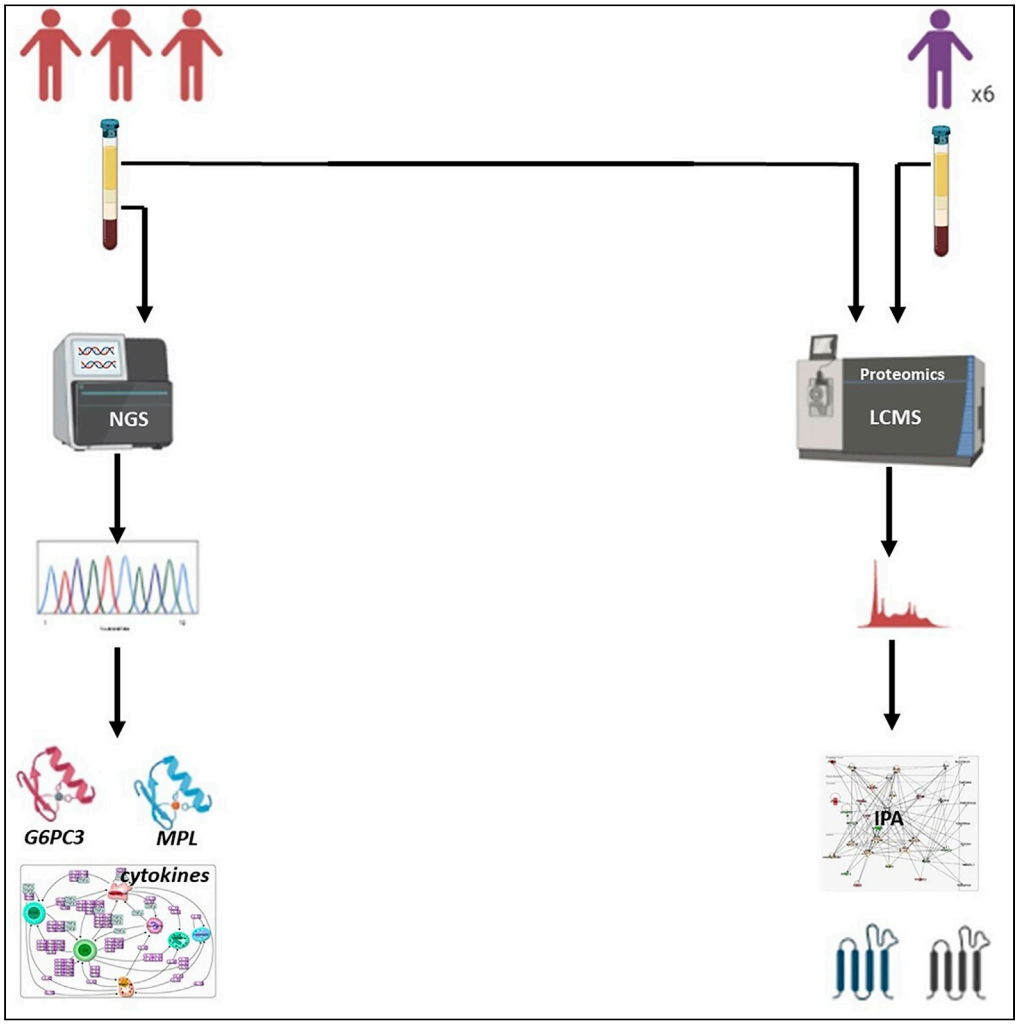
iScience

CellPress
OPEN ACCESS

Article

Comprehensive multi-omics analysis of G6PC3 deficiency-related congenital neutropenia with inflammatory bowel disease



Majed Dasouki, Ayodeele Alaiya, Tanziel ElAmin, ..., Stefan T. Arold, Mahmoud Aljurf, Syed Osman Ahmed

majed.dasouki.md@adventhealth.com

HIGHLIGHTS

Multi-omics approaches identify unique signatures

Whole-exome sequencing reveals distinct cytokine profiles

Expression of GATA4, PF4, and LST1 is dysregulated

Dasouki et al., iScience 24, 102214
March 19, 2021
https://doi.org/10.1016/j.isci.2021.102214

iScience

CellPress
OPEN ACCESS

Article

Comprehensive multi-omics analysis of G6PC3 deficiency-related congenital neutropenia with inflammatory bowel disease

Majed Dasouki,[1,3,9,*] Ayodeele Alaiya,[2] Tanziel ElAmin,[1] Zakia Shinwari,[2] Dorota Monies,[1] Mohamed Abouelhoda,[3] Amjad Jabaan,[3] Feras Almourfi,[3] Zuhair Rahbeeni,[4] Fahad Alsohaibani,[5] Fahad Almohareb,[6] Hazzaa Al-Zahrani,[6] Francisco J. Guzmán Vega,[7] Stefan T. Arold,[7,8] Mahmoud Aljurf,[6] and Syed Osman Ahmed[6]

SUMMARY

Autosomal recessive mutations in *G6PC3* cause isolated and syndromic congenital neutropenia which includes congenital heart disease and atypical inflammatory bowel disease (IBD). In a highly consanguineous pedigree with novel mutations in *G6PC3* and *MPL*, we performed comprehensive multi-omics analyses. Structural analysis of variant G6PC3 and MPL proteins suggests a damaging effect. A distinct molecular cytokine profile (cytokinome) in the affected proband with IBD was detected. Liquid chromatography-mass spectrometry-based proteomics analysis of the G6PC3-deficient plasma samples identified 460 distinct proteins including 75 upregulated and 73 downregulated proteins. Specifically, the transcription factor GATA4 and LST1 were downregulated while platelet factor 4 (PF4) was upregulated. GATA4 and PF4 have been linked to congenital heart disease and IBD respectively, while LST1 may have perturbed a variety of essential cell functions as it is required for normal cell-cell communication. Together, these studies provide potentially novel insights into the pathogenesis of syndromic congenital G6PC3 deficiency.

INTRODUCTION

Several familial syndromes can present with neutropenia including severe congenital neutropenia, or Kostmann syndrome, cyclic neutropenia, Hermansky-Pudlak syndrome, warts, hypogammaglobulinemia, infections, and myelokathexis syndrome, Barth syndrome (X-linked dilated cardiomyopathy with neutropenia), poikiloderma with neutropenia, and glycogen storage disease type 1b, among others (Banka, 2016). The severity and pattern of neutropenia, the nature of complications, and the co-existing non-hematological phenotypes all assist in making the likely correct clinical diagnosis.

Mutations in several genes including *ELANE* and *HAX1* (being the most common) have been identified as causes of severe congenital neutropenia (SCN) (Dale et al., 2000; Klein et al., 2006). The advent of next-generation sequencing (NGS)-based mutation analysis allows comprehensive testing of patients with neutropenia of different etiologies. Consequently, more patients with monogenic SCN neutropenia including G6PC3 deficiency are likely to be diagnosed.

G6PC3 (glucose-6-phosphatase, catalytic, subunit 3) mutations have been identified as a cause of congenital neutropenia, initially in 2 consanguineous pedigrees (Boztug et al., 2009). Since then, over 90 cases have been described, and the spectrum of disease phenotype has been reviewed (Banka, 2016; Banka and Newman, 2013). The glucose-6-phosphatase enzyme is involved in glycogenolysis, with 3 known catalytic subunits being present in humans. *G6PC3* encodes the glucose-6-phosphatase enzyme which is expressed ubiquitously in humans, in contrast to G6PC1 (expressed in the small intestine) and G6PC2 (expressed in the liver). The *G6PC3* gene maps to chromosome 17q21 (GRCh38:44,070,699-44,076,343) and consists of 6 exons.

G6PC3 deficiency may manifest as (i) a non-syndromic SCN, (ii) the so-called classic G6PC3 deficiency with SCN and cardiovascular and/or urogenital abnormalities, or (iii) the more severe form, known as the Dursun

[1]Department of Genetics, Department of Pathology & Laboratory Medicine King Faisal Specialist Hospital, and Research Center, Riyadh, Saudi Arabia

[2]Department of Stem Cell Therapy. Proteomics Program. King Faisal Specialist Hospital and Research Center, MBC-03-30. PO Box 3354, Riyadh 11211, Saudi Arabia

[3]Saudi Human Genome Program. King Abdulaziz Center for Science & Technology, Riyadh, Saudi Arabia

[4]Department of Medical Genetics, King Faisal Specialist Hospital, and Research Center, Riyadh, Saudi Arabia

[5]Department of Internal Medicine, King Faisal Specialist Hospital, and Research Center, Riyadh, Saudi Arabia

[6]Adult hematology/BMT, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia

[7]King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Division of Biological and Environmental Sciences and Engineering (BESE), Thuwal, 23955-6900, Saudi Arabia

[8]Centre de Biochimie Structurale, CNRS, INSERM, Université de Montpellier, 34090 Montpellier, France

[9]Lead contact

*Correspondence: majed.dasouki.md@adventhealth.com

https://doi.org/10.1016/j.isci.2021.102214

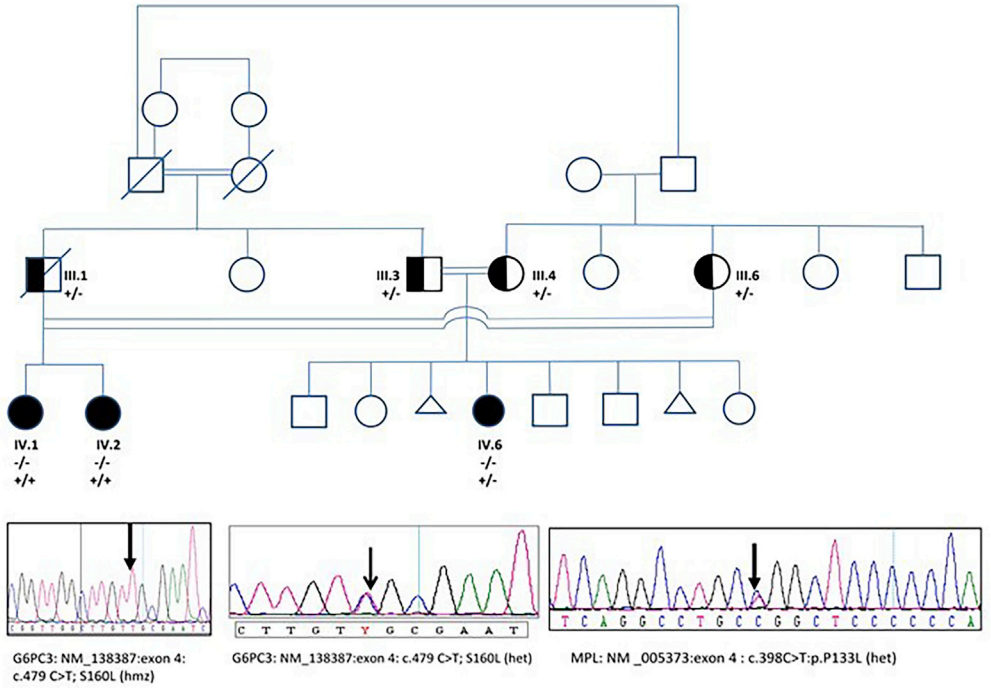


Figure 1. Family pedigree and Sanger sequencing

The large consanguineous family with G6PC3 deficiency has 3 affected patients (2 sisters: IV-1, 2 and their double cousin, IV.6). The Sanger sequencing chromatograms show homozygosity (indicated −/−) and heterozygosity (+/−) for the novel *G6PC3* mutation. The novel *MPL* heterozygous mutation is also shown as (+/−) in multiple family members including IV.6.

syndrome, involving non-myeloid hematopoietic cell lineages and neonatal pulmonary hypertension and thymic hypoplasia (Banka et al., 2010; Dursun et al., 2009). Inflammatory bowel (like) disease has been reported in some cases and small series with G6PC3 deficiency (Begin et al., 2012; Cullinane et al., 2011; Desplantes et al., 2014; Glasser et al., 2016). Hypo-glycosylation of $gp91^{phox}$, the electron-transporting component of the NADPH oxidase, had been demonstrated in patients with neutropenia and G6PC3 deficiency (Hayee et al., 2011). Failure to eliminate the phosphorylated glucose analog (1,5-anhydroglucitol-6-phosphate; 1,5AG6P) was found to cause neutropenia in patients with G6PC3 deficiency (Veiga-da-Cunha et al., 2019). Activating mutations in *THPO* and *MPL* are also known to cause thrombocytosis which may predispose to thromboembolism (Dasouki et al., 2015).

Proteomics studies are being increasingly used in the search for biomarkers for many human diseases including inflammatory bowel disease (IBD). Several proteomics studies suggested different protein signatures which were identified in serum, feces, and colonic epithelia of affected patients (summarized in Bennike et al., 2014). Examples of these diverse biomarkers include anti-*Saccharomyces cerevisiae* antibodies, perinuclear anti-neutrophil cytoplasmic antibody, C-reactive protein, calprotectin, lactoferrin, annexin A1, lymphocyte cytosolic protein 1 (aka L-plastin), proteasome activator subunit 1, and many others.

In this study, we performed extensive whole-exome and proteome analyses (with special emphasis on the cytokinome) in 3 consanguineous family members from Southern Saudi Arabia (Figure 1) with G6PC3 deficiency and secondary (atypical) IDB and thrombocytosis. We identified novel mutations in *G6PC3*, *MPL* as well as multiple cytokines and novel proteomic biomarkers. This complex and unique multi-omic fingerprint may help improve the understanding of the mechanisms and help identify novel treatment opportunities.

RESULTS

Clinical and molecular analyses

The clinical features and results of routine clinical investigations are summarized in Tables 1 and 2. While all 3 G6PC3-deficient patients had severe failure to thrive and very low body mass index, only patient IV.6 had been diagnosed with IBD (atypical Crohn disease). Recurrent chest infections and resultant bronchiectasis and

Table 1. Clinical characteristics of individuals affected with G6PC3 deficiency.

Clinical manifestations		Patient IV.1	Patient IV.2	Patient IV.6
Infections		Skin and chest	Skin and chest	Skin; otitis media; recurrent *E.coli* UTI and abdominal sepsis
GI		Childhood FTT, diarrhea/steatorrhea; no bowel symptoms in adulthood	Childhood FTT, diarrhea/steatorrhea; no bowel symptoms in adulthood	Childhood FTT, diarrhea/steatorrhea; severe IBD with stricture disease
Respiratory		Bronchiectasis, class 1 pulmonary hypertension	Bronchiectasis, respiratory failure, class IV pulmonary hypertension; lung transplantation candidate	Not significant
Cardiac		Fenestrated ASD with bidirectional flow (mainly left to right); moderately dilated RA, RV, and main PA	Dilated RV with RVH, TR, raised RVSP	ASD repaired at age 8
Skeletal		Generalized osteopenia; underdeveloped sinuses	Microcephaly, mid face hypoplasia; thoracic spine kyphosis; generalized osteopenia	Oligoarthritis, generalized osteopenia, facial dysmorphism
Growth	Weight	25.9 kg at age 17	29.4 at age 16	17kg at age 20
	BMI	14	12.7	11.6
	Height	<3rd centile	<3rd centile	<3rd centile

ASD, atrial septal defect; BMI, body mass index; FTT, failure to thrive; IBD, inflammatory bowel disease; PA, pulmonary artery; RV, right ventricle; RVH, right ventricular hypertrophy; RVSP, right ventricular systolic pressure; TR, tricuspid regurgitation.

secondary pulmonary hypertension occurred in siblings IV.1 and 2 but not their cousin (IV.6). All 3 patients had congenital heart disease (atrial septal defect) requiring repair in patient IV.6. Pan-T-cell lymphopenia was found in two patients (IV.2, and IV.6). Interestingly, serum IgE deficiency in combination with hyper-gammaglobulinemia was present in all 3 G6PC3-deficient patients in this family. However, no mutations in *IGHE* (which encodes IgE) or its receptors (*FCER1A*, *MS4A2*, *FCER1G*; *FCER2*) were found. A common variant in *MS4A2* (which encodes FCER1B) previously thought to predispose to atopy was found in 2 of the 3 patients.

Sanger sequencing of the *CFTR* and *SBDS* genes was negative. Pathogenic or likely pathogenic variants in the 3 patients' exomes are listed in Table 3. All 3 patients were homozygous for a novel homozygous missense mutation in exon 4 of G6PC3 (c.479C > T, p.Ser160Leu) which was also confirmed by Sanger sequencing and segregation analysis (Figure 1). This mutation has not previously been reported in the homozygous state in various public databases including dbSNP, 1000G, ESV, ExAC, gnomAD, and our private Saudi Genome Program database (https://www.saudigenomeprogram.org/). Only a single (African) heterozygous individual (allele frequency: 0.000004061) was recently identified in the gnomAD database (https://gnomad.broadinstitute.org/).

Also, a homozygous c.154C > T (p.Arg52Cys) missense variant in exon 1 of *MBL2* was observed in patient IV.1. In silico analysis of this variant also predicted it to be damaging and has been associated with MBL2 deficiency which is reported to be common in Caucasians (Choteau et al., 2016).

Additionally, the exome of patient IV.6 (but not her cousins IV.1 and 2) showed a very rare (allele frequency: 3.255×10^{-5}), possibly damaging heterozygous *MPL* variant (Chr1:43804948:43804948: NM_005373:exon4:c.398C > T, p.P133L) which was also confirmed by Sanger sequencing and segregation analysis (Figure 1). While all 3 patients had intermittent thrombocytosis, it was most marked in patient IV.6.

Interestingly, we also detected a novel homozygous mutation in *HBG1* (NM_000559: c.*6_*3delinsCTCT) in all 3 affected G6PC3-deficient patients in addition to another novel homozygous splicing mutation in *KLF1* (NM_006563:exon2:12996130-12996956:-14:840; GC > –) in patient IV.6 only which probably explains the elevated HbF levels observed in these patients.

In silico structural analysis of G6PC3 and MPL mutations

G6PC3 is a 346-amino acid protein, predicted to contain nine transmembrane helices that span the endoplasmic reticulum (ER) membrane, with the active site lying inside the ER lumen (Ghosh et al., 2004).

Table 2. Routine laboratory investigations

Laboratory results	Patient IV.1	Patient IV.2	Patient IV.6
Neutrophil count x10^9/L, Median (range)	Pediatric: 0.5 (0.11–2.25)	Pediatric: 0.31 (0.09–1.76)	Pediatric: 1.39 (0.8–6.2)
	Adult: 0.34(0.24–1.77)[a]	Adult 2.385 (0.4–4.57)[a]	Adult: 1.65 (0.67–17.2)
Lymphocyte x10^9/L, median (range)	2.95 (1.59–3.27)	2.4 (1.2–8.5)	0.49 (0.14–2.17)
Platelet counts, x10^9/L, median (range)	498 (344–777)	392 (179–738)	498 (150–1074)
Hb g/L, median (range)	108 (84–130)	100 (77–127)	97 (58–147)
Bone marrow	Pediatric: Unremarkable with no maturation arrest	Pediatric: marked left-shifted granulopoiesis and active erythropoiesis and megakaryopoiesis	–
	Adult: hypoplastic myelopoiesis and left shift with abnormal segmentation of myeloid precursors	Adult: myeloid hyperplasia, left shifted and minimal maturation and some with increased granulation	Adult: Left shifted myelopoiesis
Bone marrow karyotype	46, XX	46, XX	46, XX
IgA (normal: 0.7 - 4 g/L)	1.9	2.93	1.72
IgG (normal: 7–16 g/L)	27.8	56.6	23.3
IgM (normal: 0.4-2.3 g/L)	0.63	0.77	2.5
IgE (normal: 5–500 KU/L)	<2	<2	2.5
T-cell subsets	Normal T-cell subsets	T-cell lymphopenia; intact expression of MHC class I and II antigens	Absolute CD3+, CD4+, and CD8+ T-cell lymphopenia

[a]Initial (pediatric) studies were done in early childhood (3–5 years of age). Follow-up (adult) studies were done between 14 and 17 years of age.

Prediction of the 3D protein structure based on different threading templates such as the crystal structure of the transmembrane PAP2 type phosphatidylglycerol phosphate phosphatase from *Bacillus subtilis* (see transparent methods) revealed that Ser160 is located close to the end of the fourth transmembrane helix, near the ER luminal side. Ser160 is in close contact with neighboring helix 1 and is predicted to be involved in buried hydrogen bond intramolecular interactions with Arg161, Thr30, and the backbone carbonyl of Trp26 (Figure 2, G6PC3-A, B). Thus, Ser160 plays an important role in tethering helices 1 and 4 together. The mutation Ser160Leu would disrupt this interaction network and cause steric clashes (Figure 2, G6PC3-C). Hence, Ser160Leu would destabilize the interface between helices 1 and 4 and thus affect the tertiary structure of this region, which could reduce the catalytic efficiency of the enzyme or the efficiency of coupling to the transporter protein glucose 6-phosphate translocase (SLC37A4).

The amino acid residue Pro133 maps to the extracellular domain of MPL which contains two cytokine receptor homology regions (CHRs), each composed of two fibronectin type III (FNIII) domains. Pro133 is part of the linker connecting the first and second FNIII of the N-terminal CHR (CHR1). The structure of this CHR1 region can be inferred based on a ~28% sequence identity with the erythropoietin receptor (Figure 3, MPL-A). Computational modeling suggests that substituting Pro133 by a larger and more hydrophobic leucine affects the stability and dynamics of this region (Figure 3; MPL-B). Given that the relative orientation of the FNIII domains critically affects signaling through cytokine receptors (Syed et al., 1998), the Pro133Leu mutation is expected to perturb MPL activation and signaling.

Protein expression changes between G6PC3-deficient patients and control subjects

Peripheral blood plasma samples from the 3 subjects diagnosed with G6PC3 deficiency causing complex neutropenia in addition to IBD in patient IV.6, as well as 6 healthy subjects, were analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS). A total of 460 unique protein species were identified (Table S1) of which 148 were significantly differentially expressed (≥ 2 to ∞: fold change and $p < 0.05$) between G6PC3 deficient and normal control subjects (Table S2). Classification of those 148 proteins based on their biological processes and molecular functions identified 4 major groups including the immune system, homeostasis, small molecule transport, and vesicle-mediated transport, respectively (Figures 4, 5, and 6). However, a 21-protein subpanel was represented in the pathway analysis of multiple signaling networks

Table 3. Pathogenic or likely pathogenic variants in the 3 patients' exomes including cytokinome profiles

Gene[a] (mutation/variant)	Disease association[a]	Patient IV.1	Patient IV.2	Patient IV.6
G6PC3:NM_138387:exon4:c.479C > T:p.S160L	SCN4 (AR), Dursun syndrome	hmz	hmz	hmz
MPL:NM_005373:exon4:c.398C > T:p.P133L	CAMT (AR), thrombocythemia (AD), MFMM (som)	wt	wt	het
HBG1:NM_000559:c.6_[a]3delinsCTCT	HbF (QTL1)	hmz	hmz	hmz
KLF1:NM_006563:exon2:12996130-12996956:-14:840	HPFH (AD), CDA type 4	wt	wt	hmz
MS4A2 (FCER1B): NM_001256916:exon6:c.575A > G:p.E192G	susceptibility to atopy	het	wt	hmz
CCL15:NM_032965:exon1:c.5A > T:p.K2M	None	hmz	hmz	hmz
CXCR2:NM_001557:exon3:c.1075A > T:p.T359S	None	wt	het	het
IFNAR1:NM_000629:exon8:c.1143 + 12G > A	None	het	wt	wt
IFNAR2:NM_000874:exon7:c.611C > G:p.T204R	IMD45 (AR)	het	wt	wt
IFNW1:NM_002177:exon1:c.247A > G:p.M83V	None	wt	wt	het
IL10RA:NM_001558:exon1:c.67 + 11G > C	IBD28 (AR)	het	wt	het
IL16:NM_172217:exon4:c.643T > C:p.S215P	None	het	wt	het
IL18R1:NM_003855:exon6:c.626-7C > T	None	het	wt	het
IL1R1:NM_001288706:exon12:c.1211-4T > C	None	het	wt	het
IL1RL1:NM_016232:exon11:c.1501_1502delinsAG	None	het	wt	het
IL21R:NM_021798:exon9:c.1381dupG:p.A460fs	IMD56 (AR), high IgE (AD)	het	wt	wt

AD, autosomal dominant; AR, autosomal recessive; CAMT, congenital amegakaryocytic thrombocytopenia; CDA, congenital dyserythropoietic anemia; CCL15, chemokine, cc motif, ligand 15; CXCR2, chemokine, CXC motif, receptor 2; G6PC3, glucose-6-phosphatase catalytic 3; HBG1, hemoglobin gamma A; het, heterozygous mutant; hmz, homozygous mutant; HPFH, hereditary persistent fetal hemoglobin; IBD, inflammatory bowel disease; IMD, immunodeficiency; IFNAR1, interferon-alpha-beta- and omega receptor 1; IFNAR2, interferon-alpha-beta- and omega receptor 2; IFNW1, interferon, omega-1; IL10RA, interleukin 10 receptor alpha; IL16, interleukin 16; IL18R1, interleukin 18 receptor-1; IL1R1, interleukin 1 receptor-1; IL1RL1, interleukin 1 receptor-like 1; IL21R, interleukin-21 receptor; KLF1, Kruppel-like factor 1; MFMM, myelofibrosis with myeloid metaplasia; MPL, myeloproliferative virus oncogene; QTL, quantitative trait locus; SCN, severe congenital neutropenia; som, somatic; wt, wild type.

Cytokinome refers to patient-specific rare DNA variants.

[a]Disease-gene association data are based on OMIM database "www.omim.org".

using the ingenuity pathway analysis (IPA, Qiagen) (Figures 7 and 8, networks 1 and 2, respectively) and STRING network database analysis tools (Figure 9). The functional annotations and the expression profiles and other characteristics of these proteins are detailed in Figure 5.

Interestingly, several globin chain proteins (HBA1, HBB, HBD, and HBZ) as well as immune-related proteins (C6, IL6R, LST1, JCHAIN, MBL2, and TRGC2) were downregulated. However, the neutrophil defensins, DEFA1, DEFA4, and neutrophil gelatinase-associated lipocalin (LCN2) were not differentially expressed. In addition, several downregulated proteins associated with known monogenic disorders were found including C6 (complement 6 deficiency), angiotensinogen (hypertension, renal tubular dysgenesis), GATA4 (congenital heart disease), thalassemia (HBA1, HBB, HBD, HBZ), MBL2 (chronic infections), PLOD2 (Bruck syndrome 2), SERPINF1 (osteogenesis imperfecta type VI, OI6), and WASHC4 (autosomal recessive mental retardation, MR41).

Perturbation of glutathione homeostasis is also evidenced by the dysregulation of the key enzymes glutathione peroxidase (GPX3) and glutaredoxin 2 (GLXR2) which were upregulated and downregulated, respectively. GPX3, glutathione peroxidase, is a plasma protein that protects cells, enzymes, and lipids against oxidative stress (peroxidation). GLRX2, glutaredoxin 2, is a glutathione-dependent hydrogen

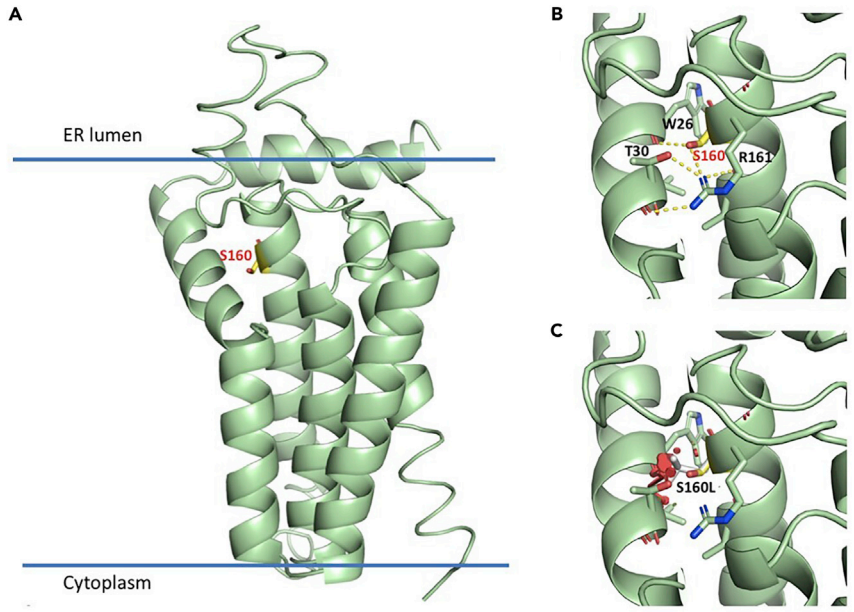


Figure 2. Computational assessment of the impact of the G6PC3 mutation

(A) Homology model produced with I-TASSER of the region comprising the residues 1–195. Ser160 is highlighted as yellow stick model, and the orientation of the protein with respect of the ER membrane is shown.

(B). Predicted intramolecular interactions between helices 1 and 4, mediated by Ser160. Hydrogen bonds are indicated with dashed lines.

(C) The substitution of Ser160 for leucine (gray sticks) would destabilize the interface between the two helices, altering the tertiary structure of the region.

donor for ribonucleotide reductase with a predominantly nuclear and a minor mitochondrial isoform (Lundberg et al., 2001); upon oxidative stress, the iron-sulfur cluster serves as a redox sensor for the activation of GLRX2 (Lillig et al., 2005).

Kynurenine formamidase (AFMID) was significantly overexpressed, and kynurenine was elevated (unpublished data) suggesting a perturbation in tryptophan catabolism.

DISCUSSION

We describe here a highly consanguineous pedigree with a novel exon 4 mutation in *G6PC3* associated with a variably severe phenotype not concordant with the severity of neutropenia. G6PC3-related neutropenia is thought to be a rare entity, but given the varying severity of the disorder, from cyclical neutropenia to a severe life-threatening immunocompromised state and the spectrum of non-hematological complications, it is almost certainly underdiagnosed. It is also likely that in patients where the cardiac, neurodevelopmental, and gastrointestinal manifestations are more severe than the degree of neutropenia, which may be mild or near normal such as in our index case-patient IV.1, SCN may not be suspected at all. The (novel) *G6PC3* exon 4 mutation we identified in this family has not, to our knowledge, been previously reported and may represent a candidate founder mutation, leading to a phenotype consistent with previous reports of *G6PC3* mutations most of which impair its enzyme activity. Lin et al. (2015) showed that 14 out of 16 missense *G6PC3* mutations abolished its enzymatic activity. Indeed, our computational structural analysis suggests that this novel mutation destabilizes the protein's 3D structure and stability at the ER luminal side. Recently, an unrelated Saudi Arabian pedigree was reported with a c.974T > G, p.Leu325Arg mutation and thrombocytopenia (Alangari et al., 2013). Conversely, we observed in our family thrombocytosis in association with neutropenia. While both thrombocytopenia and thrombocytosis may be seen clinically in sick individuals and are usually attributed to a toxic effect on myelopoiesis, alternatively, this effect may be mediated via variant(s) in the *MPL-THPO* axis (Dasouki et al., 2015). We, therefore, examined the gDNA from these 3 affected individuals for possibly damaging mutations and identified a rare heterozygous, potentially damaging, activating *MPL* mutation (c.398C > T, p.P133L). Our structural analysis localized it in a region involved in binding to its THPO ligand and hence predicted that it affects ligand recognition

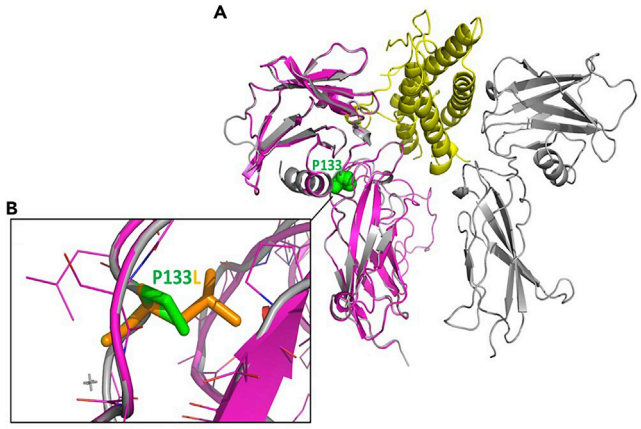


Figure 3. Computational assessment of the impact of the MPL mutation

(A) Shown is the homology model of MPL CHR1 (magenta), produced based on the 28% identical structure of the CHR of EPOR (gray) bound to erythropoietin (yellow; PDB id 4cn4). Pro133 is highlighted as a green sphere model.

(B) Zoom into the region of Pro133 (shown as a green stick model). The substitution of Pro133 by a leucine (orange) would increase the hydrophobic interaction and hence structural and dynamic coupling of both FNIII domains of CHR1. CHR, cytokine receptor homology region; EPOR, erythropoietin receptor; FNIII, fibronectin type III; MPL, myeloproliferative leukemia virus oncogene.

and signal initiation. Previously, it was shown that the deletion of the *MPL* CHR1 region has activating effects (Sabath et al., 1999), whereas point mutations in this region can have opposing effects, resulting in either gain or loss of function (Stockklausner et al., 2015). In the family reported by Alangari et al., no mutation analysis of *MPL* or *THPO* was reported.

In this progeny, we noticed striking variability in phenotype and disease severity arising from the same mutation. We suspect that other genomic and epigenetic variations likely account for these phenotypic differences. The list of genotypes for multiple loci identified in their exomes shown in Table 3 might explain additional hematological phenotypes such as elevated HbF and thrombocytosis, immune phenotypes such as low serum IgE levels and hypergammaglobulinemia, and the IBD phenotype.

Serum IgE deficiency is defined as an isolated significant decrease (<2.5 KU/L) in serum IgE level (combined with normal levels of other immunoglobulins). While selective serum IgE deficiency is still not recognized as a primary immunodeficiency disorder, it has been detected in seemingly healthy individuals, as well as in association with certain infections, sinopulmonary disease, non-allergic airway disease, autoimmunity, and various immunodeficiencies. In *Lta* (lymphotoxin) knockout mice, diminished serum IgE levels were associated with non-allergic Th1-mediated inflammatory airway disease (Kang et al., 2003). Also, IgE (a mucosal antibody) had been shown to possess antitumor activities both *in vitro* against pancreatic cancer (Fu et al., 2008; Reali et al., 2001) and in mouse models (Nigro et al., 2009; Daniels-Wells et al., 2013). These data suggest that serum IgE deficiency detected in our G6PC3-deficient patients may not be a completely benign finding.

Mutations in the mannose-binding lectin 2 (*MBL2*) have been implicated in predisposition to recurrent infections, albeit not consistently. Up to 10% of healthy individuals may have mannose-binding lectin deficiency (MBLD) and be asymptomatic. Moreover, variants have been associated with MBLD and a more severe phenotype in Crohn disease (Choteau et al., 2016).

In addition to the G6PC3 deficiency-related neutropenia and MBL2 deficiency, additional (environmental, genetic, and epigenetic) factors likely contributed to the recurrent chest infections and bronchiectasis observed in individuals IV.1 and 2. Although deaths due to respiratory failure have been reported, there are no reports to our knowledge of lung transplantation in individuals with G6PC3 deficiency (Desplantes et al., 2014).

The human neutrophil proteome contains approximately 4100 proteins, including 7 which account for 50% including the antimicrobial proteins DEF3A, S100A8, LYZ, and CTSG being among the most abundant (Grabowski et al., 2019) which we did not detect in our proteomics study. However, we detected the

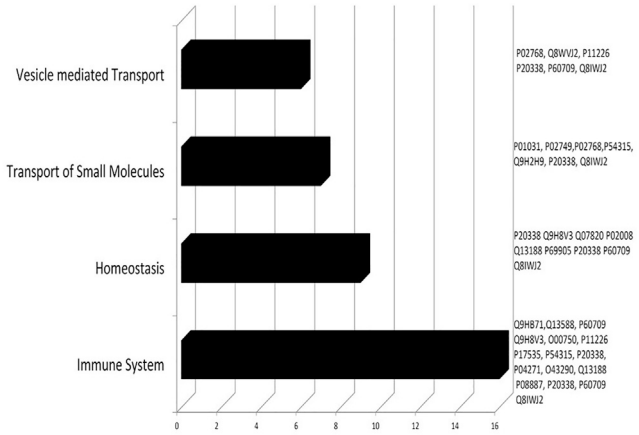


Figure 4. Biological process and molecular functional classification of the 148 differentially expressed proteins between CSN and control samples

Immune system- (16) and homeostasis (9)-related proteins were most the commonly represented.

neutrophil defensins, DEFA1, DEFA4, and LCN2 which were not differentially expressed. In addition to our study and the study of Grabowski et al. (2019), we could not find any other proteomics studies that examine monogenic neutropenia. The neutrophil-based proteomic biomarkers they detected in SCN, chronic granulomatous disease, and leukocyte adhesion deficiency 1 and 2 caused by mutations in *ELANE*, *CYBA* and *CYBB*, and *ITGB2* and *SLC35C1*, respectively, were disease specific and did not overlap with the profile caused by G6PC3 deficiency.

In our G6PC3-deficient individuals, in addition to congenital neutropenia which results in the loss of the known antimicrobial functions of neutrophils, we observed dysregulation of additional factors that may have exacerbated this primary defect, including the leukocyte-specific transcript 1 (LST1), the (myeloid specific) monocyte-specific differentiation factor (CD14), and MBL2 that have diverse immune functions.

LST1 is an extensively alternatively spliced major histocompatibility class (MHC) III gene with 13 (800-nucleotide) interferon gamma-inducible transcripts and is expressed in lymphoid tissues, T cells, macrophages, and histiocyte cell lines (Holzinger et al., 1995; Weidle et al., 2018). LST1 promotes the assembly of intercellular tunneling nanotube formation (Schiller et al., 2012) which used is for the transfer of organelles (Rustom et al., 2004), soluble markers (Watkins and Salter 2005), and electrical signals (Wang et al., 2010). Recently, the myeloid leukocyte-specific transmembrane LST1/A isoform was shown to function as an adapter protein that recruits the protein tyrosine phosphatases SHP-1 and SHP-2 to the plasma membrane, thus effecting negative regulation of signal transduction which is necessary for controlling the cell response (Draber et al., 2013). Also, the proinflammatory expression of LST1 in human IBD and non-IBD colitis was shown not to be restricted to immune cell populations (Heidemann et al., 2014). Taken together, this evidence strongly suggests a significant role for LST1 in the pathogenesis of immune dysfunction in G6PC3 deficiency.

Several patients with G6PC3 deficiency and IBD or IBD-like disorder have been described, although the mechanism(s) of this association is not known (Begin et al., 2013; Chandrakasan et al., 2017). In Table 3, we hypothesize that the novel genomics-based immune/cytokine signature profile which we identified in patient IV.6 but not in her two cousins without IBD is potentially IBD related. However, additional confirmatory studies are needed. Infliximab, a chimeric anti-TNF alpha monoclonal antibody, is used to treat patients with various autoimmune disorders including IBD (Crohn disease). Similarly, as reported by Begin et al., infliximab resulted in clinical improvement, albeit transient. Platelet factor 4 (PF4, CXCL4) is an index of platelet activation and thromboembolic risk (Simi et al., 1987). Elevated plasma and colonic mucosa PF4 levels correlated with disease activity (Simi et al., 1987; Ye et al., 2017; Sobhani et al., 1992) and response to treatment with infliximab in patients with IBD (Crohn disease) (Meuwis et al., 2008). Consequently, the up-regulation of plasma PF4 levels we observed in this study may have a similar implication.

The GATA factors (1–6) are zinc-finger 'WGATAR' DNA motif binding proteins that regulate diverse pathways associated with embryonic morphogenesis and cellular differentiation (Patient and McGhee, 2002).

Immune System

Accession	Anova (p)	Max fold cl	Highest me	Lowest me	Description	GN
Q9HB71	5,51E-05	Infinity	NORMAL	CSCN	Calcyclin-binding protein	CACYBP
Q13588	0,003768	8,398633	CSCN	NORMAL	GRB2-related adapter protein	GRAP
P60709	0,000701	2,42359	CSCN	NORMAL	Actin_ cytoplasmic 1	ACTB
Q9H8V3	0,001539	Infinity	NORMAL	CSCN	Protein ECT2	ECT2
O00750	0,001355	2,056049	NORMAL	CSCN	Phosphatidylinositol 4-phosphate 3-kinas	PIK3C2B
P11226	0,000314	2,503113	NORMAL	CSCN	Mannose-binding protein C	MBL2
P08887	1,06E-05	4,18871	NORMAL	CSCN	Interleukin-6 receptor subunit alpha	IL6R
Q13188	0,000171	74,58913	NORMAL	CSCN	Serine/threonine-protein kinase 3	STK3
Q8IWJ2	0,000174	2,247177	CSCN	NORMAL	GRIP and coiled-coil domain-containing p	GCC2
P60709	0,000701	2,42359	CSCN	NORMAL	Actin_ cytoplasmic 1	ACTB
O43290	0,001147	6,114737	NORMAL	CSCN	U4/U6.U5 tri-snRNP-associated protein 1	SART1
P17535	0,002783	5,11978	NORMAL	CSCN	Transcription factor jun-D	JUND
P20338	0,005719	2,14666	CSCN	NORMAL	Ras-related protein Rab-4A	RAB4A
P54315	0,035214	2,009928	CSCN	NORMAL	Inactive pancreatic lipase-related protein 1	PNLIPRP1
P04271	0,046813	4,434763	NORMAL	CSCN	Protein S100-B	S100B

Vesicle Mediated Transport

Accession	Anova (p)	Max fold cl	Highest me	Lowest me	Description	GN
P02768	8,99E-05	3,722523	NORMAL	CSCN	Serum albumin	ALB
Q8IWJ2	0,000174	2,247177	CSCN	NORMAL	GRIP and coiled-coil domain-containing protein 2	GCC2
P11226	0,000314	2,503113	NORMAL	CSCN	Mannose-binding protein C	MBL2
P60709	0,000701	2,42359	CSCN	NORMAL	Actin_ cytoplasmic 1	ACTB
P20338	0,005719	2,14666	CSCN	NORMAL	Ras-related protein Rab-4A	RAB4A
Q8WVJ2	0,032164	2,342885	CSCN	NORMAL	NudC domain-containing protein 2 OS=Homo sapiens GN=NUDCD2 PE=1 SV=1	NUDCD2

HOMEOSTASIS

Accession	Anova (p)	Max fold cl	Highest me	Lowest me	Description	GN
P20338	0,005719	2,14666	CSCN	NORMAL	Ras-related protein Rab-4A	RAB4A
Q8WVJ2	0,032164	2,342885	CSCN	NORMAL	NudC domain-containing protein 2	NUDCD2
P60709	0,000701	2,42359	CSCN	NORMAL	Actin_ cytoplasmic 1	ACTB
Q07820	9,81E-06	2,540953	CSCN	NORMAL	Induced myeloid leukemia cell differentiation protein Mcl-1	MCL1
P69905	9,22E-05	3,792674	NORMAL	CSCN	Hemoglobin subunit alpha	HBA1
Q13188	0,000171	74,58913	NORMAL	CSCN	Serine/threonine-protein kinase 3	STK3
P02008	0,000264	4,385323	NORMAL	CSCN	Hemoglobin subunit zeta	HBZ
Q9H8V3	0,001539	Infinity	NORMAL	CSCN	Protein ECT2	ECT2

TRANSPORT OF SMALL MOLECULES

Accession	Anova (p)	Max fold cl	Highest me	Lowest me	Description	
P20338	0,005719	2,14666	CSCN	NORMAL	Ras-related protein Rab-4A	RAB4A
Q8WVJ2	0,032164	2,342885	CSCN	NORMAL	NudC domain-containing protein 2	NUDCD2
P02768	8,99E-05	3,722523	NORMAL	CSCN	Serum albumin	ALB
Q9H2H9	0,000976	53,24604	CSCN	NORMAL	Sodium-coupled neutral amino acid transporter 1	SLC38A1
P01031	0,001422	2,635753	CSCN	NORMAL	Complement C5	C5
P02749	0,0091	2,026687	NORMAL	CSCN	Beta-2-glycoprotein 1	APOH
P54315	0,035214	2,009928	CSCN	NORMAL	Inactive pancreatic lipase-related protein 1	PNLIPRP1

Figure 5. Biological process and molecular functional classification of the 148 differentially expressed proteins between CSN and control samples

GATA4 is essential for cardiac morphogenesis as well as the development of the liver, pancreas, and swim bladder (Holtzinger and Evans, 2005). Heterozygous loss-of-function mutations in *GATA4* are associated with a variety of cardiac malformations including (atrial, ventricular, and atrioventricular) septal defects and tetralogy of Fallot (Garg et al., 2003). The finding of low expression of GATA4 in the plasma of our 3 patients with G6PC3 deficiency and congenital heart disease suggests a link which requires additional studies.

Conclusions

G6PC3 deficiency is a complex multisystem disorder with neutropenia as the main phenotype and IBD-like disease having been described in some patients. Proteomics profiling in monogenic neutropenia is disease

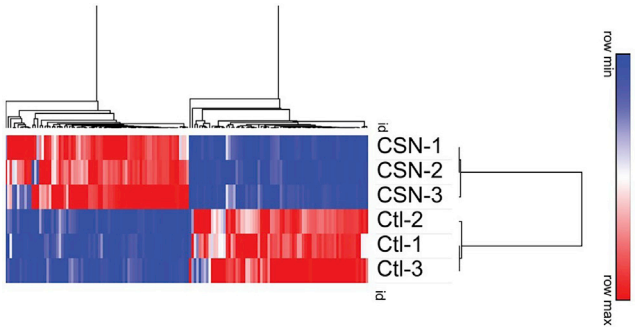


Figure 6. Differential protein expression in patients with G6PC3 deficiency

Heatmap of protein expression in 3 patients with severe congenital neutropenia (CSN) caused by G6PC3 deficiency and 6 healthy controls. Patient plasma samples and controls were pooled separately and run in triplicates.

specific. Using extensive genomics (including genomic cytokine profiling) and proteomics analyses, we identified a unique IBD-related profile including perturbations in several cytokines, LST1. We also observed the downregulation of GATA4 which may potentially explain the congenital heart disease seen in some patients with G6PC3 deficiency. Further studies are needed to elucidate the underlying mechanisms of soluble and membrane-expressed LST1 and GATA-mediated phenotypic defects as well as the role of kynurenine in the regulation of mucosal intestinal immunity and inflammation.

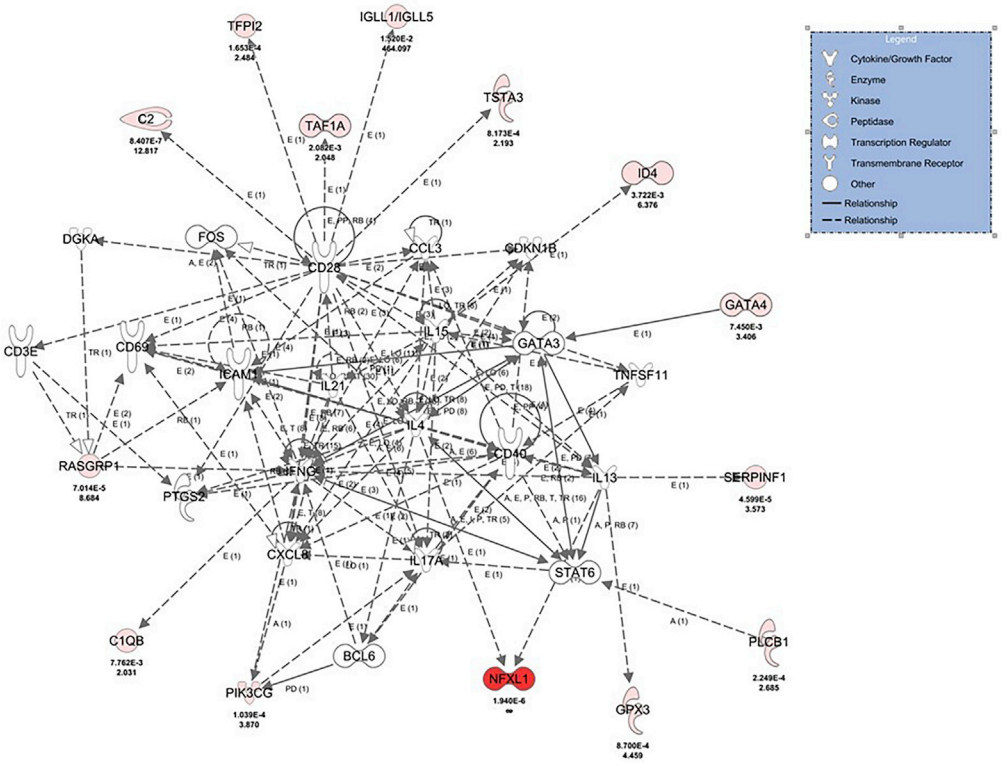


Figure 7. IPA network analysis of differentially expressed proteins in G6PC3 deficiency

Ingenuity Pathway Analysis (IPA)-based network 1 of the 148 differentially expressed protein panel. The network is displayed graphically as nodes (gene/gene products) and edges (the biological relationship between nodes). The node color intensity indicates the expression of genes: red is upregulated and green is downregulated in plasma. The fold change values of differentially expressed proteins are indicated under each corresponding node. Node shapes indicate the functional class of the gene product as shown in the key.

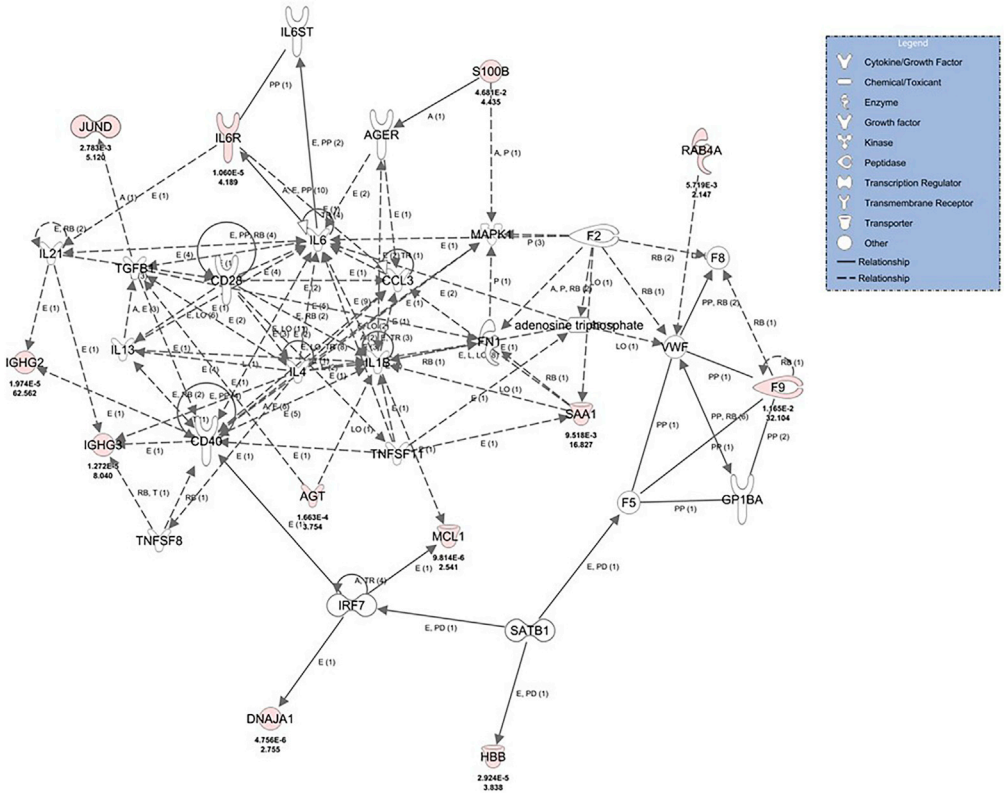


Figure 8. IPA networks analysis of differentially expressed proteins in G6PC3 deficiency

Ingenuity Pathway Analysis (IPA)-based network 2 of the 148 differentially expressed protein panel. The network is displayed graphically as nodes (gene/gene products) and edges (the biological relationship between nodes). The node color intensity indicates the expression of genes: red is upregulated and green is downregulated in plasma. The fold change values of differentially expressed proteins are indicated under each corresponding node. Node shapes indicate the functional class of the gene product as shown in the key.

Limitations of the study

In this study, we used genome-wide (exome) and MS-based plasma proteomics analyses to identify genomic and proteomic alterations in a highly consanguineous family with G6PC3 deficiency-related complex phenotypes. Apart from the novel G6PC3 and MPL variants which are likely to be directly related to the patients phenotypes and given the small sample size and incomplete knowledge about the other genomic variants shown in Table 3 and potential complex interactions between them, their exact role in this complex disease cannot be confirmed. Although MS-based comparative proteomics has increasingly become a powerful tool for biomedical research of complex monogenic and polygenic disorders, its analytical depth is still limited and several biochemical pathways playing a role in neutrophil dysfunction may have not been detected. Also, changes in protein abundance alone cannot characterize the entire pathology or underlying mechanism(s) of a complex disease like syndromic G6PC3 deficiency. The use of fully characterized healthy controls, which we could not use in our proteomics study, is highly desirable as it is more likely to reveal real differences in expression not related to genetic background. Also, the pooling of patients' plasma samples prevented us from potentially correlating specific differentially expressed proteins with varying clinical phenotypes.

Resource availability

Lead contact

Majed Dasouki, MD, Department of Genetics, Department of Pathology & Laboratory Medicine, King Faisal Specialist Hospital & Research Center, MBC-03-30. PO Box 3354. Riyadh, 11,211. Saudi Arabia, Current email: majed.dasouki.md@adventhealth.com.

Material availability

This study did not generate new unique reagents.

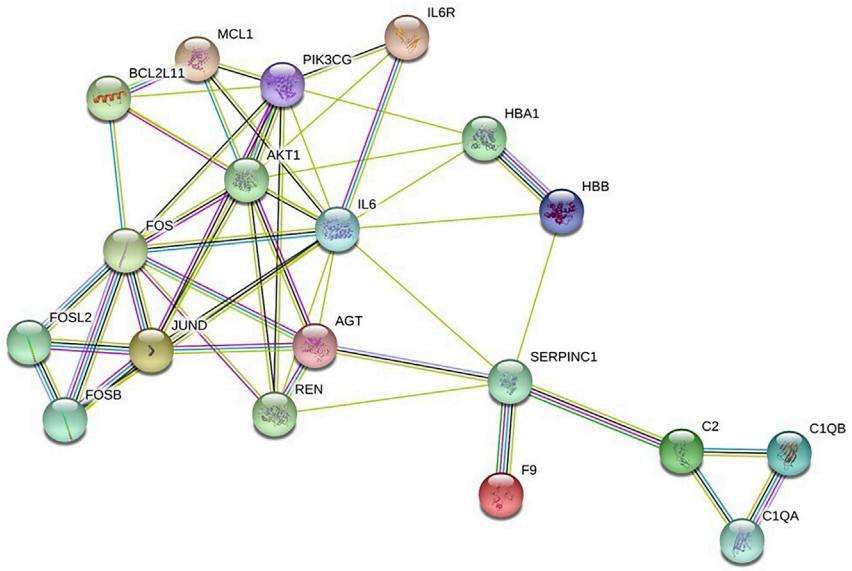


Figure 9. The inter-relationships and functional characteristics of some of the 21 identified proteins as mapped in STRINGS pathway analysis database

Some of these molecules function as enzymes, coagulation factors, transporters, transcription regulators, and homeostasis and differentiation regulators. Others act as kinases, peptidase or growth factors, and cytokines. (The connecting network analyses were done, and the above figures were generated in Ingenuity Pathway Analysis program [IPA, version 8.7] and string, respectively). AGT, angiotensinogen; AKT1, AKT serine/threonine kinase 1; BCL2L11, BCL2-like 11; C1QA, complement component 1, q subcomponent, A chain; C1QB, complement component 1, q subcomponent, B chain; C2, complement component 2; F9, coagulation factor IX; FOS, FOS protooncogene; FOSB, FOSB protooncogene; FOSL1, FOS-like antigen 1; FOSL2, FOS-like antigen 2; HBA1, hemoglobin-alpha locus 1; HBB, hemoglobin-beta locus; IL6, interleukin 6; IL6R, interleukin 6 receptor; JUND, oncogene JUN-D; MCL1, myeloid leukemia sequence 1; PIK3CG, phosphatidylinositol 3-kinase catalytic gamma; REN, renin; SERPINC1, serpin peptidase inhibitor clade C (antithrombin) member 1.

Data and code availability

The published article includes all data sets/code generated or analyzed during this study.

METHODS

All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102214.

ACKNOWLEDGMENTS

The authors wish to thank all the clinicians who provided routine clinical care for these patients and thank the patients and their families who participated in this project. We also acknowledge the Saudi Human Genome Project for infrastructure and informatics support relating to the NGS work. The work by M.D., F.A., H.A., M.A., S.O.A. was supported by the National Science, Technology, and Innovation Plan program of King Abdulaziz City for Science & Technology (KACST), Saudi Arabia, award number: 13-BIO1978-20. The work by S.T.A. and F.J.G.V. was supported by the King Abdullah University of Science and Technology (KAUST), Saudi Arabia through the Award No. FCC/1/1976-25 from the Office of Sponsored Research (OSR).

AUTHOR CONTRIBUTIONS

M.J.D. and S.O.A. contributed in study design, supervision, and manuscript writing; A.A.A., Z.S., and M.J.D. contributed in proteomic studies; T.E., M.A., D.M., A.J., F.A., and M.J.D. contributed in genomic molecular

and bioinformatics studies; S.T.A. and F.J.G.V. contributed in protein structural analysis; M.S.A., H.A.S., and S.O.A. contributed in clinical studies. All authors reviewed and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 2, 2020
Revised: August 29, 2020
Accepted: February 17, 2021
Published: March 19, 2021

REFERENCES

Alangari, A.A., Alsultan, A., Osman, M.E., Anazi, S., and Alkuraya, F.S. (2013). A novel homozygous mutation in G6PC3 presenting as cyclic neutropenia and severe congenital neutropenia in the same family. J. Clin. Immunol. *33*, 1403–1406.

Banka, S. (2016). G6PC3 deficiency. In GeneReviews(R), R.A. Pagon, M.P. Adam, H.H. Ardinger, S.E. Wallace, A. Amemiya, L.J.H. Bean, T.D. Bird, C.T. Fong, H.C. Mefford, and R.J.H. Smith, et al., eds., https://www.ncbi.nlm.nih.gov/books/NBK285321/.

Banka, S., and Newman, W.G. (2013). A clinical and molecular review of ubiquitous glucose-6-phosphatase deficiency caused by G6PC3 mutations. Orphanet J. Rare Dis. *8*, 84.

Banka, S., Newman, W.G., Özgül, R.K., and Dursun, A. (2010). Mutations in the G6PC3 gene cause Dursun syndrome. Am. J. Med. Genet. A *152A*, 2609–2611.

Bégin, P., Patey, N., Mueller, P., Rasquin, A., Sirard, A., Klein, C., Haddad, É., Drouin, É., and Le Deist, F. (2012). Inflammatory bowel disease and T cell lymphopenia in G6PC3 deficiency. J. Clin. Immunol. *33*, 520–525.

Bennike, T., Birkelund, S., Stensballe, A., and Andersen, V. (2014). Biomarkers in inflammatory bowel diseases: Current status and proteomics identification strategies. World J. Gastroenterol. *20*, 3231.

Boztug, K., Appaswamy, G., Ashikov, A., Schäffer, A.A., Salzer, U., Diestelhorst, J., Germeshausen, M., Brandes, G., Lee-Gossler, J., Noyan, F., et al. (2009). A syndrome with congenital neutropenia and mutations in G6PC3. New Engl. J. Med. *360*, 32–43.

Chandrakasan, S., Venkateswaran, S., and Kugathasan, S. (2017). Nonclassic inflammatory bowel disease in young infants. Pediatr. Clin. North America *64*, 139–160.

Choteau, L., Vasseur, F., Lepretre, F., Figeac, M., Gower-Rousseau, C., Dubuquoy, L., Poulain, D., Colombel, J.-F., Sendid, B., and Jawhara, S. (2016). Polymorphisms in the mannose-binding lectin gene are associated with defective mannose-binding lectin functional activity in Crohn's disease patients. Sci. Rep. *6*, 29636.

Cullinane, A.R., Vilboux, T., O'Brien, K., Curry, J.A., Maynard, D.M., Carlson-Donohoe, H., Ciccone, C., NISC Comparative Sequencing Program, Markello, T.C., Gunay-Aygun, M., Huizing, M., and Gahl, W.A. (2011). Homozygosity mapping and whole-exome sequencing to detect SLC45A2 and G6PC3 mutations in a single patient with oculocutaneous albinism and neutropenia. J. Invest. Dermatol. *131*, 2017–2025.

Dale, D.C., Person, R.E., Bolyard, A.A., Aprikyan, A.G., Bos, C., Bonilla, M.A., Boxer, L.A., Kannourakis, G., Zeidler, C., Welte, K., et al. (2000). Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia. Blood *96*, 2317–2322.

Daniels-Wells, T.R., Helguera, G., Leuchter, R.K., Quintero, R., Kozman, M., Rodríguez, J.A., Ortiz-Sánchez, E., Martínez-Maza, O., Schultes, B.C., Nicodemus, C.F., and Penichet, M.L. (2013). A novel IgE antibody targeting the prostate-specific antigen as a potential prostate cancer therapy. BMC Cancer *13*, 195.

Dasouki, M., Saadi, I., and Ahmed, S.O. (2015). THPO–MPL pathway and bone marrow failure. Hematol. Oncol. Stem Cell Ther. *8*, 6–9.

Desplantes, C., Fremond, M.L., Beaupain, B., Harousseau, J.L., Buzyn, A., Pellier, I., Roques, G., Morville, P., Paillard, C., Bruneau, J., et al. (2014). Clinical spectrum and long-term follow-up of 14 cases with G6PC3 mutations from the French severe congenital neutropenia registry. Orphanet J. Rare Dis. *9*, 183.

Draber, P., Stepanek, O., Hrdinka, M., Drobek, A., Chmatal, L., Mala, L., Ormsby, T., Angelisova, P., Horejsi, V., and Brdicka, T. (2013). LST1/A is a myeloid leukocyte-specific transmembrane adaptor protein recruiting protein tyrosine phosphatases SHP-1 and SHP-2 to the plasma membrane. J. Biol. Chem. *288*, 28309.

Dursun, A., Ozgul, R.K., Soydas, A., Tugrul, T., Gurgey, A., Celiker, A., Barst, R.J., Knowles, J.A., Mahesh, M., and Morse, J.H. (2009). Familial pulmonary arterial hypertension, leucopenia, and atrial septal defect: a probable new familial syndrome with multisystem involvement. Clin. Dysmorphol. *18*, 19–23.

Fu, S.L., Pierre, J., Smith-Norowitz, T.A., Hagler, M., Bowne, W., Pincus, M.R., Mueller, C.M., Zenilman, M.E., and Bluth, M.H. (2008). Immunoglobulin E antibodies from pancreatic cancer patients mediate antibody-dependent cell-mediated cytotoxicity against pancreatic cancer cells. Clin. Exp. Immunol. *153*, 401–409.

Garg, V., Kathiriya, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K., et al. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature *424*, 443–447.

Ghosh, A., Shieh, J.-J., Pan, C.-J., and Chou, J.Y. (2004). Histidine 167 is the phosphate acceptor in glucose-6-phosphatase-β forming a phosphohistidine enzyme intermediate during catalysis. J. Biol. Chem. *279*, 12479–12483.

Glasser, C.L., Picoraro, J.A., Jain, P., Kinberg, S., Rustia, E., Gross Margolis, K., Anyane-Yeboa, K., Iglesias, A.D., and Green, N.S. (2016). Phenotypic heterogeneity of neutropenia and gastrointestinal illness associated with G6PC3 founder mutation. J. Pediatr. Hematol. Oncol. *38* (7), e243–e247.

Grabowski, P., Hesse, S., Hollizeck, S., Rohlfs, M., Behrends, U., Sherkat, R., Tamary, H., Ünal, E., Somech, R., Patıroğlu, T., et al. (2019). Proteome analysis of human neutrophil granulocytes from patients with monogenic disease using data-independent acquisition. Mol. Cell Proteomics *18*, 760–772.

Hayee, B., Antonopoulos, A., Murphy, E.J., Rahman, F.Z., Sewell, G., Smith, B.N., McCartney, S., Furman, M., Hall, G., Bloom, S.L., et al. (2011). G6PC3 mutations are associated with a major defect of glycosylation: a novel mechanism for neutrophil dysfunction. Glycobiology *21*, 914–924.

Heidemann, J., Kebschull, M., Tepasse, P.R., and Bettenworth, D. (2014). Regulated expression of leukocyte-specific transcript (LST) 1 in human intestinal inflammation. Inflamm. Res. *63*, 513–517.

Holtzinger, A., and Evans, T. (2005). Gata4 regulates the formation of multiple organs. Development *132*, 4005–4014.

Holzinger, I., de Baey, A., Messer, G., Kick, G., Zwierzina, H., and Weiss, E.H. (1995). Cloning and genomic characterization of LST1: a new gene in the human TNF region. Immunogenetics *42*, 315–322.

Kang, H.-S., Blink, S.E., Chin, R.K., Lee, Y., Kim, O., Weinstock, J., Waldschmidt, T., Conrad, D., Chen, B., Solway, J., et al. (2003). Lymphotoxin is required for maintaining physiological levels of serum IgE that minimizes Th1-mediated airway inflammation. J. Exp. Med. *198*, 1643–1652.

Klein, C., Grudzien, M., Appaswamy, G., Germeshausen, M., Sandrock, I., Schäffer, A.A., Rathinam, C., Boztug, K., Schwinzer, B., Rezaei,

N., et al. (2006). HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). Nat. Genet. *39*, 86–92.

Lillig, C.H., Berndt, C., Vergnolle, O., Lonn, M.E., Hudemann, C., Bill, E., and Holmgren, A. (2005). Characterization of human glutaredoxin 2 as iron-sulfur protein: a possible role as redox sensor. Proc. Natl. Acad. Sci. *102*, 8168–8173.

Lin, S.R., Pan, C.J., Mansfield, B.C., and Chou, J.Y. (2015). Functional analysis of mutations in a severe congenital neutropenia syndrome caused by glucose-6-phosphatase-β deficiency. Mol. Genet. Metab. *114*, 41–45.

Lundberg, M., Johansson, C., Chandra, J., Enoksson, M., Jacobsson, G., Ljung, J., Johansson, M., and Holmgren, A. (2001). Cloning and expression of a novel human glutaredoxin (Grx2) with mitochondrial and nuclear isoforms. J. Biol. Chem. *276*, 26269–26275.

Meuwis, M.-A., Fillet, M., Lutteri, L., Marée, R., Geurts, P., de Seny, D., Malaise, M., Chapelle, J.-P., Wehenkel, L., Belaiche, J., et al. (2008). Proteomics for prediction and characterization of response to infliximab in Crohn's disease: a pilot study. Clin. Biochem. *41*, 960–967.

Nigro, E.A., Brini, A.T., Soprana, E., Ambrosi, A., Dombrowicz, D., Siccardi, A.G., and Vangelista, L. (2009). Antitumor IgE adjuvanticity: key role of FcεRI. J. Immunol. *183*, 4530–4536.

Patient, R.K., and McGhee, J.D. (2002). The GATA family (vertebrates and invertebrates). Curr. Opin. Genet. Dev. *12*, 416–422.

Reali, E., Greiner, J.W., Corti, A., Gould, H.J., Bottazzoli, F., Paganelli, G., Schlom, J., and Siccardi, A.G. (2001). IgEs targeted on tumor cells: therapeutic activity and potential in the design of tumor vaccines. Cancer Res. *61*, 5517–5522.

Rustom, A., Saffrich, R., Markovic, I., Walther, P., and Gerdes, H.H. (2004). Nanotubular highways for intercellular organelle transport. Science *303*, 1007–1010.

Sabath, D.F., Kaushansky, K., and Broudy, V.C. (1999). Deletion of the extracellular membrane-distal cytokine receptor homology module of mpl results in constitutive cell growth and loss of thrombopoietin binding. Blood *94*, 365–367.

Schiller, C., Diakopoulos, K.N., Rohwedder, I., Kremmer, E., von Toerne, C., Ueffing, M., Weidle, U.H., Ohno, H., and Weiss, E.H. (2012). LST1 promotes the assembly of a molecular machinery responsible for tunneling nanotube formation. J. Cell Sci. *126*, 767–777.

Simi, M., Leardi, S., Tebano, M.T., Castelli, M., Costantini, F.M., and Speranza, V. (1987). Raised plasma concentrations of platelet factor 4 (PF4) in Crohn's disease. Gut *28*, 336–338.

Sobhani, I., Hochlaf, S., Denizot, Y., Vissuzaine, C., Rene, E., Benveniste, J., Lewin, M.M., and Mignon, M. (1992). Raised concentrations of platelet activating factor in colonic mucosa of Crohn's disease patients. Gut *33*, 1220–1225.

Stockklausner, C., Klotter, A.-C., Dickemann, N., Kuhlee, I.N., Duffert, C.M., Kerber, C., Gehring, N.H., and Kulozik, A.E. (2015). The thrombopoietin receptor P106L mutation functionally separates receptor signaling activity from thrombopoietin homeostasis. Blood *125*, 1159–1169.

Syed, R.S., Reid, S.W., Li, C., Cheetham, J.C., Aoki, K.H., Liu, B., Zhan, H., Osslund, T.D., Chirino, A.J., Zhang, J., et al. (1998). Efficiency of signaling through cytokine receptors depends critically on receptor orientation. Nature *395*, 511–516.

Veiga-da-Cunha, M., Chevalier, N., Stephenne, X., Defour, J.-P., Paczia, N., Ferster, A., Achouri, Y., Dewulf, J.P., Linster, C.L., Bommer, G.T., and Van Schaftingen, E. (2019). Failure to eliminate a phosphorylated glucose analog leads to neutropenia in patients with G6PT and G6PC3 deficiency. Proc. Natl. Acad. Sci. *116*, 1241–1250.

Wang, X., Veruki, M.L., Bukoreshtliev, N.V., Hartveit, E., and Gerdes, H.H. (2010). Animal cells connected by nanotubes can be electrically coupled through interposed gap-junction channels. Proc. Natl. Acad. Sci. U S A *107*, 17194–17199.

Watkins, S.C., and Salter, R.D. (2005). Functional connectivity between immune cells mediated by tunneling nanotubules. Immunity *23*, 309–318.

Weidle, U.H., Rohwedder, I., Birzele, F., Weiss, E.H., and Schiller, C. (2018). LST1: a multifunctional gene encoded in the MHC class III region. Immunobiology *223*, 699–708.

Ye, L., Zhang, Y.-P., Yu, N., Jia, Y.-X., Wan, S.-J., and Wang, F.-Y. (2017). Serum platelet factor 4 is a reliable activity parameter in adult patients with inflammatory bowel disease. Medicine *96*, e6323.

iScience, Volume *24*

Supplemental information

Comprehensive multi-omics analysis of G6PC3 deficiency-related congenital neutropenia with inflammatory bowel disease

Majed Dasouki, Ayodeele Alaiya, Tanziel ElAmin, Zakia Shinwari, Dorota Monies, Mohamed Abouelhoda, Amjad Jabaan, Feras Almourfi, Zuhair Rahbeeni, Fahad Alsohaibani, Fahad Almohareb, Hazzaa Al-Zahrani, Francisco J. Guzmán Vega, Stefan T. Arold, Mahmoud Aljurf, and Syed Osman Ahmed

Materials and Methods

The 3 affected young adult patients (IV.1, 2, 6) and their parents (Fig.1) were enrolled in an IRB approved protocol (RAC 2060-021) at the King Faisal Specialist Hospital & Research Center.

Patient clinical reports

Patient IV.1 whose gestation was complicated by polyhydramnios and born at full term, was first seen at 8 months of age with history of diarrhea and steatorrhea, failure to thrive (weight & height $<5^{th}$ centile) and neutropenia that was not responding to GCSF. Shwachman-Diamond syndrome was suspected, however, molecular genetic testing was negative (table 1). She was also commenced on pancreatic enzyme replacement with some improvement. She also had recurrent skin and chest infections and needed IV antibiotics at times. She was followed clinically until the age of 6 and was subsequently lost to follow up. She left school at 7th grade due to her recurrent medical problems, mainly breathlessness. Currently, she is 20 years old.

She is still under-weight with a low BMI of 11.6. She had a similar facial phenotype to her paternal cousin (patient IV.6) with a depressed nasal bridge and full lips. She had no significant bowel symptoms. She has been relatively well and had the mildest phenotype, with an adequate ANC being maintained on GCSF 2-3 times per week. There was mild hypoxemia with saturations of 93-97%. She had absolute neutrophil count of <0.5 x10^9/L and positive (IgG and IgM) anti-neutrophil antibodies. Bone marrow evaluation revealed hypercellularity (90%) with hypoplastic myelopoiesis and left shift with abnormal segmentation of myeloid precursors. Cytogenetic studies including FISH for myelodysplastic syndrome (MDS) panel (5q-, -5, 7q-, -7, trisomy 8, MLL, 20q) as well as Fanconi anemia (DEB) chromosome breakage analyses were negative. Chest CT scan examination showed bronchiectasis and mild mediastinal lymphadenopathy with mucoid impaction, cystic changes, bronchial wall thickening and air trapping.

Patient IV.2 is the younger sister to patient IV.1, and had severe congenital neutropenia, with neutrophil count of <0.1 x10^9/L during childhood. She had a similar history of failure to thrive, chronic

malodourous diarrhea, steatorrhea which improved with pancreatic enzyme replacement therapy, and recurrent skin and chest infections requiring hospitalization. She had developmental delay and was also noted to have right arm and leg hemi-hypertrophy. She also was lost to follow up for several years. On return to clinic, there was an intervening history of severe recurrent chest infections. Her nutritional condition was poor and she had a low BMI. She has bilateral chest crepitation and a CT scan that confirmed bronchiectasis, which had been complicated by recurrent severe infections, right heart failure and hypoxia and was therefore considered as a candidate for lung transplantation. She is 19 years old.

Patient IV.6, the cousin of patients IV.1 and IV.2 presented since early childhood with failure to thrive, steatorrhea and diarrhea for which she received pancreatic enzyme replacement therapy. She was noted to have mild dysmorphic features including low set ears, depressed nasal bridge, full lips and long fingers. Although she had recurrent urinary tract infections, imaging revealed no urogenital malformations. She was also found to have moderate to large secundum atrial septal defect (ASD) for which she underwent repair at 8 years of age. She also gave a history of recurrent skin infections requiring oral and systemic antibiotics. She had only a mild intermittent neutropenia, with a median ANC of 1.39 (0.8 – 6.2 x 10^9) which did not require GCSF support. Having been lost to follow up for 11 years, mainly due to social reason, she then presented at 16 years of age weighing only 13 kg with a BMI of 12.7 and history of frequent diarrhea, fever and an oligo-arthritis. Colonoscopy revealed extensive ulceration and strictures. She also had esophageal strictures requiring dilatation. She was managed as having “atypical” Crohn’s disease, and was treated initially with steroids, azathioprine and 5-aminosalicylic acid (5-ASA). Numerous upper and lower GI biopsies revealed gastritis, colitis inflammatory exudate, ulceration; but no granulomas were noted. Abdominal MRI imaging revealed stricture disease of the bowel and hepatosplenomegaly, and transfusion related iron overload, but no urogenital abnormalities. She required bowel resection due to iatrogenic bowel perforation during endoscopy. The bowel resection specimen revealed transmural inflammation, with acute inflammation and bacterial growth at the site of perforation. More recently, she received infliximab, with transient improvement in symptoms. As an adult, she

presented with mild intermittent neutropenia requiring occasional therapy with GCSF. Neutrophils were morphologically normal. Other investigations revealed T-cell lymphopenia, elevated IgG, low IgE and an elevated HbF (table 2). Bone marrow was normocellular with myelopoiesis that was mildly prominent without significant dysplasia. Bone marrow karyotype was normal (46,XX) and peripheral blood chromosome breakage with DEB (diepoxybutane) was negative as well. Following recent infections, she was given GCSF 5 ug/kg 3 days per week. She died at 23 years of age at home, presumably due to complications secondary to neutropenia and infection.

Molecular studies

Standard Sanger sequencing of the entire coding sequence of *CFTR* and *SBDS* genes was done in patients IV.1 and IV.6.

Whole Exome Sequencing (WES)

WES was performed as described previously (Saudi Mendeliome Group, 2015). Briefly, one hundred nanograms of each DNA sample were treated to obtain the Ion Proton AmpliSeq library. DNA was then amplified in twelve separate wells using Exome Primer Pools, AmpliSeq HiFi mix (Thermo Fisher, Carlsbad, CA, USA), and 10 amplification cycles. All twelve PCR pools were combined in one well and subjected to primer digestion performing incubation with FuPa reagent (Thermo Fisher, Carlsbad, CA, USA). Amplified Exome targets were ligated with Ion P1 and Ion Xpress Barcode adapters. After purification, libraries were quantified using qPCR with the Ion Library Quantification Kit (ThermoFisher, Carlsbad, CA, USA). The prepared exome library was further used for emulsion PCR on an Ion OneTouch System and templated Ion Sphere particles were enriched using Ion OneTouch ES, both procedures following the manufacturer's instructions (ThermoFisher, Carlsbad, CA, USA). The template-positive Ion PI Ion Sphere particles were processed for sequencing on the Ion Proton instrument (ThermoFisher, Carlsbad, CA, USA). Approximately 15-17 Gb of DNA sequence was generated per sequencing run. Reads

were mapped to the UCSC hg19 (http://genome.ucsc.edu/) and variants identified using the Saudi Human Genome Program (SHGP) pipeline (Saudi Mendeliome Group, 2015.; Mustafa et al., 2018).

Computational structural analysis of mutants

Protein sequences of G6PC3 and MPL were retrieved from the Uniprot database (accession codes Q9BUM1 and P40238, respectively). I-TASSER (Yang and Zhang, 2015) and SwissModel (Waterhouse et al., 2018) were used to produce tertiary structure predictions. G6PC3 was modelled with I-TASSER (C-score of -3.02, estimated TM-score of 0.37±0.13) based on several threading templates, including the PAP2 type phosphatidylglycerol phosphate phosphatase from *Bacillus subtilis* (PDB ID: 5JKI), with a sequence identity of 23% to G6PC3. The 3D structure of residues 1-195 could be inferred with highest confidence, with a mean estimated local accuracy of 6.00 Å vs 11.71 Å for the rest of the sequence.

The tertiary and quaternary structures of the first two fibronectin type-III (FNIII) domains of MPL (residues 28-283) were modelled using SwissModel, based on ~28% sequence identity with the human erythropoietin receptor EpoR. Modelling using the entire fragment 28-283 resulted in a relatively poor QMEAN score of -5.0 for the monomer, based on several experimental templates (e.g. PDB IDs 1EER, 1CN4, 6MOJ, 6MOE). Deleting a long MPL-specific loop (residues 196-236) in the second domain increased the score to an acceptable QMEAN of -3.7 (e.g. for template 6MOE), showing that the low score for the entire fragment is largely due to the poorly modelled insertion. Models were manually inspected, and mutations were evaluated, using the Pymol program (pymol.org).

Plasma proteomics analysis in G6PC3 deficiency

Sample preparation

Peripheral blood samples obtained from the 3 subjects (IV.1, 2, 6) affected with G6PC3 deficiency and 6 healthy individuals without a history of hematological disorders were subjected to expression proteome analysis using label-free quantitative liquid chromatography-mass spectrometry (LC-MS) analysis.

Approximately 5 ml of peripheral blood was subjected to Ficoll centrifugation and the extracted plasma was depleted of high abundance proteins using the human albumin removal kit (Agilent Technologies). 2-DE gels of individual samples were run in duplicate as a quality control measure and assessed for both qualitative and quantitative changes prior to pooling. Only samples with high degree of similarity as judged by 2-DE signatures were pooled together into the analysis group while samples with marked variations were excluded. However; all samples that were left out of the pools together with other samples were used in the validation for individual variability of the observed changes in the pooled samples. The G6PC3 deficient samples (group 1) were pooled as well as the 6 healthy control samples (group 2). Each pool was run in triplicate. For each of the two study sample groups, 100 μg complex protein mixture was subjected to in-solution digestion and protein concentrations were adjusted to 1 μg/μl at the end of digestion as previously described (Alayia et al., 2016; Colak et al., 2016). For absolute quantitation, all samples were spiked with yeast alcohol dehydrogenase (ADH; P00330) and approximately 3 μg protein digests were loaded onto the analytical column for LC/MS analysis.

Expression Proteomics analysis by mass spectrometry

The digested peptides were run on a Nano-Acquity liquid chromatography Synapt G2 tandem mass spectrometry (Waters Corporation). Quantitative proteomics data were generated for the sample groups as previously described (Alayia et al., 2016; Colak et al., 2016). Data acquisition using ion mobility separation experiments (HDMSE) was performed and data were acquired over a range of m/z 50-2000 Da over 120 min acquisition time. All samples were analyzed in triplicate runs using the MassLynx programs (version. 4.1, SCN833; Waters) and Progenesis QI for proteomics (Progenesis QIfp version 2.0.5387; Nonlinear Dynamics, Waters Corporation) was used for all automated data processing and database searching. The generated peptide masses were searched against the unified non-redundant databases (Uniprot/SwissProt Human protein sequence database).

Proteome bioinformatics and data analysis

Progenesis QI for proteomics (QIfP) V3.0 was used to identify proteins that are significantly differentially expressed between the sample groups followed by filtering for only statistically (ANOVA), significantly regulated proteins ($P \leq 0.05$ and a fold change >1.5). While the Progenesis QI algorithm partially takes into consideration the issue of multiple testing and the False Discovery Rate, we used adjusted p-value or q-value and have applied power threshold as well as a minimum of 2-fold change as criteria to define our differentially expressed proteins. Also, 'Hi3' absolute quantification was performed using ADH as an internal standard to give an absolute amount of each identified protein.

For network analysis of differentially expressed proteins, we used the Ingenuity Pathway Analysis (IPA) software (Qiagen) where sample source and species of sample origin were selected while "other diseases or cell types or tissues" were excluded to search for protein-protein interactions.

Supplemental References

Alaiya, A.A., Aljurf, M., Shinwari, Z., Almohareb, F., Malhan, H., Alzahrani, H., Owaidah, T., Fox, J., Alsharif, F., Mohamed, S.Y., Rasheed, W., Aldawsari, G., Hanbali, A., Ahmed, S.O. and Chaudhri, N. (2016). Protein signatures as potential surrogate biomarkers for stratification and prediction of treatment response in chronic myeloid leukemia patients. International Journal of Oncology, 49(3), pp.913–933.

Colak, D., Alaiya, A.A., Kaya, N., Muiya, N.P., AlHarazi, O., Shinwari, Z., Andres, E. and Dzimiri, N. (2016). Integrated Left Ventricular Global Transcriptome and Proteome Profiling in Human End-Stage Dilated Cardiomyopathy. PLOS ONE, 11(10), p.e0162669.

Mustafa A., Faquih T., Baz B., Kattan R., Al-Issa A., Tahir A., Imtiaz F., Ramzan K., Al-Sayed M., Alowain M., Al-Hassnan Z., Al-Zaidan H., Abouelhoda M., Al-Mubarak B., Al Tassan N. (2018). Validation of Ion Torrent TM Inherited Disease Panel with the PGM TM Sequencing Platform for Rapid and Comprehensive Mutation Detection. Genes (Basel), 9(5):267.

Saudi Mendeliome Group. (2015). Comprehensive Gene Panels Provide Advantages Over Clinical Exome Sequencing for Mendelian Diseases. Genome Biol. Jun 26;16(1):134.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research, 46(W1), pp.W296–W303.

Yang, J. and Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Research, 43(W1), pp.W174–W181.