; Kawashima et al., 2019). MS1 spectra were collected in the range, 495–785 m/z, at 30,000 resolution, to set an automatic gain control target of  $3 \times 10^6$  and maximum injection time of “auto”. MS2 spectra were collected in the range, 200–1800 m/z, at 30,000 resolution to set an automatic gain control target of  $3 \times 10^6$ , maximum injection time of “auto”, and stepped normalized collision energies of 22%, 26%, and 30%. The MS2 isolation width was set to 4 m/z, and overlapping window patterns of 500–780 m/z used for window placements optimized by Skyline (20.2.0.34, University of Washington, Seattle, WA) (MacLean et al., 2010).

## **MS data analysis**

MS files were compared with human spectral libraries using Scaffold DIA (2.2.1, Proteome Software, Inc., Portland, OR). Human spectral libraries were generated from the human protein sequence database (UniProt

id UP000005640, reviewed, canonical) using ProSIT (<https://www.proteomicsdb.org/prosit/>) (Gessulat et al., 2019; Searle et al., 2020). Scaffold DIA search parameters were as follows: experimental data search enzyme, trypsin; maximum missed cleavage sites, 1; precursor mass tolerance, 9 ppm; and fragment mass tolerance, 9 ppm. The protein identification threshold was set at < 1% for both peptide and protein false discovery rates. Proteins and peptides were quantified using the EncyclopeDIA algorithm in Scaffold DIA (Searle et al., 2018).

### **Prediction of effects of missense mutations on protein stability**

Crystal structures of the proteins encoded by the disease-causing genes ADA, CYBA, CYBB, IL2RG, and STXBP2 were downloaded from Protein Data Bank (<https://pdj.org/>) (PDB codes: 3iar for ADA; 8gz3 for CYBA and CYBB; 5m5e for IL2RG; 4cca for STXBP2). Where there were no experimentally determined structures for proteins with mutated residues (namely, AP3B1, PRF1, UNC13D, and WAS), protein model structures were obtained from the AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>). Computational predictions of the impacts of missense mutations on protein stability were evaluated using FoldX, which is a state-of-the-art method for stability estimation based on the three-dimensional protein structures (Delgado et al., 2019). If the total energy stability change value of the mutant protein was > 1.58 kcal/mol, the missense mutation was predicted to affect protein stability, following a previous report (Gerasimavicius et al., 2020).

## Statistical analysis

Protein expression levels were standardized using those of  $\beta$ -actin (% $\beta$ -actin). Exploratory analysis, including  $\log_2$  transformation, multiple scatter plotting, calculation of Pearson correlation coefficient values between two sets of data, data imputation (width = 0.3, Down shift = 2.4), Z-score normalization, and hierarchical comparison were performed using the Perseus software platform (<https://maxquant.net/perseus/>). Volcano plots were created by multiple t test using the two-stage step up method of Benjamini, Krieger, and Yekutieli in GraphPad Prism (version 9.0.2) software (GraphPad Software, Inc., La Jolla, CA). DEPs between two groups were selected for further analysis using a cutoff of adjusted P value (q value)  $< 0.01$  and fold-change  $\geq 2$ . Statistical differences among groups were analyzed using one-way ANOVA on ranks (Kruskal-Wallis test), followed by Dunn's multiple comparisons test using GraphPad Prism. A P value  $< 0.05$  was considered significant.

## Supplemental material

**Fig. S1** shows a Wikipathway map of the Complement and coagulation cascades highlighting downregulated proteins ( $P < 0.05$ ) in healthy newborns relative to adults. **Fig. S2** illustrates quantification of proteins responsible for various genetic disorders in neonatal DBS samples. **Fig. S3** demonstrates that Munc13-4 is conspicuously reduced in DBS samples from patients with FHL type 3. **Fig. S4** presents flow cytometric

analysis of perforin expression in natural killer cells from patients with familial hemophagocytic lymphohistiocytosis type 2. **Table S1** lists proteins identified in DBS samples analyzed in this study. **Table S2** lists proteins identified in DBS samples that are cataloged in the OMIM database. Table S1 and S2 are available at <https://doi.org/10.5281/zenodo.10539971>.

### **Data availability**

All of the data used to generate the figures are available in the published article and its online supplemental material.

### **Acknowledgments**

The authors are grateful to all participating patients and their families. This work was supported by JSPS KAKENHI under Grant Numbers 19K17328 and 21K07795; by AMED under Grant Number JP23ek0109586; and by the “Research on Measures for Intractable Diseases” project from the Japanese Ministry of Health, Labor, and Welfare.

### **Authorship**

Contributions: H.S., O.O., Y.K., and T.Y. designed the research; H.S., M.H., S.S., T. Miyamoto, and N-I.M. performed flow cytometric analyses of patients with FHL and collected their DBS samples; D.N., R.K., and Y.K. performed proteome analyses; K.O. and O.O. performed genetic analyses; A.H. and O.O. performed *in silico* prediction of missense mutation effects on protein stability; H.O. provided DBS samples from healthy babies; K. Izawa, H.O., M.I., M.Y., H.K., Y.N., S.N., T.U., M.O., and R.N. attended to patients and provided their samples; H.S., M.H., E.H., K. Izawa, J.T., T.H., K. Imai, T. Morio, O.O., Y.K., and T.Y. analyzed the data and discussed the results; and H.S., Y.K., and T.Y. wrote the paper.

**Conflict of interest disclosure:** The authors declare no competing financial interests.

**Correspondence:** Takahiro Yasumi, Department of Pediatrics, Kyoto University Graduate School of Medicine, 54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto, 606-8507 JAPAN, e-mail: [yasumi@kuhp.kyoto-u.ac.jp](mailto:yasumi@kuhp.kyoto-u.ac.jp); Yusuke Kawashima, Kazusa DNA Research Institute, 2-5-23 Kazusa-Kamatari, Kisarazu, 292-0818, JAPAN, e-mail: [ykawashi@kazusa.or.jp](mailto:ykawashi@kazusa.or.jp)



## References

- Adhikari, A.N., R.C. Gallagher, Y. Wang, R.J. Currier, G. Amatuni, L. Bassaganyas, F. Chen, K. Kundu, M. Kvale, S.D. Mooney, R.L. Nussbaum, S.S. Randi, J. Sanford, J.T. Shieh, R. Srinivasan, U. Sunderam, H. Tang, D. Vaka, Y. Zou, B.A. Koenig, P.Y. Kwok, N. Risch, J.M. Puck, and S.E. Brenner. 2020. The role of exome sequencing in newborn screening for inborn errors of metabolism. *Nat Med* 26:1392-1397.
- Almarza Novoa, E., S. Kasbekar, A.J. Thrasher, D.B. Kohn, J. Sevilla, T. Nguyen, J.D. Schwartz, and J.A. Bueren. 2018. Leukocyte adhesion deficiency-I: A comprehensive review of all published cases. *J Allergy Clin Immunol Pract* 6:1418-1420.e1410.
- Amodei, D., J. Egertson, B.X. MacLean, R. Johnson, G.E. Merrihew, A. Keller, D. Marsh, O. Vitek, P. Mallick, and M.J. MacCoss. 2019. Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. *J Am Soc Mass Spectrom* 30:669-684.
- Bakhtiar, S., E. Salzmann-Manrique, H.J. Blok, D.J. Eikema, S. Hazelaar, M. Ayas, A. Toren, G. Goldstein, D. Moshous, F. Locatelli, P. Merli, G. Michel, G. Öztürk, A. Schulz, C. Heilmann, M. Ifversen, R.F. Wynn, O. Aleinikova, Y. Bertrand, A. Tbakhi, P. Veys, M. Karakukcu, A. Kupesiz, A. Ghavamzadeh, R. Handgretinger, E. Unal, A. Perez-Martinez, M. Gokce, F. Porta, T. Aksu, G. Karasu, I. Badell, P. Ljungman, E. Skorobogatova, A. Yesilipek, T. Zuckerman, R.R.G. Bredius, P. Stepensky, B. Shadur, M. Slatter, A.R. Gennery, M.H. Albert, P. Bader, and A. Lankester. 2021. Allogeneic hematopoietic stem cell transplantation in leukocyte adhesion deficiency type I and III. *Blood Adv* 5:262-273.
- Berg, J.S., P.B. Agrawal, D.B. Bailey, A.H. Beggs, S.E. Brenner, A.M. Brower, J.A. Cakici, O. Ceyhan-Birsoy, K. Chan, F. Chen, R.J. Currier, D. Dukhovny, R.C. Green, J. Harris-Wai, I.A. Holm, B. Iglesias, G. Joseph, S.F. Kingsmore, B.A. Koenig, P.Y. Kwok, J. Lantos, S.J. Leeder, M.A. Lewis, A.L. McGuire, L.V. Milko, S.D. Mooney, R.B. Parad, S. Pereira, J. Petrikin, B.C. Powell, C.M. Powell, J.M. Puck, H.L. Rehm, N. Risch, M. Roche, J.T. Shieh, N. Veeraraghavan, M.S. Watson, L. Willig, T.W. Yu, T. Urv, and A.L. Wise. 2017. Newborn Sequencing in Genomic Medicine and Public Health. *Pediatrics* 139:
- Bergsten, E., A. Horne, I. Hed Myrberg, M. Aricó, I. Astigarraga, E. Ishii, G. Janka, S. Ladisch, K. Lehmberg, K.L. McClain, M. Minkov, V. Nanduri, D.A. Rosso, E. Sieni, J. Winiarski, and J.I. Henter. 2020. Stem cell transplantation for children with hemophagocytic lymphohistiocytosis: results from the HLH-2004 study. *Blood Adv* 4:3754-3766.
- Borte, S., U. von Döbeln, A. Fasth, N. Wang, M. Janzi, J. Winiarski, U. Sack, Q. Pan-Hammarström, M. Borte, and L. Hammarström. 2012. Neonatal screening for severe primary immunodeficiency diseases using high-throughput triplex real-time PCR. *Blood* 119:2552-2555.
- Bousfiha, A., A. Moundir, S.G. Tangye, C. Picard, L. Jeddane, W. Al-Herz, C.C. Rundles, J.L. Franco, S.M. Holland, C. Klein, T. Morio, E. Oksenhendler, A. Puel, J. Puck, M.R.J. Seppänen, R. Somech, H.C. Su, K.E. Sullivan, T.R. Torgerson, and I. Meyts. 2022. The 2022 Update of IUIS Phenotypical Classification for Human Inborn Errors of Immunity. *J Clin Immunol* 42:1508-1520.

- Burnett, D., C.J. Ward, R.A. Stockley, R.G. Dalton, A.J. Cant, S. Hoare, and J. Crocker. 1995. Neutrophil elastase and cathepsin G protein and messenger RNA expression in bone marrow from a patient with Chediak-Higashi syndrome. *Clin Mol Pathol* 48:M28-34.
- Burroughs, L.M., A. Petrovic, R. Brazauskas, X. Liu, L.M. Griffith, H.D. Ochs, J.J. Bleesing, S. Edwards, C.C. Dvorak, S. Chaudhury, S.E. Prockop, R. Quinones, F.D. Goldman, T.C. Quigg, S. Chandrakasan, A.R. Smith, S. Parikh, B.J. Dávila Saldaña, M.S. Thakar, R. Phelan, S. Shenoy, L.R. Forbes, C. Martinez, D. Chellapandian, E. Shereck, H.K. Miller, N. Kapoor, J.L. Barnum, H. Chong, D.C. Shyr, K. Chen, R. Abu-Arja, A.J. Shah, K.G. Weinacht, T.B. Moore, A. Joshi, K.B. DeSantes, A.P. Gillio, G.D.E. Cuvelier, M.D. Keller, J. Rozmus, T. Torgerson, M.A. Pulsipher, E. Haddad, K.E. Sullivan, B.R. Logan, D.B. Kohn, J.M. Puck, L.D. Notarangelo, S.Y. Pai, D.J. Rawlings, and M.J. Cowan. 2020. Excellent outcomes following hematopoietic cell transplantation for Wiskott-Aldrich syndrome: a PIDTC report. *Blood* 135:2094-2105.
- Certain, S., F. Barrat, E. Pastural, F. Le Deist, J. Goyo-Rivas, N. Jabado, M. Benkerrou, R. Seger, E. Vilmer, G. Beullier, K. Schwarz, A. Fischer, and G. de Saint Basile. 2000. Protein truncation test of LYST reveals heterogenous mutations in patients with Chediak-Higashi syndrome. *Blood* 95:979-983.
- Chan, K., and J.M. Puck. 2005. Development of population-based newborn screening for severe combined immunodeficiency. *J Allergy Clin Immunol* 115:391-398.
- Chiesa, R., J. Wang, H.J. Blok, S. Hazelaar, B. Neven, D. Moshous, A. Schulz, M. Hoenig, F. Hauck, A. Al Seraihy, J. Gozdzik, P. Ljungman, C.A. Lindemans, J.F. Fernandes, K. Kalwak, B. Strahm, U. Schanz, P. Sedlacek, K.W. Sykora, S. Aksoylar, F. Locatelli, P. Stepensky, R. Wynn, S.H. Lum, M. Zecca, F. Porta, M. Taskinen, B. Gibson, S. Matthes, M. Karakukcu, M. Hauri-Hohl, P. Veys, A.R. Gennery, G. Lucchini, M. Felber, M.H. Albert, D. Balashov, A. Lankester, T. Güngör, and M.A. Slatter. 2020. Hematopoietic cell transplantation in chronic granulomatous disease: a study of 712 children and adults. *Blood* 136:1201-1211.
- Collins, C.J., F. Yi, R. Dayuha, J.R. Whiteaker, H.D. Ochs, A. Freeman, H.C. Su, A.G. Paulovich, G.R.S. Segundo, T. Torgerson, and S.H. Hahn. 2020. Multiplexed Proteomic Analysis for Diagnosis and Screening of Five Primary Immunodeficiency Disorders From Dried Blood Spots. *Front Immunol* 11:464.
- Conti, F., S.O. Lugo-Reyes, L. Blancas Galicia, J. He, G. Aksu, E. Borges de Oliveira, C. Deswarte, M. Hubeau, N. Karaca, M. de Suremain, A. Guérin, L.A. Baba, C. Prando, G.G. Guerrero, M. Emiroglu, F.N. Öz, M.A. Yamazaki Nakashimada, E. Gonzalez Serrano, S. Espinosa, I. Barlan, N. Pérez, L. Regairaz, H.E. Guidos Morales, L. Bezrodnik, D. Di Giovanni, G. Dbaibo, F. Ailal, M. Galicchio, M. Oleastro, J. Chemli, S. Danielian, L. Perez, M.C. Ortega, S. Soto Lavin, J. Hertecant, O. Anal, N. Kechout, E. Al-Idrissi, G. ElGhazali, A. Bondarenko, L. Chernyshova, P. Ciznar, R.M. Herbigneaux, A. Diabate, S. Ndaga, B. Konte, A. Czarna, M. Migaud, S. Pedraza-Sánchez, M.B. Zaidi, G. Vogt, S. Blanche, I. Benmustapha, D. Mansouri, L. Abel, S. Boisson-Dupuis, N. Mahlaoui, A.A. Bousfiha, C. Picard, R. Barbouche, S. Al-Muhsen, F.J. Espinosa-Rosales, N. Kütükçüler, A. Condino-Neto, J.L.

- Casanova, and J. Bustamante. 2016. Mycobacterial disease in patients with chronic granulomatous disease: A retrospective analysis of 71 cases. *J Allergy Clin Immunol* 138:241-248.e243.
- Delgado, J., L.G. Radusky, D. Cianferoni, and L. Serrano. 2019. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics* 35:4168-4169.
- Dezfouli, M., S. Bergström, L. Skattum, H. Abolhassani, M. Neiman, M. Torabi-Rahvar, C. Franco Jarava, A. Martin-Nalda, J.M. Ferrer Balaguer, C.A. Slade, A. Roos, L.M. Fernandez Pereira, M. López-Trascasa, L.I. Gonzalez-Granado, L.M. Allende-Martinez, Y. Mizuno, Y. Yoshida, V. Friman, Å. Lundgren, A. Aghamohammadi, N. Rezaei, M. Hernández-Gonzalez, U. von Döbeln, L. Truedsson, T. Hara, S. Nonoyama, J.M. Schwenk, P. Nilsson, and L. Hammarström. 2020. Newborn Screening for Presymptomatic Diagnosis of Complement and Phagocyte Deficiencies. *Front Immunol* 11:455.
- Elsink, K., J.M. van Montfrans, M.E. van Gijn, M. Blom, P.M. van Hagen, T.W. Kuijpers, and G.W.J. Frederix. 2020. Cost and impact of early diagnosis in primary immunodeficiency disease: A literature review. *Clin Immunol* 213:108359.
- Gerasimavicius, L., X. Liu, and J.A. Marsh. 2020. Identification of pathogenic missense mutations using protein stability predictors. *Sci Rep* 10:15387.
- Gessulat, S., T. Schmidt, D.P. Zolg, P. Samaras, K. Schnatbaum, J. Zerweck, T. Knaute, J. Rechenberger, B. Delanghe, A. Huhmer, U. Reimer, H.C. Ehrlich, S. Aiche, B. Kuster, and M. Wilhelm. 2019. Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods* 16:509-518.
- GUTHRIE, R., and A. SUSI. 1963. A SIMPLE PHENYLALANINE METHOD FOR DETECTING PHENYLKETONURIA IN LARGE POPULATIONS OF NEWBORN INFANTS. *Pediatrics* 32:338-343.
- Guzman, U.H., A.M.D. Val, Z. Ye, E. Damoc, T.N. Arrey, A. Pashkova, E., Denisov, J. Petzoldt, A.C. Peterson, F. Harking, O. Østergaard, H., Stewart, Y. Xuan, D. Hermanson, C. Hock, A. Makarov, V. Zabrouskov, and J.V. Olsen. 2023. Narrow-window DIA: Ultra-fast quantitative analysis of comprehensive proteomes with high sequencing depth. *bioRxiv* doi.org/10.1101/2023.06.02.543374 (Preprint posted June 06, 2023)
- Ishikawa, M., R. Konno, D. Nakajima, M. Gotoh, K. Fukasawa, H. Sato, R. Nakamura, O. Ohara, and Y. Kawashima. 2022. Optimization of Ultrafast Proteomics Using an LC-Quadrupole-Orbitrap Mass Spectrometer with Data-Independent Acquisition. *J Proteome Res* 21:2085-2093.
- Jung, J., G. Bohn, A. Allroth, K. Boztug, G. Brandes, I. Sandrock, A.A. Schäffer, C. Rathinam, I. Köllner, C. Beger, R. Schilke, K. Welte, B. Grimbacher, and C. Klein. 2006. Identification of a homozygous deletion in the AP3B1 gene causing Hermansky-Pudlak syndrome, type 2. *Blood* 108:362-369.
- Kanakoudi, F., V. Drossou, V. Tzimouli, E. Diamanti, T. Konstantinidis, A. Germenis, and G. Kremenopoulos. 1995. Serum concentrations of 10 acute-phase proteins in healthy term and preterm infants from birth to age 6 months. *Clin Chem* 41:605-608.

- Kawashima, Y., H. Nagai, R. Konno, M. Ishikawa, D. Nakajima, H. Sato, R. Nakamura, T. Furuyashiki, and O. Ohara. 2022. Single-Shot 10K Proteome Approach: Over 10,000 Protein Identifications by Data-Independent Acquisition-Based Single-Shot Proteomics with Ion Mobility Spectrometry. *J Proteome Res* 21:1418-1427.
- Kawashima, Y., E. Watanabe, T. Umeyama, D. Nakajima, M. Hattori, K. Honda, and O. Ohara. 2019. Optimization of Data-Independent Acquisition Mass Spectrometry for Deep and Highly Sensitive Proteomic Analysis. *Int J Mol Sci* 20:5932.
- King, J.R., and L. Hammarström. 2018. Newborn Screening for Primary Immunodeficiency Diseases: History, Current and Future Practice. *J Clin Immunol* 38:56-66.
- King, J.R., L.D. Notarangelo, and L. Hammarström. 2021. An appraisal of the Wilson & Jungner criteria in the context of genomic-based newborn screening for inborn errors of immunity. *J Allergy Clin Immunol* 147:428-438.
- Kuhns, D.B., W.G. Alvord, T. Heller, J.J. Feld, K.M. Pike, B.E. Marciano, G. Uzel, S.S. DeRavin, D.A. Priel, B.P. Soule, K.A. Zarembler, H.L. Malech, S.M. Holland, and J.I. Gallin. 2010. Residual NADPH oxidase and survival in chronic granulomatous disease. *N Engl J Med* 363:2600-2610.
- Lucchini, G., R. Marsh, K. Gilmour, A. Worth, Z. Nademi, A. Rao, C. Booth, P. Amrolia, J. Silva, R. Chiesa, R. Wynn, K. Lehmborg, I. Astigarraga, T. Güngör, J. Stary, D. Moshous, M. Ifversen, D. Zinn, M. Jordan, A. Kumar, T. Yasumi, P. Veys, and K. Rao. 2018. Treatment dilemmas in asymptomatic children with primary hemophagocytic lymphohistiocytosis. *Blood* 132:2088-2096.
- MacLean, B., D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, and M.J. MacCoss. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966-968.
- Murata, Y., T. Yasumi, R. Shirakawa, K. Izawa, H. Sakai, J. Abe, N. Tanaka, T. Kawai, K. Oshima, M. Saito, R. Nishikomori, O. Ohara, E. Ishii, T. Nakahata, H. Horiuchi, and T. Heike. 2011. Rapid diagnosis of FHL3 by flow cytometric detection of intraplatelet Munc13-4 protein. *Blood* 118:1225-1230.
- Nakagawa, N., K. Imai, H. Kanegane, H. Sato, M. Yamada, K. Kondoh, S. Okada, M. Kobayashi, K. Agematsu, H. Takada, N. Mitsuiki, K. Oshima, O. Ohara, D. Suri, A. Rawat, S. Singh, Q. Pan-Hammarström, L. Hammarström, J. Reichenbach, R. Seger, T. Ariga, T. Hara, T. Miyawaki, and S. Nonoyama. 2011. Quantification of  $\kappa$ -deleting recombination excision circles in Guthrie cards for the identification of early B-cell maturation defects. *J Allergy Clin Immunol* 128:223-225.e222.
- Nakajima, D., Y. Kawashima, H. Shibata, T. Yasumi, M. Isa, K. Izawa, R. Nishikomori, T. Heike, and O. Ohara. 2020. Simple and Sensitive Analysis for Dried Blood Spot Proteins by Sodium Carbonate Precipitation for Clinical Proteomics. *J Proteome Res* 19:2821-2827.
- Parkos, C.A., M.C. Dinauer, A.J. Jesaitis, S.H. Orkin, and J.T. Curnutte. 1989. Absence of both the 91kD and 22kD subunits of human neutrophil cytochrome b in two genetic forms of chronic granulomatous disease. *Blood* 73:1416-1420.
- Porter, C.D., M.H. Parkar, A.J. Verhoeven, R.J. Levinsky, M.K. Collins, and C. Kinnon. 1994.

- p22-phox-deficient chronic granulomatous disease: reconstitution by retrovirus-mediated expression and identification of a biosynthetic intermediate of gp91-phox. *Blood* 84:2767-2775.
- Reverdiau-Moalic, P., B. Delahousse, G. Body, P. Bardos, J. Leroy, and Y. Gruel. 1996. Evolution of blood coagulation activators and inhibitors in the healthy human fetus. *Blood* 88:900-906.
- Sakura, F., K. Noma, T. Asano, K. Tanita, E. Toyofuku, K. Kato, M. Tsumura, H. Nihira, K. Izawa, K. Mitsui-Sekinaka, R. Konno, Y. Kawashima, Y. Mizoguchi, S. Karakawa, S. Hayakawa, H. Kawaguchi, K. Imai, S. Nonoyama, T. Yasumi, H. Ohnishi, H. Kanegane, O. Ohara, and S. Okada. 2023. A complementary approach for genetic diagnosis of inborn errors of immunity using proteogenomic analysis. *PNAS Nexus* 2:pgad104.
- Salz, R., R. Bouwmeester, R. Gabriels, S. Degroeve, L. Martens, P.J. Volders, and P.A.C. 't Hoen. 2021. Personalized Proteome: Comparing Proteogenomics and Open Variant Search Approaches for Single Amino Acid Variant Detection. *J Proteome Res* 20:3353-3364.
- Searle, B.C., L.K. Pino, J.D. Egerton, Y.S. Ting, R.T. Lawrence, B.X. MacLean, J. Villén, and M.J. MacCoss. 2018. Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat Commun* 9:5128.
- Searle, B.C., K.E. Swearingen, C.A. Barnes, T. Schmidt, S. Gessulat, B. Küster, and M. Wilhelm. 2020. Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nat Commun* 11:1548.
- Sheller, R., J. Ojodu, E. Griffin, S. Edelman, C. Yusuf, T. Pigg, A. Huston, B. Fitzek, J.G. Boyle, and S. Singh. 2020. The Landscape of Severe Combined Immunodeficiency Newborn Screening in the United States in 2020: A Review of Screening Methodologies and Targets, Communication Pathways, and Long-Term Follow-Up Practices. *Front Immunol* 11:577853.
- Murata, Y., T. Yasumi, R. Shirakawa, K. Izawa, H. Sakai, J. Abe, N. Tanaka, T. Kawai, K. Oshima, M. Saito, R. Nishikomori, O. Ohara, E. Ishii, T. Nakahata, H. Horiuchi, and T. Heike. 2011. Rapid diagnosis of FHL3 by flow cytometric detection of intraplatelet Munc13-4 protein. *Blood* 118:1225-1230.
- Shibata, H., T. Yasumi, S. Shimodera, E. Hiejima, K. Izawa, T. Kawai, R. Shirakawa, T. Wada, R. Nishikomori, H. Horiuchi, O. Ohara, E. Ishii, and T. Heike. 2018. Human CTL-based functional analysis shows the reliability of a munc13-4 protein expression assay for FHL3 diagnosis. *Blood* 131:2016-2025.
- Spiekerkoetter, U., D. Bick, R. Scott, H. Hopkins, T. Krones, E.S. Gross, and J.R. Bonham. 2023. Genomic newborn screening: Are we entering a new era of screening? *J Inherit Metab Dis* 46:778-795.
- Strand, J., K.A. Gul, H.C. Erichsen, E. Lundman, M.C. Berge, A.K. Trømborg, L.K. Sjørgjerd, M. Ytre-Arne, S. Hogner, R. Halsne, H.J. Gaup, L.T. Osnes, G.A.B. Kro, H.S. Sorte, L. Mørkrid, A.D. Rowe, T. Tangeraas, J.V. Jørgensen, C. Alme, T.E.H. Bjørndalen, A.E. Rønnestad, A.M. Lang, T. Rootwelt, J. Buechner, T. Øverland, T.G. Abrahamsen, R.D. Pettersen, and A. Stray-Pedersen. 2020. Second-Tier Next Generation Sequencing Integrated in Nationwide Newborn Screening Provides Rapid Molecular Diagnostics of Severe Combined Immunodeficiency. *Front Immunol* 11:1417.

- Ulph, F., and R. Bennett. 2023. Psychological and Ethical Challenges of Introducing Whole Genome Sequencing into Routine Newborn Screening: Lessons Learned from Existing Newborn Screening. *New Bioeth* 29:52-74.
- Voskoboinik, I., M.C. Thia, and J.A. Trapani. 2005. A functional analysis of the putative polymorphisms A91V and N252S and 22 missense perforin mutations associated with familial hemophagocytic lymphohistiocytosis. *Blood* 105:4700-4706.
- Wilcken, B., V. Wiley, J. Hammond, and K. Carpenter. 2003. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* 348:2304-2312.
- Yue, P., Z. Li, and J. Moulton. 2005. Loss of protein structure stability as a major causative factor in monogenic disease. *J Mol Biol* 353:459-473.
- Zhu, Q., C. Watanabe, T. Liu, D. Hollenbaugh, R.M. Blaese, S.B. Kanner, A. Aruffo, and H.D. Ochs. 1997. Wiskott-Aldrich syndrome/X-linked thrombocytopenia: WASP gene mutations, protein expression, and phenotype. *Blood* 90:2680-2689.
- zur Stadt, U., J. Rohr, W. Seifert, F. Koch, S. Grieve, J. Pagel, J. Strauss, B. Kasper, G. Nurnberg, C. Becker, A. Maul-Pavicic, K. Beutel, G. Janka, G. Griffiths, S. Ehl, and H.C. Hennies. 2009. Familial hemophagocytic lymphohistiocytosis type 5 (FHL-5) is caused by mutations in Munc18-2 and impaired binding to syntaxin 11. *Am J Hum Genet* 85:482-492.

## Abbreviations

- CHS, Chédiak-Higashi syndrome
- DBS, dried blood spot
- DEPs, differentially expressed proteins
- DIA, data-independent acquisition
- FHL, familial hemophagocytic lymphohistiocytosis
- GO, Gene Ontology
- HPS, Hermansky-Pudlak syndrome
- IEIs, inborn errors of immunity
- IEMs, inborn errors of metabolism
- KRECs, kappa-deleting element recombination circles
- LC-MS/MS, liquid chromatography-assisted mass spectrometry
- NBS, newborn screening
- OMIM, Online Mendelian Inheritance in Man
- TRECs, T cell receptor excision circles
- WAS, Wiskott-Aldrich syndrome
- XLA, X-linked agammaglobulinemia

**Table 1. Details of 45 dried blood spot samples from 43 patients with inborn errors of immunity analyzed in the current study.**

Patient	Sample	Disease	Causative Gene	Causative Protein	Mutation		Protein levels encoded by the mutated gene (relative to ACTB)		
					Allele 1	Allele 2	Patient	Healthy newborns	Healthy adults
1	1	FHL type 3	UNC13D	Munc13-4	c.1596+1G>C	c.1255delC, p.Leu419Serfs*23	0.143	Mean, 10.9; SD, 2.6; Min, 5.9	Mean, 8.2; SD, 3.2; Min, 3.3
2	2				c.1596+1G>C	c.1240C>T, p.Arg414Cys	0.192		
3	3 <sup>b</sup>				c.1596+1G>C	c.118-308C>T, dysregulated transcription	0.357		
4	4 <sup>b</sup>				c.1596+1G>C	c.1545-2A>G	0.182		
5	5				c.754-1G>C	c.1596+1G>C	0.132		
	6 <sup>b</sup>						0.120		
6	7				c.1849-1G>C	c.1849-1G>C	0.270		
	8 <sup>b</sup>						0.360		
7 <sup>a</sup>	9	c.1596+1G>C	c.1596+1G>C	0.280					
8	10 <sup>b</sup>	c.754-1G>C	c.1596+1G>C	0.190					
9	11	FHL type 5	STXB2	Munc18-2	c.1197delC, p.Ala400Profs*18	c.1430C>T, p.Pro477Leu	1.120	Mean, 10.9; SD, 2.1; Min, 4.0	Mean, 7.8; SD, 2.1; Min, 4.8
10	12	FHL type 2	PRF1	perforin	c.658G>A, p.Gly220Arg	c.1090_1091delCT, p.Leu364Glufs*93	0.290	Mean, 0.33; SD, 0.28; Min, 0.28	Mean, 0.30; SD, 0.35; Min, 0.044
11	13				c.853_855delAAG, p.Lys285del	c.1090_1091delCT, p.Leu364Glufs*93	0.330		
12	14	HPS type 2	AP3B1	AP3B1	c.2546T>A, p.Leu849X	c.1122_1123insAG, p.Phe375Serfs*11	0.183	Mean, 2.93; SD, 0.88; Min, 1.08	Mean, 0.91; SD, 0.25; Min, 0.58
13	15				c.364C>T, p.Arg122X	c.2810-1G>T	1.080		
14	16				c.2546T>A, p.Leu849X	c.188T>A, p.Met63Lys	0.439		
15	17				c.1528_1529delTT, p.Leu510Valfs*8	NA	0.010		
16	18	CGD	CYBB	p91-phox	c.121dupT, p.Tyr41Leufs*62	NA	0.020	Mean, 2.63; SD, 1.36; Min, 0.72	Mean, 0.78; SD, 0.51; Min, 0.30
17	19				c.1031C>T, p.Ser344Phe	NA	0.010		
18	20				c.252G>A	NA	0.024		
19	21				c.810G>A, p.Trp270X	NA	0.015		
20	22				c.271C>T, p.Arg91X	NA	0.058		
21	23				CGD	CYBA	p22-phox		
22	24	CGD	NCF1	p47-phox	c.75_76delGT, p.Tyr26Hisfs*26	c.75_76delGT, p.Tyr26Hisfs*26	0.010	Mean, 0.18; SD, 0.099; Min, 0.025	Mean, 0.10; SD, 0.062; Min, 0.057
23	25	WAS	WAS	WASP	c.777+3_777+6delGAGT, splice error	NA	Below detection limit	Mean, 0.098; SD, 0.039; Min, 0.022	Mean, 0.056; SD, 0.025; Min, 0.024
24	26				c.132+1G>T, splice error	NA	Below detection limit		
25	27				c.982delC, p.Arg328Glyfs*117	NA	0.011		
26	28				c.961C>T, p.Arg321X	NA	Below detection limit		
27	29				c.1075C>A, p.Pro359Thr	NA	Below detection limit		
28	30				c.1574G>A, Arg525Gln	NA	0.005		
29	31	XLA	BTK	BTK	c.1856C>T, p.Pro619Leu	NA	0.004	Mean, 0.60; SD, 0.16; Min, 0.25	Mean, 0.39; SD, 0.17; Min, 0.26
30	32				c.862C>T, p.Arg288Trp	NA	0.017		
31	33				c.1921C>T, p.Arg641Cys	NA	0.029		
32	34				c.902-904delAAG, p.Glu301del	NA	0.055		
33	35				c.95T>C, p.Leu32Ser	NA	0.002		
34	36				c.1856C>T, p.Pro619Leu	NA	0.035		
35	37				SCID	ADA	ADA1		
36	38	X-SCID	IL2RG	IL2Rγ	c.218+2T>G	c.632G>A, p.Arg211His	0.108	NA	NA
37	39				c.865C>T, p.Arg289X	NA	Below detection limit		
38	40	CHS	LYST	LYST	c.374A>G, p.Tyr125Cys	NA	Below detection limit	Mean, 0.056; SD, 0.035; Min, 0.01	Mean, 0.025; SD, 0.018; Min, 0.075
39	41 <sup>b</sup>				c.1673dupT, p.Leu558Phefs*22	c.5506C>T, p.Arg1836X	0.003		
40	42				c.3393+1G>T	c.3393+1G>T	0.016		
41	43				c.5541_5542delAA, p.Arg1848Serfs*3	unknown	0.012		
42	44				c.5541_5542delAA, p.Arg1848Serfs*3	unknown	0.009		
43	45	c.5541_5542delAA, p.Arg1848Serfs*3	unknown	0.026					

FHL, familial hemophagocytic lymphohistiocytosis; HPS, Hermansky-Pudlak syndrome; CGD, chronic granulomatous disease; WAS, Wiskott-Aldrich syndrome; XLA, X-linked agammaglobulinemia; X-SCID, X-linked SCID; CHS, Chédiak-Higashi syndrome; NA, not applicable

<sup>a</sup>Fetal-onset case

<sup>b</sup>Samples obtained pre-symptomatically and have been used for public screening of metabolic diseases



**Table 2. Predicted and observed levels of mutated proteins.**

Disease	Causative gene	Mutated allele	Mutation type	Pathogenicity determined by Mutation effect predictor <sup>a</sup>	NMD probability <sup>b</sup>	Predicted structural stability <sup>c</sup>	Observed protein loss
FHL type 3	UNC13D	c.1240C>T, p.Arg414Cys	Missense mutation	Deleterious (aggregated score, 0.856)		3.28	Yes
		c.1255delC, p.Leu419Serfs*23	Frameshift		NMD		Yes
		c.1596+1G>C	Splicing anomaly				Yes
		c.118-308C>T	Transcriptional dysregulation		Predicted benign		Yes
		c.754-1G>C	Splicing anomaly		Exon skip 105 bp		Yes
		c.1849-1G>C	Splicing anomaly		Exon skip 144 bp		Yes
FHL type 5	STXBP2	c.1545-2A>G	Splicing anomaly		Exon skip 54 bp		Yes
		c.1430C>T, p.Pro477Leu	Missense mutation	Deleterious (aggregated score, 0.857)		4.29	Yes
FHL type 2	PRF1	c.1197delC, p.Ala400Profs*18	Frameshift		NMD		Yes
		c.658G>A, p.Gly220Arg	Missense mutation	Deleterious (aggregated score, 0.98)		22.8	No
		c.853_855delAAG, p.Lys285del	deletion	benign			No
HPS type 2	AP3B1	c.1090_1091delCT, p.Leu364Glufs*93	Frameshift		NMD escaping		No
		c.188T>A, p.Met63Lys	Missense mutation	Uncertain (aggregated score, 0.67)		5.13	No
		c.1122_1123insAG, p.Phe375Ser*11	Frameshift		NMD		Yes
		c.2546T>A, p.Leu849X	Nonsense mutation		NMD		Yes
		c.364C>T, p.Arg122X	Nonsense mutation		NMD		Reduced but detectable
CGD	CYBB	c.2810-1G>T	Splicing anomaly		Exon skip 85 bp		Reduced but detectable
		c.1031C>T, p.Ser344Phe	Missense mutation	Deleterious (aggregated score, 0.99)		12.5	Yes
		c.1528_1529delTT, p.Leu510Valfs*8	Frameshift		NMD		Yes
		c.121dupT, p.Tyr41Leufs*62	Frameshift		NMD		Yes
		c.810G>A, p.Trp270X	Nonsense mutation		NMD		Yes
		c.271C>T, p.Arg91X	Nonsense mutation		NMD		Yes
		c.252G>A, p. Ala84Ala	Splicing anomaly	Deleterious (aggregated score, 0.80)	splicing anomaly		Yes
CGD	CYBA	c.70G>A, p.Gly24Arg	Missense mutation	Deleterious (aggregated score, 0.87)		6.54	Yes
		c.7C>T, p.Gln3X	Nonsense mutation		NMD escaping		Yes
CGD	NCF1	c.75_76delGT, p.Tyr26Hisfs*26	Frameshift		NMD escaping		Reduced but detectable
WAS	WAS	c.1075C>A, p.Pro359Thr	Missense mutation	Uncertain (aggregated score, 0.54)		0.37	Yes
		c.982delC, p.Arg328Glyfs*117	Frameshift		NMD		Reduced but detectable
		c.961C>T, p.Arg321X	Nonsense mutation		NMD		Yes
		c.777+3_777+6delGAGT	Splicing anomaly				Yes
		c.132+1G>T	Splicing anomaly				Yes
XLA	BTK	c.95T>C, p.Leu32Ser	Missense mutation	Deleterious (aggregated score, 0.99)		5.26	Yes
		c.862C>T, p.Arg288Trp	Missense mutation	Deleterious (aggregated score, 0.88)		1.81	Yes
		c.1574G>A, Arg525Gln	Missense mutation	Deleterious (aggregated score, 0.88)		0.02	Yes
		c.1856C>T, p.Pro619Leu	Missense mutation	Deleterious (aggregated score, 0.87)		6.19	Yes
		c.1921C>T, p.Arg641Cys	Missense mutation	Deleterious (aggregated score, 0.88)		2.95	Yes
		c.902-904delAAG, p.Glu301del	Deletion				Yes
SCID	ADA	c.632G>A, p.Arg211His	Missense mutation	Deleterious (aggregated score, 0.99)		9.20	No
		c.218+2T>G	Splicing anomaly				No
X-SCID	IL2RG	c.374A>G, p.Tyr125Cys	Missense mutation	Deleterious (aggregated score, 0.86)		3.36	ND
		c.865C>T, p.Arg289X	Nonsense mutation		NMD escaping		ND
CHS	LYST	c.5506C>T, p.Arg1836X	Nonsense mutation		NMD		Reduced but detectable
		c.1673dupT, p.Leu558Phefs*22	Frameshift		NMD		Reduced but detectable
		c.5541_5542delAA, p.Arg1848Serfs*3	Frameshift		NMD		Reduced but detectable
		c.3393+1G>T	Splicing anomaly				No

FHL, familial hemophagocytic lymphohistiocytosis; HPS, Hermansky-Pudlak syndrome; CGD, chronic granulomatous disease; WAS, Wiskott-Aldrich syndrome; XLA, X-linked agammaglobulinemia; X-SCID, X-linked SCID; CHS, Chédiak-Higashi syndrome

<sup>a</sup>Aggregated prediction score obtained using Franklin (<https://franklin.genoox.com/clinical-db/home>).

<sup>b</sup>Nonsense mediated decay (NMD) prediction using Variant Effect Predictor.

<sup>c</sup>Protein structure was predicted to be destabilized with a score > 1.

## Figure legends

### **Figure 1. Proteomics analysis distinguishes healthy newborn dried blood spot (DBS) samples from those of adults.**

(A) Clustered heatmap of Pearson correlation coefficient values between all healthy newborn and adult DBS sample pairs. Dark red denotes higher correlation and dark blue denotes lower correlation (Pearson  $r = 0.82-0.96$ ). (B) Principal component analysis of healthy newborn and adult DBS sample pairs. Red denotes newborn (NB) and blue denotes adult DBS samples.

### **Figure 2. Proteins differentially expressed between healthy newborn and adult samples.**

(A) Heatmap of hierarchical clustering of all differentially expressed proteins (DEPs) detected by comparison between healthy newborn and adult samples. Red, upregulated DEPs; blue, downregulated DEPs. (B) Volcano plot showing DEPs in dried blood spot samples from healthy newborns compared with those from healthy adults. (C) Protein-protein interaction network analysis illustrating the top 100 upregulated and downregulated proteins ( $q\text{-value} < 0.05$ ), contrasting healthy newborns and adults; network generated in Cytoscape using the STRING database. Node size corresponds to P value, while colors (red and blue) signify DEPs; red and blue circles represent the 'Hemoglobin complex' and 'Complement and coagulation' Gene Ontology (GO) terms, respectively. (D) GO enrichment analysis of DEPs. Bar charts showing the top 10 GO terms for biological processes upregulated (left) and downregulated (right) in newborn samples relative to adult samples.

### **Figure 3. Expression of disease-responsible proteins in samples from patients with inborn errors of immunity (IEIs).**

Proteomics evaluation of disease-responsible proteins in samples from healthy newborns (HNB), healthy adults (HA), and patients with IEIs, including: familial hemophagocytic lymphohistiocytosis (FHL) type 2, 3, and 5 (FHL2:  $n = 2$ , FHL3:  $n = 8$ , and FLH5:  $n = 1$ , respectively); p91-phox, p22-phox, and p47-phox deficiencies ( $n = 6$ , 1, and 1, respectively); Hermansky-Pudlak syndrome type 2 (HPS2:  $n = 3$ ); Wiskott-Aldrich syndrome (WAS:  $n = 5$ ); X-linked agammaglobulinemia (XLA:  $n = 7$ ); Chédiak-Higashi

syndrome (CHS: n = 5); adenosine deaminase deficiency (ADA: n = 2); common  $\gamma$  chain deficiency (IL2RG: n = 2). All data were standardized relative to the expression of  $\beta$ -actin in the same sample. Levels of (A) Munc13-4, (B) Munc18-2, (C) Syntaxin-11, (D) perforin, (E) p91-phox, (F) p22-phox, (G) p47-phox, (H) WASP, (I) BTK, (J) LYST, (K) AP3B1, and (L) ADA in samples from HNB, HA, and patients with IEs. Open symbols, samples obtained from pre-symptomatic newborn patients with IEs; half-closed symbols, results under the limit of detection plotted at approximate lower detection limits. Statistical differences among groups were analyzed using one-way ANOVA on ranks (the Kruskal-Wallis test), followed by Dunn's multiple comparisons. Data represent the mean  $\pm$  SD. \*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001.

**Figure 4. Identification of marker proteins associated with molecular phenotypes of inborn errors of immunity (IEs).**

Gene ontology enrichment analysis of differentially expressed proteins (DEPs) (left) and volcano plots showing DEPs (right) in dried blood spot (DBS) samples from patients with (A) Hermansky-Pudlak syndrome type 2 (HPS2: n = 3), (B) Chédiak-Higashi syndrome (CHS: n = 5), (D) SCID: n = 4, and (F) Wiskott-Aldrich syndrome (WAS: n = 5) relative to samples from healthy newborns (HN) (left) and healthy adults (HA) (right). Levels of (C) cathepsin G, myeloperoxidase, and ELANE, (E) CD3E, and (G) ITGA2B (CD41) and Rab27B in DBS samples from HNB, HA, and patients with IEs. Statistical differences among groups were analyzed using one-way ANOVA on ranks (the Kruskal-Wallis test), followed by Dunn's multiple comparisons. Data represent the mean  $\pm$  SD. \*P < 0.05, \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001. #, Top Gene Ontology terms commonly altered in patient samples compared to both newborn and adult control groups. Red dots in volcano plots indicate proteins shared in biological pathways commonly altered in patient samples relative to those from HN and HA.

**Figure 5. Comparison of proteomics data from patients with inborn errors of immunity with those from healthy newborns and adults.**

Heatmap showing hierarchical clustering of the levels of 42 proteins that are disease-responsive or associated

with molecular phenotypes in dried blood spot samples from healthy newborns, healthy adults, and patients with inborn errors of immunity. Red, relatively higher expression; blue, relatively lower expression.

### **Supplementary figure legends**

#### **Figure S1. Wikipathway map of the Complement and coagulation cascades.**

Pathways were retrieved and visualized from WikiPathways using Cytoscape, highlighting downregulated proteins ( $P < 0.05$ ) in healthy newborns relative to adults. Blue node colors denote differentially expressed proteins and fold changes in their expression. Blue characters indicate metabolite molecules.

#### **Figure S2. Quantification of proteins responsible for various genetic disorders in neonatal dried blood spot (DBS) samples.**

(A) Venn diagram of proteins identified in DBS samples. OMIM, Online Mendelian Inheritance in Man. (B) Levels of 124 proteins listed in the 2022 International Union of Immunological Societies classifications of inborn errors of immunity in DBS samples from healthy newborns (blue) and healthy adults (red).

#### **Figure S3. Munc13-4 is conspicuously reduced in dried blood spot (DBS) samples from patients with familial hemophagocytic lymphohistiocytosis type 3 (FHL3).**

Volcano plots showing the differentially expressed proteins in DBS samples from patients with FHL3 that were obtained (A) before and (B) after disease onset compared to samples from healthy newborns (left) and from healthy adults (right).

#### **Figure S4. Flow cytometric analysis of perforin expression in natural killer cells from patients with familial hemophagocytic lymphohistiocytosis type 2.**

Fluorescence-activated cytometric analysis of perforin expression in CD3<sup>+</sup> peripheral lymphocytes obtained from patients 10 and 11 relative to healthy controls.

## Supplemental table legends

### **Table S1. List of proteins identified in dried blood spot samples analyzed in this study.**

The list includes identified proteins with peptide and protein false discovery rates < 1%. Proteins shown in black are those identified by unique peptide sequences that were clearly distinguishable from others. Proteins shown in blue represent grouped master proteins. Proteins shown in grey represent grouped proteins. A quantitative value of “0” indicates that the protein was detected, but did not reach a level sufficient for quantification.

### **Table S2. Proteins identified in dried blood spot samples that are listed in the Online Mendelian Inheritance in Man (OMIM) database.**

List of proteins identified in every healthy newborn sample that are included in the OMIM database. The list includes identified proteins with peptide and protein false discovery rates < 1%. AD, autosomal dominant; AR, autosomal recessive; ARLOF, autosomal recessive loss of function; XL, X-linked; IUIS2022, 2022 Classification of Human Inborn Errors of Immunity by the International Union of Immunological Societies Expert Committee.

Figure 1

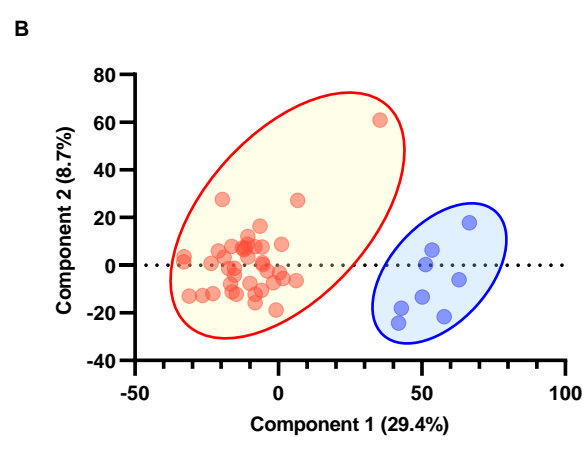
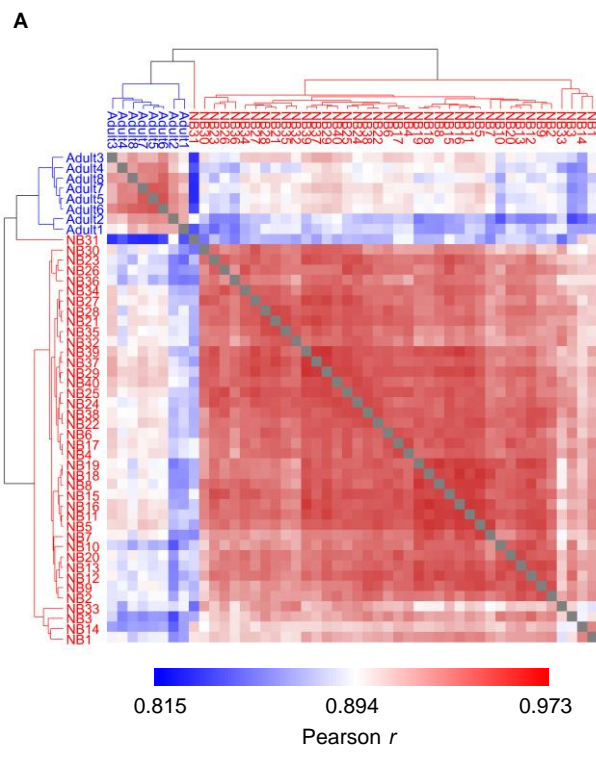
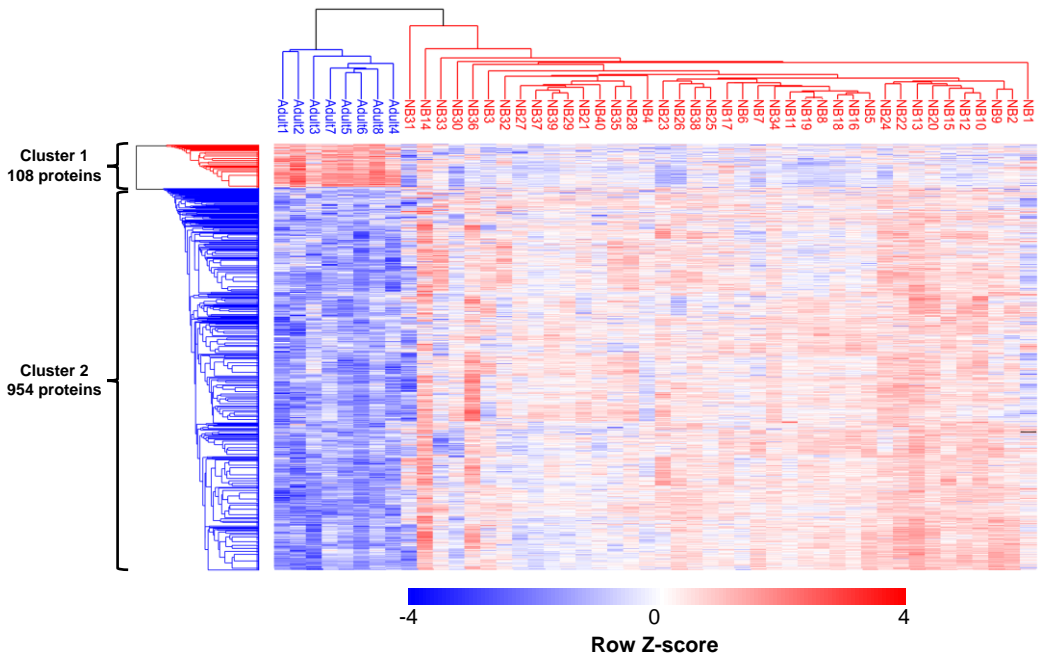
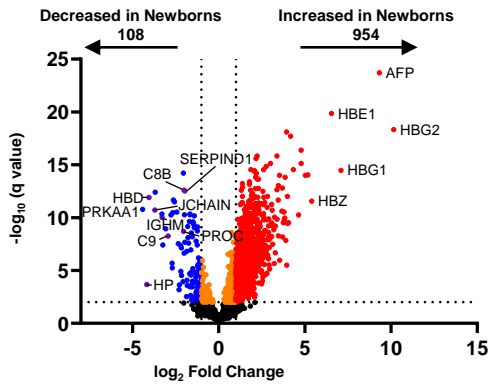


Figure 2

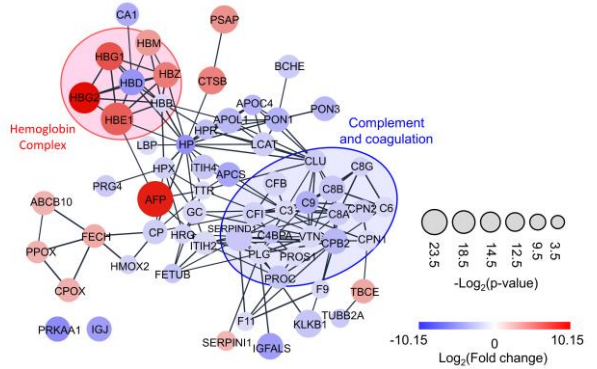
A



B



C



D

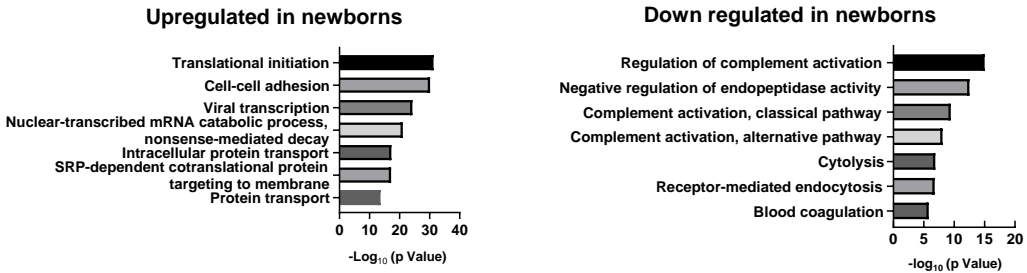
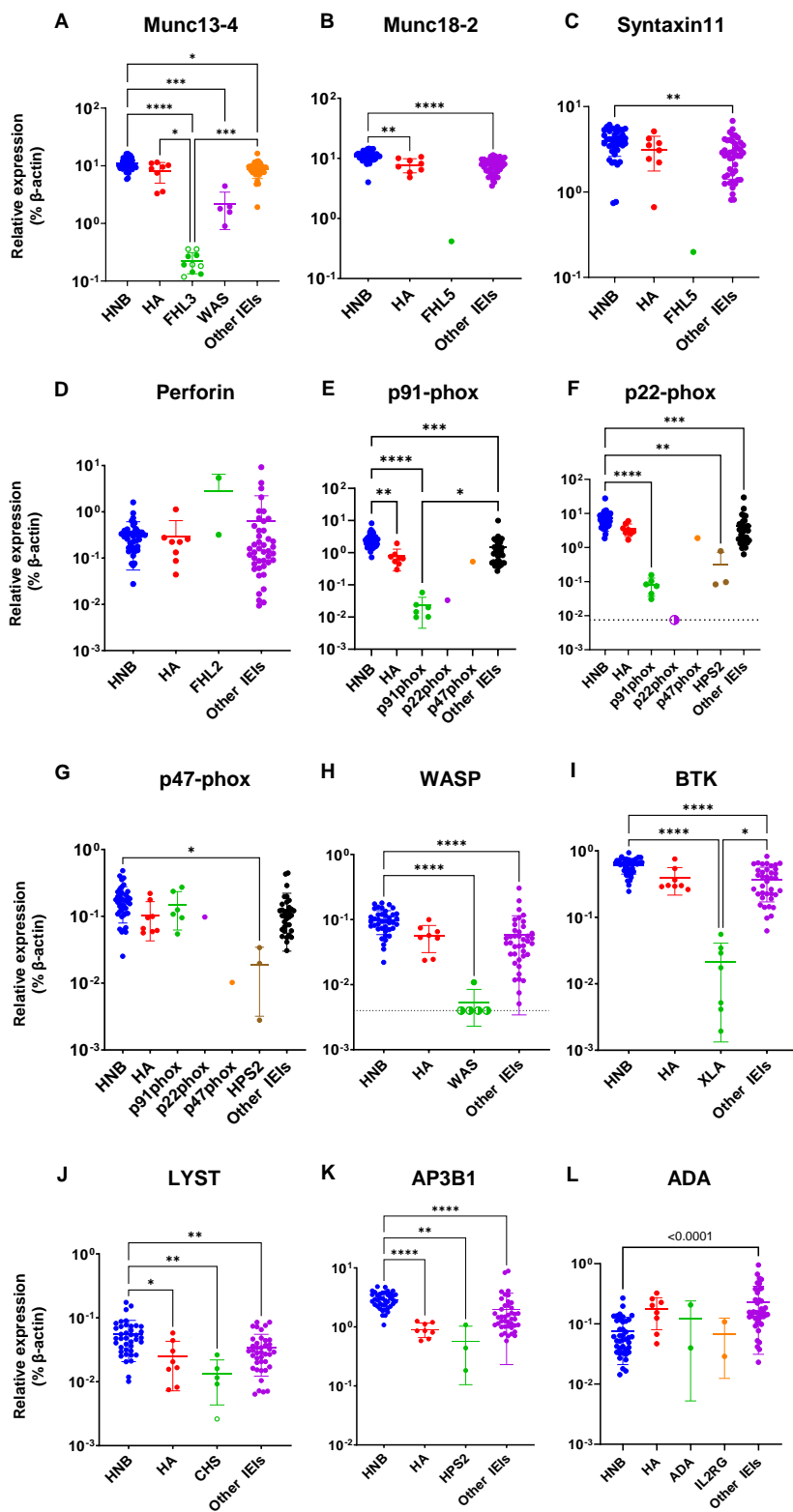




Figure 3



**Figure 4**

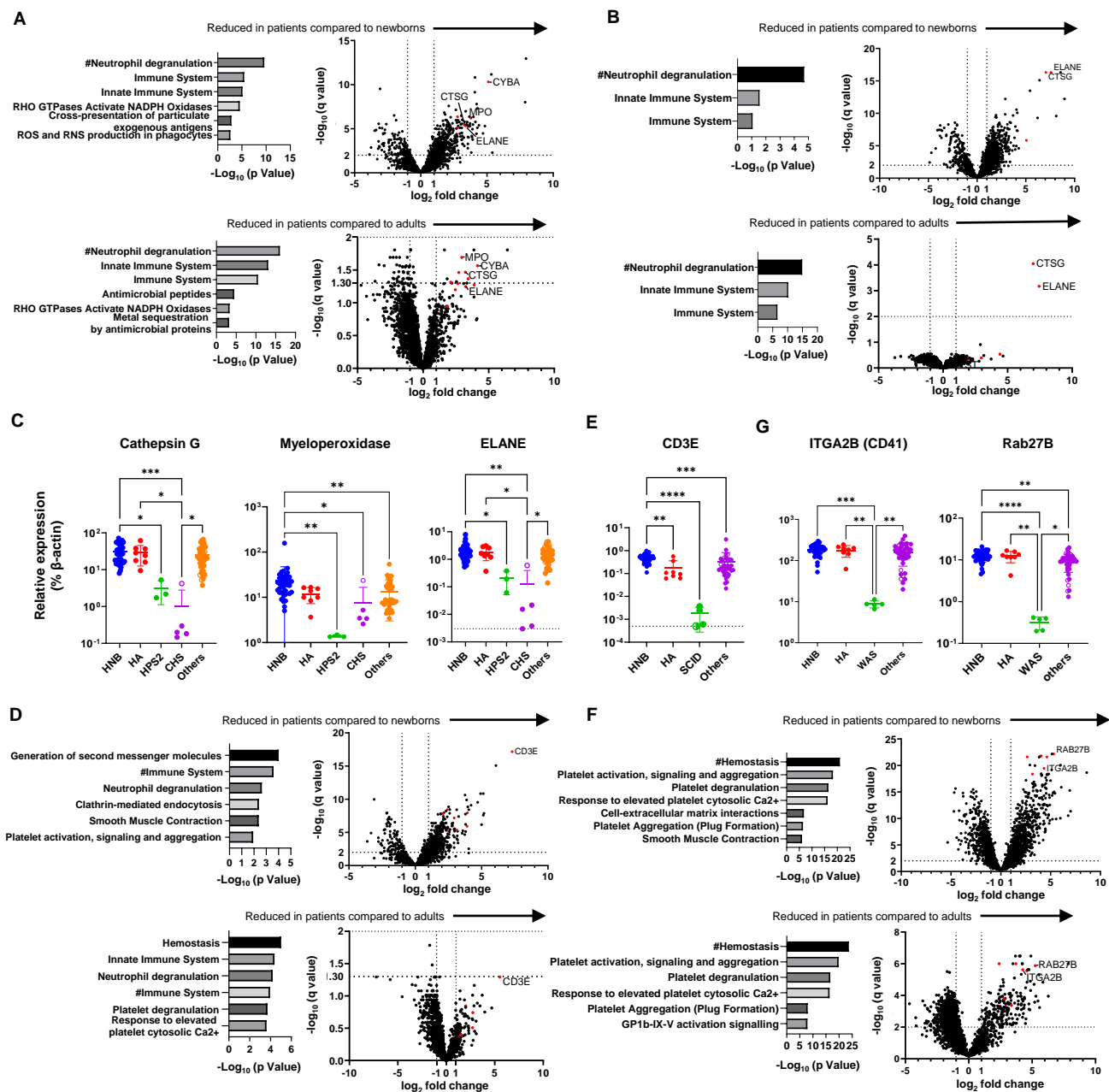
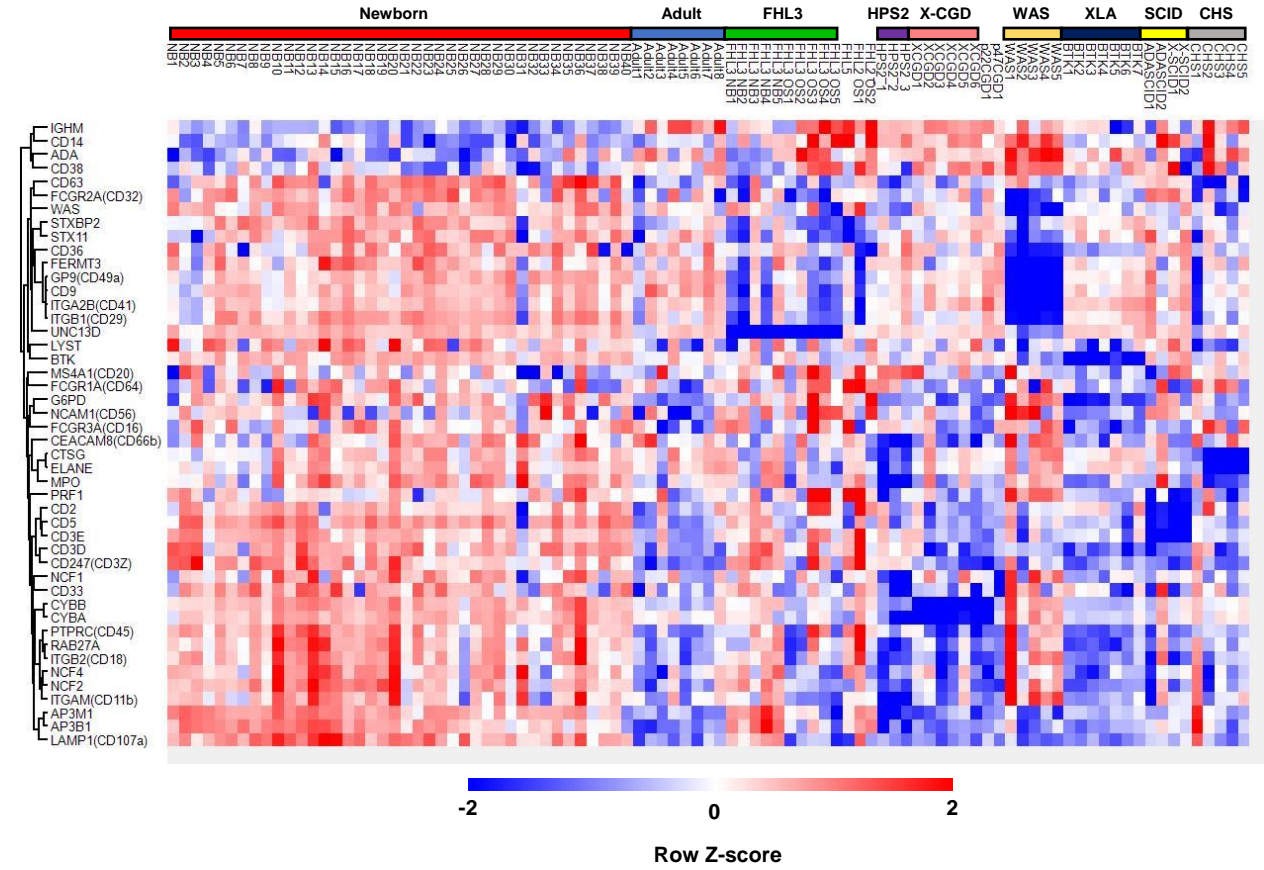


Figure 5



**Figure S1**

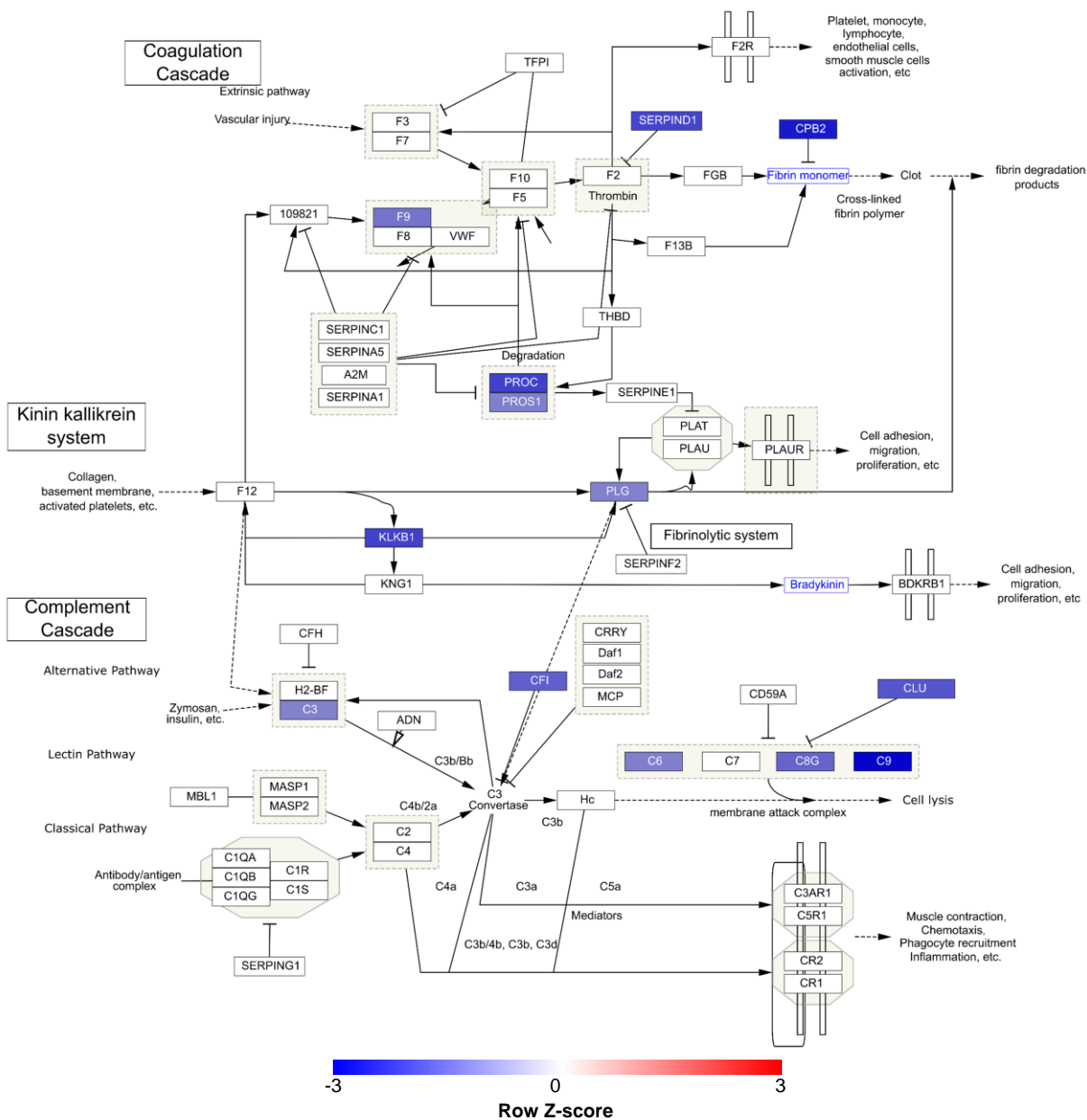
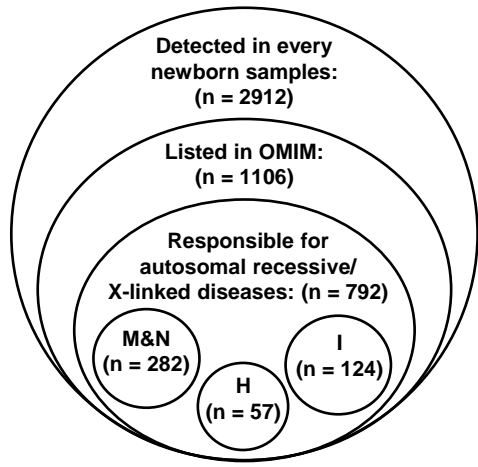


Figure S2

A



M&N: Metabolic/Neurologic disease  
H: Hematologic disease  
I: Listed in 2022 IUIS Classification for Human Inborn Errors of Immunity

B

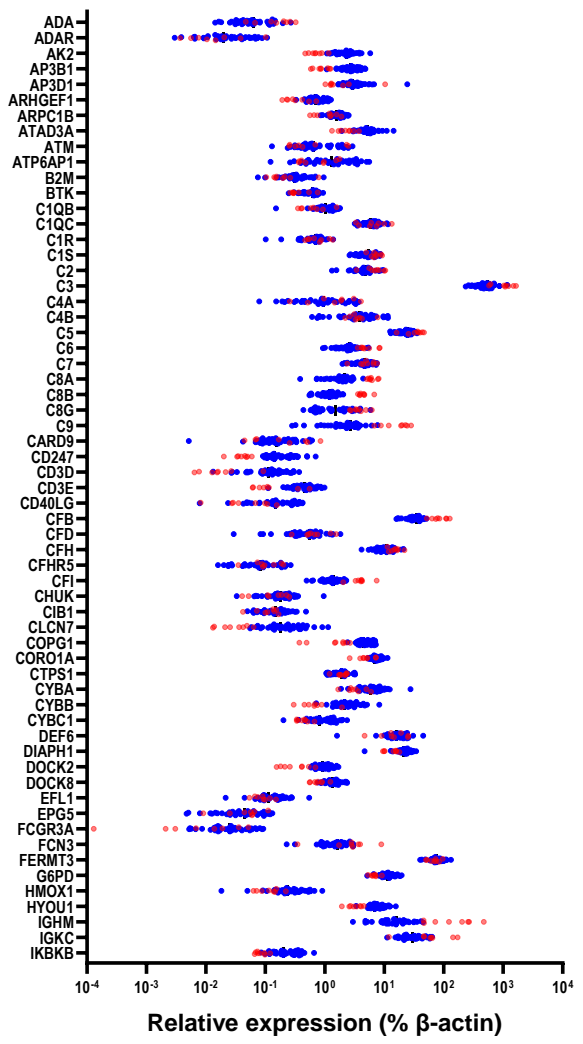
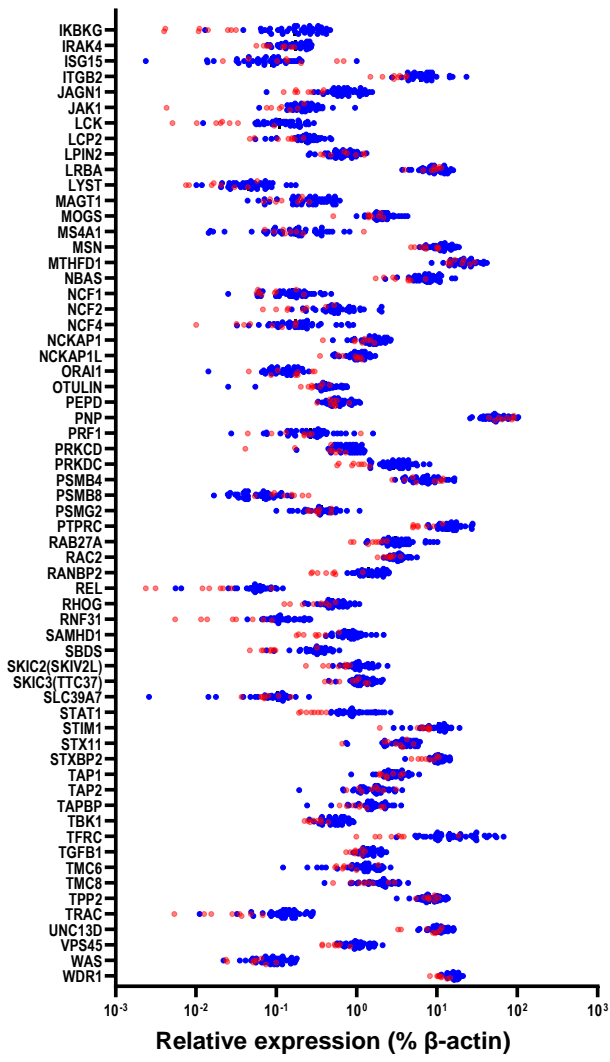
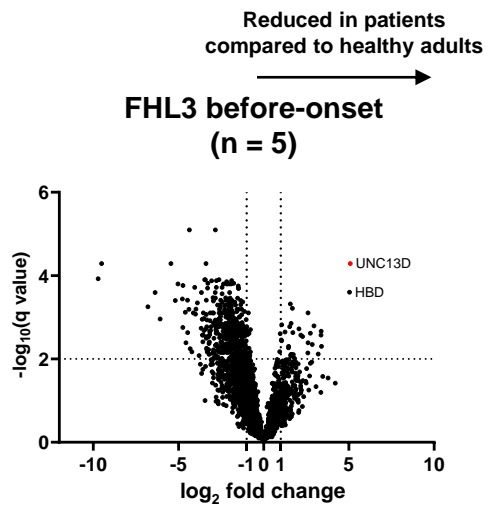
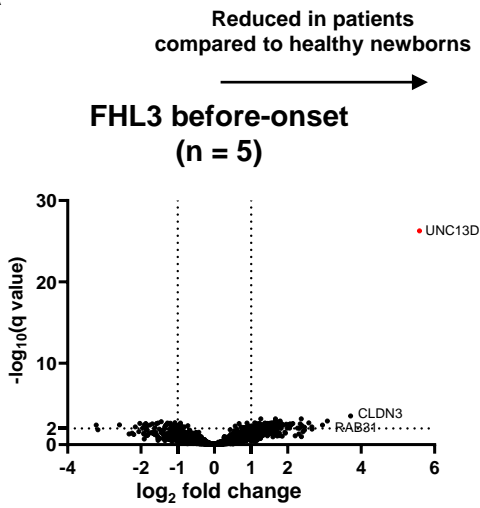


Figure S3

A



B

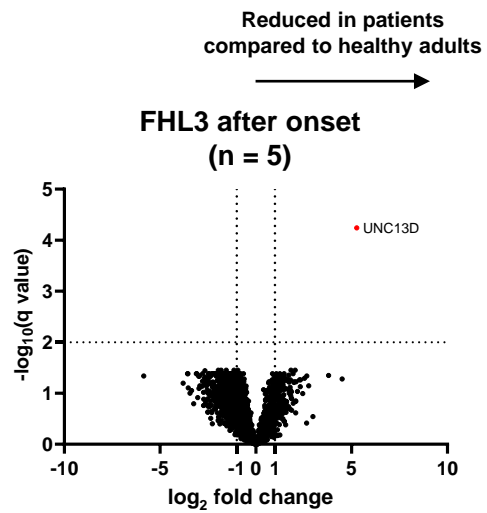
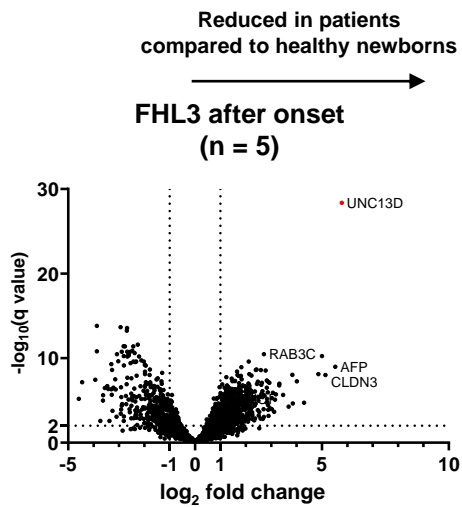


Figure S4

Gated on CD3<sup>+</sup> lymphocyte population

