

Detection of new pathogenic mutations in patients with congenital haemolytic anaemia using next-generation sequencing

R. DEL ORBE BARRETO ¹, B. ARRIZABALAGA ^{1,2}, A. B. DE LA HOZ ¹, A. GARCIA-ORAD ^{1,3}, M. I. TEJADA ^{1,4}, J. C. GARCIA-RUIZ ^{1,2}, T. FIDALGO ⁵, C. BENTO ⁵, L. MANCO ⁶, M. L. RIBEIRO ⁵

¹ BioCruces Health Research Institute, Barakaldo, Spain

² Department of Haematology, Cruces University Hospital, Barakaldo, Spain

³ Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Medicine and Odontology, University of the Basque Country (UPV/EHU), Bilbao, Spain

⁴ Molecular Genetics Laboratory, Genetics Service, Cruces University Hospital, Barakaldo, Spain

⁵ Serviço de Hematologia Clínica, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

⁶ Research Centre for Anthropology and Health (CIAS), Department of Life Sciences, University of Coimbra, Coimbra, Portugal

Correspondence:

Rafael Del Orbe Barreto,
BioCruces Health Research
Institute, Hematologic Cancer
Study Group, Plaza de Cruces,
48903 Barakaldo, Spain.
Tel.: +34 946006000;
Fax: +34 946006172;
E-mail: rafaelandres.delorbebar
reto@osakidetza.eus

SUMMARY

Introduction: Congenital haemolytic anaemia (CHA) refers to a group of genetically heterogeneous disorders, mainly caused by changes in genes encoding globin chains, cytoskeletal proteins and red cell enzymes, in which accurate diagnosis can be challenging with conventional techniques.

Methods: To set-up a comprehensive assay for detecting mutations that could improve aetiological diagnosis, we designed a custom panel for sequencing coding regions from 40 genes known to be involved in the pathogenesis of CHA, using the Ion Torrent™ (Thermo Fisher Scientific, S.L. Waltham, MA, USA) Personal Genome Machine (PGM) Sequencer. A control group of 16 samples with previously known mutations and a test group of 10 patients with unknown mutations were included for assay validation and application, respectively.

Results: In the test group, we identified pathogenic mutations in all cases: four patients had novel mutations in genes related to membrane defects (SPTB, ANK1, SLC4A1 and EPB41), four were homozygous or compound heterozygous for mutations in genes related to enzyme deficiencies (GPI, TPI1 and GSS), one had a mutation in the HBB gene and another presented a homozygous mutation in the ADAMTS13 gene.

Conclusions: Ion PGM sequencing with our custom panel is a highly efficient way to detect mutations causing haemolytic anaemia, including new variations. It is a high-throughput detection method that is ready for application in clinical laboratories.

INTRODUCTION

Congenital haemolytic anaemia (CHA) refers to a diverse group of genetically heterogeneous disorders that are characterized by increased red blood cell (RBC) destruction, with consequences ranging from relatively harmless to severe life-threatening anaemia [1]. They can be due to defects in haemoglobin (Hb), RBC enzymes or RBC cytoskeletal proteins [2]. Occasionally, diagnosis of CHA is a challenge because clinical features are similar in cases with different aetiologies and it is not possible to distinguish between them using conventional diagnostic techniques [1, 2].

Generally, b-haemoglobinopathies are identified using chromatography and electrophoresis, but there are some cases such as those due to unstable haemoglobin or other Hb variants that can be misdiagnosed; hence, genetic analysis is required to detect variations in the HBB gene. Lately, next-generation sequencing (NGS) has demonstrated its usefulness as a diagnostic tool and in genetic counselling [3–5].

In relation to RBC enzyme deficiencies, most are inherited in an autosomal recessive form with haemolysis occurring only in homozygous or compound heterozygous individuals and some enzyme deficiencies are X-linked. The most common are glucose-6-phosphate dehydrogenase (G6PD) and pyruvate kinase (PK) deficiencies; however, there are a number of other much less frequent enzyme disorders that cause haemolytic anaemia without specific abnormalities in the morphology of the RBC. Diagnosis is based on detection of reduced specific enzyme activity in RBCs, but errors can be made in the interpretation of results because of instability of the enzyme, high reticulocyte count or recent blood transfusions. Hence, DNA sequence analysis may be the only way to establish a diagnosis [6].

RBC cytoskeletal disorders, including hereditary spherocytosis (HS), hereditary elliptocytosis (HE), hereditary pyropoikilocytosis and hereditary stomatocytosis, are caused by defects in various membrane proteins. These disorders are typically inherited as autosomal dominant traits or de novo but can also present as recessive forms, frequently severe. Diagnosis is based on RBC morphology, screening tests for membrane defects (osmotic fragility, acid glycerol lysis time and eosin-50-maleimide binding tests), ion flux measurement and analysis of RBC deformability by ektacytometry [7]. However, in neonates or transfused individuals, diagnosis can be difficult due to unclear morphological features, and screening tests may be unreliable [8]. In such cases, gene sequencing provides a powerful diagnostic tool [7, 8] and NGS has shown to be feasible [9]. Occasionally, it can be difficult to make a differential diagnosis between HS and congenital dyserythropoietic anaemias (CDAs), inherited disorders affecting red cell lineage, especially CDA-II. They are characterized by an ineffective erythropoiesis with specific morphological features in the bone marrow that have served as a basis for their classification; however, recent identification of several causative genes could help reclassify these disorders. Indeed, some authors have suggested that the use of NGS technologies will improve

molecular diagnosis in CDAs [10, 11]. In addition, cases of congenital thrombotic microangiopathy (TM) due to defects in the ADAMTS13 gene or in the complement system can be underestimated as a cause of CHA. Diagnosis is often hindered by the need for plasma exchange as an urgent treatment as this can produce an artefact in activity levels of the ADAMTS13 enzyme or complement fractions. In these patients, genetic studies are helpful to make an accurate diagnosis [12, 13].

Diagnosis of CHA based on genetic mutations is a useful strategy and provides additional insight into the mechanisms of the disease. On the other hand, the conventional Sanger sequencing-based approach for mutation screening in CHA is time-consuming and costly, and most importantly, may not detect all types of disease-causing alterations [9, 14, 15]. Thus, the aim of our study was to evaluate the feasibility of a custom panel based on NGS technology (Ion Torrent PGM™ Sequencer) for sequencing of the genes most frequently involved in the pathogenesis of CHA.

METHODS

This study was approved by the Euskadi Research Ethics Committee in accordance with the principles of the Declaration of Helsinki.

Patients

A total of 26 samples of genomic DNA were included, 16 patients (validation samples) with a previous CHA diagnosis by conventional biochemical methods and a genetic mutation identified by Sanger sequencing and 10 patients without a clear diagnosis. To explore the flexibility of our assay, we evaluated different types of pathogenic mutations (missense, nonsense and splice site mutations, as well as deletions and insertions) in the validation sample cohort. We investigated nine patients referred for CHA mutation screening and one patient with a haemoglobinopathy, without prior knowledge of their gene defect, to explore whether our NGS method also identified mutations in unclassified CHA patients. Of the former nine patients, four showed morphological and biochemical characteristics of an RBC membrane protein defect, one had no measurable ADAMTS13 activity and TM and four had chronic haemolysis but no distinctive morphological features, while the patient with a haemoglobinopathy had an Hb variant unidentified by capillary electrophoresis.

Gene panel

We developed an NGS assay to analyse a panel of 40 genes known to be involved in the pathogenesis of CHA (Table 1). Using Ampliseq Designer software (Thermo Fisher Scientific, S.L.), we selected the regions of interest to be sequenced, including all exons, 30- and 50-

untranslated regions, exon–intron boundaries and promoter regions. Additionally, various intronic regions of several genes were included to analyse copy number variation.

Library preparation and sequencing

For each sample, a double reaction was required for two pools of primers. To generate a multiplexed PCR using the Ion Ampliseq Library Kit 2.0 (Thermo Fisher Scientific, S.L.), 10 ng of DNA was mixed with each pool of primers and polymerase, as per the manufacturer's protocol. After that, Ion Xpress™-specific index paired-end barcode adapters (Thermo Fisher Scientific, S.L.) were ligated to the 50 and 30 ends of DNA fragments by incubation with DNA ligase for 30 min at 22 °C and 10 min at 72 °C. The adapter ligated DNA fragments were purified with Agencourt AMPure XP Reagent. The quantity of material in each library was assessed by qPCR using a specific quantification kit from Thermo Fisher Scientific, S.L. Pooled barcoded libraries were clonally amplified using the Ion OneTouch2™ (Thermo Fisher Scientific, S.L.) system according to the manufacturer's instructions. Ion sphere particles were enriched using the E/S module, and subsequently, the enriched template positive particles were loaded onto an Ion 316™ or 318™ chip and sequenced using Ion Hi-Q sequencing kit chemistry on an Ion Torrent PGM (Thermo Fisher Scientific, S.L.) [15–18].

Analysis

Postrun QC/QA filtering was performed using Torrent Suite (v3.6; Thermo Fisher Scientific, S.L.), and sequences aligned to the human genome version 19 using T-MAP (version 3.6.58977). Variants were called using the Torrent Variant Caller (version 3.6.59049, with Germ Line – Low Stringency settings) and annotated using wANNOVAR. The potential functional impact of identified variants was assessed using the predictive algorithms available in wANNOVAR. Finally, potentially causative mutations identified as being novel or rare and with some predictive indicators were confirmed by Sanger sequencing.

Detection of single nucleotide variations and small insertions/deletions (indels)

The data analysis pipeline developed for the detection of single nucleotide variations and small indels is comprised of tools freely available on the web. An initial quality check of the sequence reads is followed by mapping the paired-end reads with the Burrows-Wheeler Aligner to the National Centre for Biotechnology Information human genome version 19 reference genome. Subsequently, single nucleotide polymorphisms and small indels are called using the Ion Reporter™ software (Thermo Fisher Scientific, S.L.). The resulting list of variations is annotated with Annovar, which utilizes data from external databases to generate context around the mutations, such as the consequences of amino acid changes, location with respect to canonical splice-site regions and information related to the Short Genetic Variations database (dbSNP) and frequencies if available. Finally, a manual filtering step is executed to prioritize relevant mutations. Low-frequency frameshifts and truncating mutations are considered likely to be pathogenic. Unreported nonsynonymous amino acid variations were analysed in silico by the

pathogenicity prediction programs, Mutation Taster, PolyPhen-2, PROVEAN and SIFT, to help assess the damaging effect.

Mutation confirmation

Sanger sequencing was performed using an ABI 3130 genetic analyser (Applied Biosystems, Foster City, CA, USA) to confirm the DNA variants derived from the NGS. Sequence variants were compared with those on the dbSNP, 1000 Genomes, PAH mutation and PNDdb (formerly BIOMDB) databases. Additionally, enzyme activity analysis and SDS-PAGE were performed to confirm the phenotype in newly studied cases.

RESULTS

After NGS, an average of 300 000 unique reads were obtained per sample resulting in a median sequence depth of about 100-fold, with an average enrichment efficiency of >90%.

Several types of disease-causing genetic aberrations were present in the assayed DNA samples including single nucleotide substitutions, multiple nucleotide substitutions and small indels (1–8 nucleotides).

In the 16 samples from the validation cohort with known mutations, we detected the same mutations by NGS, except in one case and this was attributable to an error (discussed below) in the technique (Table 2).

On the other hand, we detected pathogenic mutations in all nine samples from patients with unclassified CHA and the carrier of the Hb variant (Table 3). In four patients with a suspicion of RBC membrane defects (Table 3, Cases 1–4), we found a mutation in related genes (SPTB, ANK1, SLC4A1 and EPB41). In the patient with abnormal ADAMTS13 activity and TM (Table 3, Case 5), we detected a mutation in the ADAMTS13 gene. Further, the four patients with CHA and no distinctive morphologic features (Table 3, Cases 7–10) showed mutations in various different enzyme genes, two of them previously reported (Cases 7 and 9). Lastly, the patient with a variant in HBB (Table 3, Case 6) was identified as having Hb-D Punjab [19]. Table 4 summarizes the characteristics of the patients included in the study.

In the four sequenced patients with clinical features of RBC membrane defects (Cases 1–4), we found novel pathogenic mutations. Case 1 had a mutation in the SPTB (spectrin beta) gene: c.167A>G (p.Gln56Arg), in heterozygosis. Morphological features in the proband (Case 1) were consistent with HS and cryohaemolysis, and eosin 5-maleimide binding tests were also positive. In Cases 2, 3 and 4, mutations at the heterozygous state were detected in ankyrin (ANK1), band-3 (SLC4A1) and protein 4.1 (EPB41), respectively (ANK1: p.Arg935Ter; SLC4A1: p.Leu417-Pro; and EPB41: p.Tyr442Ter). All these cases were confirmed by Sanger sequencing and SDS-PAGE of membrane proteins. Moreover, affected family members with CHA were tested and the same mutation was observed in them.

Case 5 was a 3-year-old girl with anaemia and thrombocytopenia. A blood smear revealed the presence of schistocytes and ADAMTS13 activity levels below 5%, without anti-ADAMTS13 antibodies. We detected the c.3070T>G (p.Cys1024Gly) mutation in the ADAMTS13 gene in homozygosis, previously reported to be causative of Upshaw–Schulman syndrome [13]. Her parents were cousins and were also sequenced by the Sanger method. They both had the same variation in heterozygosis in the ADAMTS13 gene.

In Case 6, NGS enabled us to identify the Hb-D variant, due to a specific mutation in the HBB gene (p.Glu122Gln).

In the remaining four cases (Cases 7–10), we identified mutations related to enzyme-deficiency haemolytic anaemia. All these were confirmed by Sanger sequencing and measurement of low enzyme activity. Cases 7 and 9 showed mutations in homozygosis corresponding to previously reported variations in the GPI (p.Lys550Glu) and TPI1 (p.Glu105Asp) genes, respectively. In Case 9, we confirmed the presence of variation in the TPI1 gene in both parents in heterozygosity. Cases 8 and 10 showed novel mutations in compound heterozygosity in the GPI (p.Arg81Gln; Gln413Argfs*24) and GSS (p.Arg125Leu; Phe218Cys) genes. None of the parents of proband cases showed features of haemolytic anaemia. In Cases 7 and 8, a glucose phosphate isomerase (GPI) assay was also performed, finding around 10-fold lower activity in the patients than in a healthy control (Case 7: 4.16 IU/g Hb, Case 8: 5.9 IU/g Hb, control: 54.0 IU/g Hb). In Case 10, levels of 5-oxoproline in urine and reduced glutathione in blood were measured and found to be notably lower than normal (46.4 mmol/mol creatinine, normal range: 26–63; and 9.37 mg/dL RBC, normal range: 54.7–82.7, respectively).

DISCUSSION

We have successfully applied our NGS method, using a novel targeted gene panel to identify the underlying genetic cause in cases of nondiagnosed CHA where standard testing had been unable to provide a definitive diagnosis or was not available. The current approach to the genetic diagnosis of CHA disorders involves Sanger sequencing of the coding region of genes known to be involved in pathogenicity. This process is slow, iterative, costly, and in many cases, still fails to identify the underlying genetic cause. Recent developments in NGS now allow its utilization in a practical, manageable and cost-efficient manner and have enabled our group to establish a platform for sequencing 40 candidate genes related to CHA in parallel, for the first time, and hence in a faster and cheaper way. We confirmed the presence of 15/16 of known pathogenic mutations in the validation samples with the correct homozygous / heterozygous state, representing a high sensitivity of mutant allele detection (94%). In the case in which the mutation was not detected (C05), the patient had a compound heterozygous mutation located in exon 3 of

the PKLR gene, but the specific primer failed and no amplicons were generated. This amplification error was observed repeatedly through the samples analysed, and this was resolved by using Sanger sequencing for the corresponding region. It is important to mention that the gene panel we have designed does not allow evaluation of some coding regions that remain uncovered or deep intronic mutations, and nor does it enable the identification of gene variants located in homopolymer stretches, large deletions or crossover variants, a well-known weakness of Ion PGM technology.

Molecular characterization of defects at the DNA level is essential for confirmation of the diagnosis, and identification of a causative mutation enables proper genetic counselling and prenatal diagnosis.

In some cases, DNA sequencing may be the only way to establish a diagnosis, especially in enzyme defects, as there are no characteristic morphological features, analysis of the activity of each RBC enzyme can be laborious and high reticulocyte count or recent blood transfusions can lead to misinterpretation of results. Moreover, recessive inheritance does not allow us to identify deficiency in parents.

Likewise, there could be suspected cases of RBC membrane protein defects in which genetic studies may be helpful to rule out stomatocytosis and avoid surgical procedures such as splenectomy when they would not be appropriate.

The present work demonstrates the usefulness of our NGS panel for diagnosing CHA. Among the patients with unclassified CHA, we found four causative mutations in the ADAMTS13, GPI and TPI1 genes (Cases 5, 6, 7 and 9) which had previously been classified as deleterious [13, 20, 21]. Eight novel variants were detected in SPTB, ANK1, SLC4A1, EPB41, GPI and GSS (GPI and GSS, compound heterozygous) (Table 3). Segregation of these variants was analysed by Sanger sequencing.

We conclude that the variant in SPTB (Case 1) is a de novo mutation, because parents (clinically normal) were studied and no mutations were found; while for the ANK1 (Case 2), SLC4A1 (Case 3) and EPB41 (Case 4) genes, the same variants were found in family members with CHA, indicating a dominant hereditary pattern. Two compound heterozygous cases were found in the GPI and GSS genes (Case 8 and 10, respectively), and in both cases, each variant was present in the parental genomes in heterozygosity.

In Case 1, morphological features were consistent with HS; a cryohaemolysis and eosin-50-maleimide binding tests were also positive. Gln56 is highly conserved in mammals, and all computational tools that were checked, namely Mutation Taster, PolyPhen and SIFT, predicted this variant to be damaging. The p.Gln56Arg mutation is located in the calponin homology domain, an actin-binding domain that is found in proteins essential for the regulation of cell shape and signalling proteins [22]. This variant is not on the 1000 Genomes or dbSNP databases. All the available data lead us to believe that this variant is deleterious.

The ANK1 variant (c.2803C>T, Case 2) is also present in the affected mother in heterozygosity. It has not been reported previously as a single nucleotide polymorphism, and all computational tools predicted that it would be ‘disease causing’. This variant leads to a premature stop codon, and hence, the protein loses the 945 amino acids in the C-terminal region, including the ZU5 domain which can function as a spectrin-binding domain in ankyrin [23, 24].

The variant found in SLC4A1 (c.1250T>C, Case 3) is present in the proband and affected members of the family (sister and daughter), and thus, we conclude that it is a novel variant not previously reported with an autosomal dominant hereditary pattern. In addition, all computational tools used predicted it to be a damaging variant [25]. The p.Leu417Pro change has not been described previously. It takes place in the transmembrane domain 1, and it has been suggested that this is one of the transmembrane domains of this protein that provides access to the extracellular medium, being negatively affected the anion exchange. Leucine is a hydrophobic amino acid that prefers to be buried in protein hydrophobic cores and in this case it is within α -helices; however, proline is unable to easily occupy many of the main-chain conformations and it can also introduce kinks into α -helices, as it is unable to adopt a normal helical conformation. Leu417 is highly conserved with a PhyloP score of 4.651 and a PhastCons score of 1. The PROVEAN score is -6.535 , indicating that the change has a damaging effect (-2.5 being considered the threshold for deleterious variants).

The EPB41 protein, together with spectrin and actin, constitutes the red cell membrane cytoskeletal network. The resulting complex plays a critical role in erythrocyte shape and deformability [26]. The variant found in EPB41 (c.1326T>A, Case 4) is not reported in the most widely used databases and variant effect predictors indicated that it was deleterious, due to a premature stop codon that leads to the loss of 422 amino acids, obtaining a truncated protein that lacks function.

In Cases 8 and 10 with GPI and GSS deficiency, respectively, two heterogeneous alleles – compound heterozygous pattern – have been found to cause disease. This is the condition of having two heterogeneous recessive alleles at a particular locus that can cause genetic disease in a heterozygous state. Each variant was detected in the patient’s parents in heterozygosis, and none of them showed features of haemolytic anaemia. However, GPI and GSS deficiency were detected in both patients (Case 8 and 10, respectively).

In Case 8, a frameshift (c.1238_1238delA; p.Gln413Arg*24) and a missense (c.212G>A; p.Arg81Gln) variation have been found in GPI. The frameshift variation leads to a premature stop codon causing the loss of 111 amino acids in the C-terminal region of the protein that is a highly conserved region. It is known that C-terminal domains are crucial to the sugar ring-opening step catalysed by this enzyme. As the isomerization activity occurs at the dimer interface, the dimer structure of this enzyme is critical to its catalytic function [27]. In addition, this variant was not

found in either the Exome Aggregation Consortium or 1000 Genomes databases and all predictive software used classified it as deleterious. In this patient, a missense variant of the other allele was found, this conversion of arginine 81 to glutamine has a pretty low allele frequency (Minor Allele Frequency: 0.0002) and was classified as damaging by the bioinformatics tools that were used, with a PROVEAN score of -3.163 .

Finally, in Case 10, the GSS gene has two variants, both missense, in a heterozygous state. Variants in arginine 125 were previously reported to be pathological in the Human Gene Mutation Database (CM960786 – p.Arg125Cis- and CM052884 – p.Arg125His-). This amino acid is highly conserved; being a substrate binding site with PhyloP and Phast-Cons scores of 6.003 and 0.987, respectively, the variation p.Arg125Leu gives a PROVEAN score of -6.007 . The other mutation p.Phe218Cys has not previously been described and nor is it found in the Exome Aggregation Consortium ExAC or 1000 Genomes databases, but it is predicted to be pathogenic by all in silico predictive tools that were checked. This amino acid is also highly conserved with PhyloP and Phast-Cons scores of 4.648 and 1, respectively.

In conclusion, we have presented an innovative specific targeted gene panel based on NGS platforms for the diagnosis of patients with CHA and described several cases illustrating the utility of this approach. The findings of this study demonstrate that our custom NGS panel is an efficient and effective method for the diagnosis of CHA and that this technology has matured sufficiently for use in clinical laboratories.

ACKNOWLEDGEMENTS

The author was supported by a grant from BBK – Bio-Cruces (BBK-BC/POST-MIR1250T/2015/001).

CONFLICT OF INTERESTS

The authors declare no conflict of interests.

REFERENCES

1. Clarke GM, Higgins T. Laboratory investigation of hemoglobinopathies and thalassemias: review and update. *Clin Chem* 2000;46:1284–90.
2. Christensen RD, Nussenzveig RH, Yaish HM, Henry E, Eggert LD, Agarwal AM. Causes of hemolysis in neonates with extreme hyperbilirubinemia. *J Perinatol* 2014;34:616–9.

3. Shooter C, Rooks H, Lay S, Clark B. Next generation sequencing identifies a novel rearrangement in the HBB cluster permitting to-the-base characterization. *Hum Mutat* 2015;36:142–50.
4. Hassan S, Vosses R, Chessa R, de Dunnes J, Bakker E, Giordano P, Hartevelde C. Molecular diagnostics of HBB gene in an Omani cohort using bench-top DNA Ion Torrent PGM technology. *Blood Cells Mol Dis* 2014;53:133–7.
5. Cao A, Galanello R. Beta-thalassemia. *Genet Med* 2010;12:61–76.
6. Koralkova P, Van Solinge WW, Van Wijk R. Rare hereditary red blood cell enzymopathies associated with hemolytic anemia – pathophysiology, clinical aspects, and laboratory diagnosis. *Int J Lab Hematol* 2014;36:388–97.
7. King MJ, Garc_on L, Hoyer JD, Iolascon A, Picard V, Stewart G, Bianchi P, Lee SH, Zanella A, International Council for Standardization in Haematology ICSH guidelines for the laboratory diagnosis of nonimmune hereditary red cell membrane disorders. *Int J Lab Hematol* 2015;37:304–25.
8. Bolton-Maggs P, Langer J, Iolascon A, Tittensor P, King M. Guidelines for the diagnosis and management of hereditary spherocytosis – 2011 update. *Br J Haematol* 2012;156:37–49.
9. Ameziane N, Sie D, Dentro S, Ariyurek Y, Kerkhoven L, Joenje H, Dorsman JC, Ylstra B, Gille JJ, Sistermans EA, de Winter JP. Diagnosis of fanconi anemia: mutation analysis by next-generation sequencing. *Anemia* 2012;2012:132856.
10. Iolascon A, Esposito MR, Russo R. Clinical aspects and pathogenesis of congenital dyserythropoietic anemias: from morphology to molecular approach. *Haematologica* 2012;97:1786–94.
11. Iolascon A, Russoa R, Delaunay J. Congenital dyserythropoietic anemias. *Curr Opin Hematol* 2011;18:146–51.
12. Campistol JM, Arias M, Ariceta G, Blasco M, Espinosa L, Espinosa M, Griny_o JM, Mac_ia M, Mendiz_abal S, Praga M, Roman E, Torra R, Vald_es F, Vilalta R, Rodriguez de C_ordoba S. An update for atypical haemolytic uraemic syndrome: diagnosis and treatment. A consensus document. *Nefrologia* 2015;35:421–47.

13. Lotta L, Garagiola I, Palla R, Cairo A, Peyvandi F. ADAMTS13 mutations and polymorphisms in congenital thrombotic thrombocytopenic purpura. *Hum Mutat* 2010;31:11–9.
14. McDonald CJ, Ostini L, Wallace DF, Lyons A, Crawford DH, Subramaniam VN. Next-generation sequencing: application of a novel platform to the analysis of atypical iron disorders. *J Hepatol* 2015;63:1288–93.
15. Zhang J, Barbaro P, Guo Y, Alodaib A, Li J, Gold W, Ades L, Keating BJ, Xu X, Teo J, Hakonarson H, Christodoulou J. Utility of next-generation sequencing technologies for the efficient genetic resolution of haematological disorders. *Clin Genet* 2016; 89:163–72.
16. Buermans HPJ, den Dunnen JT. Next generation sequencing technology: advances and applications. *Biochim Biophys Acta* 2014;1842:1932–41.
17. Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet* 2014;30:418–26.
18. Smith A, Boycott KM, Jarinova O. Lake Louise mutation detection meeting 2013: clinical translation of next-generation sequencing requires optimization of workflows and interpretation of variants. *Hum Mutat* 2014;35:265–9.
19. Torres LS, Okumura JV, Silva DG, Bonini-Domingos CR. Hemoglobin D-Punjab: origin, distribution and laboratory diagnosis. *Rev Bras Hematol Hemoter* 2015;37:120–6.
20. Repiso A, Oliva B, Vives-Corrons JL, Beutler E, Carreras J, Climent F. Red cell glucose phosphate isomerase (GPI): a molecular study of three novel mutations associated with hereditary nonspherocytic haemolytic anemia. *Hum Mutat* 2006;27:1159.
21. Pekrun A, Neubauer BA, Eber SW, Lakomek M, Seidel H, Schröter W. Triosephosphate isomerase deficiency: biochemical and molecular genetic analysis for prenatal diagnosis. *Clin Genet* 1995;47:175–9.
22. An X, Debnath G, Guo X, Liu S, Lux SE, Baines A, Gratzler W, Mohandas N. Identification

and functional characterization of protein 4.1R and actin-binding sites in erythrocyte beta spectrin: regulation of the interactions by phosphatidylinositol-4,5-bisphosphate. *Biochemistry* 2005;44:10681–8.

23. Mohler PJ, Yoon W, Bennett V. Ankyrin-B targets beta2-spectrin to an intracellular compartment in neonatal cardiomyocytes. *J Biol Chem* 2004;279:40185–93.

24. Steiner JP, Bennett V. Ankyrin-independent membrane protein-binding sites for brain and erythrocyte spectrin. *J Biol Chem* 1988;263:14417–25.

25. Barneaud-Rocca D, Etchebest C, Guizouarn H. Structural model of the anion exchanger 1 (SLC4A1) and identification of transmembrane segments forming the transport site. *J Biol Chem* 2013;288:26372–84.

26. Parra M, Gascard P, Walensky LD, Snyder SH, Mohandas N, Conboy JG. Cloning and characterization of 4.1G (EPB41L2), a new member of the skeletal protein 4.1 (EPB41) gene family. *Genomics* 1998;49:298–306.

27. Cordeiro AT, Godoi PH, Silva CH, Garratt RC, Oliva G, Thiemann OH. Crystal structure of human phosphoglucose isomerase and analysis of the initial catalytic steps. *Biochim Biophys Acta* 2003;1645:117–22.

Table 1. Genes included in the panel.

Gene symbol	Disease
<i>CDAN1</i>	Congenital dyserythropoietic anaemia type I
<i>SEC23B</i>	Congenital dyserythropoietic anaemia type II
<i>KLF1, GATA1</i>	Others types of congenital dyserythropoietic anaemia
<i>HBB</i>	Haemoglobinopathies and beta thalassaemia
<i>SPTA1, SPTB</i>	Hereditary elliptocytosis, spherocytosis, pyropoikilocytosis (alpha or beta spectrin)
<i>ANK1</i>	Hereditary spherocytosis (ankyrin)
<i>EPB72, PIEZO1</i>	Hereditary stomatocytosis (stomatatin, piezo1)
<i>GYPB, GYPC</i>	Hereditary elliptocytosis (glycophorin A, C)
<i>SLC4A1</i>	Hereditary spherocytosis (band 3)
<i>EPB41</i>	Hereditary elliptocytosis (protein 4.1)
<i>EPB42</i>	Hereditary spherocytosis (protein 4.2)
<i>G6PD</i>	Anaemia due to glucose-6-phosphate dehydrogenase deficiency
<i>HK1</i>	Anaemia due to hexokinase deficiency
<i>GPI</i>	Anaemia due to glucose-6-phosphate isomerase deficiency
<i>AK1</i>	Anaemia due to adenylate kinase deficiency
<i>GSS</i>	Anaemia due to glutathione synthetase deficiency
<i>PKLR</i>	Anaemia due to pyruvate kinase deficiency
<i>PGK</i>	Anaemia due to phosphoglycerate kinase deficiency
<i>TP11</i>	Anaemia due to triose phosphate isomerase deficiency
<i>ALDOA</i>	Anaemia due to aldolase deficiency
<i>PFKFB1</i>	Anaemia due to phosphofructokinase deficiency
<i>BPGM</i>	Anaemia due to biphosphoglyceromutase deficiency
<i>NT5C3A</i>	Anaemia due to pyrimidine 5'-nucleotidase deficiency
<i>CFH, CFHR1, CFHR3, CFHR4, CFI, CFB, CD46, THBD, C3, DGKE</i>	Atypical haemolytic uraemic syndrome
<i>ADAMTS13</i>	Upshaw–Schulman syndrome
<i>PIGA, PIGT</i>	Paroxysmal nocturnal haemoglobinuria

Table 2. Previously known mutations in validation samples

Sample ID	Clinical diagnosis	Affected gene	Location	Previously known mutation	Amino acid change	Zygoty	Found by NGS
003	G6PD deficiency	<i>G6PD</i>	chrX:153762634	c.653C>T	p.Ser218Phe	Heterozygous	+
006	G6PD deficiency	<i>G6PD</i>	chrX:153763492	c.292G>A	p.Val98Met	Compound	+
			chrX:153764217	c.466A>G	p.Asn156Asp	heterozygous	
008	G6PD deficiency	<i>G6PD</i>	chrX:153764154	c.352_354delTTC	p.Phe118del	Hemizygous	+
010	Hb D	<i>HBB</i>	chr11:5246908	c.364G>C	p.Glu122Gln	Heterozygous	+
011	β-Thalassaemia minor	<i>HBB</i>	chr11:5247992	c.126_129delCTTT	p.Phe42Leufs*19	Heterozygous	+
012	Hb S	<i>HBB</i>	chr11:5248232	c.20A>T	p.Glu7Val	Homozygous	+
022	β-Thalassaemia major	<i>HBB</i>	chr11:5248004	c.118C>T	p.Gln40Ter	Compound	+
			chr11:5248050	c.93-21G>A	Intronic Variation	heterozygous	
026	Hb Johnstown	<i>HBB</i>	chr11:5246944	c.328G>C	p.Val110Leu	Heterozygous	+
029	G6PD deficiency	<i>G6PD</i>	chrX:153762604	c.683G>A	p.Arg228His	Hemizygous	+
030	G6PD deficiency	<i>G6PD</i>	chrX:153760909	c.1250G>A	p.Arg417His	Hemizygous	+
033	HbC/β-thalassaemia minor	<i>HBB</i>	chr11:5248154	c.92+6T>C	Intronic variation	Compound	+
			chr11:5248233	C.19G>A	P.Glu7Lys	heterozygous	
C01	β-Thalassaemia	<i>HBB</i>	chr11:5246857	c.404_405delTTG	p.Val135Glyfs*5	Compound	+
			chr11:5246866	c.411_414delTGTG	p.Val138Leufs*158	heterozygous	
C03	Pyruvate kinase deficiency	<i>PKLR</i>	chr11:55271258	c.-72A>G	Promoter region	Homozygous	+
C04	β-Thalassaemia minor	<i>HBB</i>	chr11:5248154	c.92+6T>C	Intronic variation	Heterozygous	+
C05	Pyruvate kinase deficiency	<i>PKLR</i>	chr11:55264079	c.1063A>G	p.Met355Val	Compound	ND
			chr11:55264080	c.1064T>C	p.Met355Thr	heterozygous	ND
C06	Congenital dyserythropoietic anaemia	<i>GATA1</i>	chrX:48651700	c.866A>G	p.His289Arg	Heterozygous	+

Table 3. Mutations identified in patients with unclassified CHA

Case	Gene	Location	Transcript	Mutation (Human Gene Mutation Database accession)	Amino acid change	Zygoty
1	<i>SPTB</i>	chr14:65271790	NM_001024858.2	c.167A>G	p.Gln56Arg	Heterozygous
2	<i>ANK1</i>	chr8:41554038	NM_000037	c.2803C>T	p.Arg935Ter	Heterozygous
3	<i>SLC4A1</i>	chr17:42335386	NM_000342.3	c.1250T>C	p.Leu417Pro	Heterozygous
4	<i>EPB41</i>	chr1:29359718	NM_001166005.1	c.1326T>A	p.Tyr442Ter	Heterozygous
5*	<i>ADAMTS13</i>	chr9:136319562	NM_139025.4	c.3070T>G	p.Cys1024Gly	Homozygous
				(CM012893)		
7*	<i>GPI</i>	chr19:34890912	NM_000175.3	c.1648A>G	p.Lys550Glu	Homozygous
				(CM065245)		
8	<i>GPI</i>	chr19:34857716 chr19:34887530	NM_000175.3	c.242G>A c.1238_1238delA	p.Arg81Glnp. Gln413Arg fs*24	Compound heterozygous
9*	<i>TP11</i>	chr12:6978338	NM_000365.5	c.315G>C	p.Glu105Asp	Homozygous
				(CM860016)		
10	<i>GSS</i>	chr20:33524782	NM_000178.2	c.374G>T	p.Arg125Leu	Compound heterozygous
		chr20:33530408		c.653T>G	p.Phe218Cys	

*Previously reported to be pathogenic.

Table 4. Characteristics of the patients included in the study.

Case	Hb level (g/dL)	Reticulocyte count (%)	Lactic acid dehydrogenase level (U/L)	Total bilirubin (mg/dL)	RBC morphology	Haemoglobin electrophoresis pH 8.6	Eosin-5'-maleimide binding test	Osmotic fragility	Enzyme levels
1	11.1	11.4	301	2	Spherocytosis	Normal	Positive	High	PK: Normal G6PD: Normal
2	7.3	14.1	577	1.1	Spherocytosis	Normal	Positive	High	NP
3	16	2.3	240	1.7	Spherocytosis	Normal	positive	High	NP
4	14	2.7	215	1	Elliptocytosis	Normal	Negative	Slightly high	NP
5	9.7	1.6	685	1	Schistocytes	Normal	NP	NP	ADAMTS13 undetectable
6	13.6	1.5	221	0.9	No alterations	Slow migration band not identified 41%	NP	NP	NP
7	7.1	11.3	1497	7.9	Basophilic stippling	Normal	NP	Normal	GPI: 4.16 UI/g Hb (46-66)
8	13.5	4.1	169	1	No alterations	Normal	NP	Normal	GPI: 5.9 UI/g Hb (46-66)
9	9.2	18	386	2.6	Equinocytes, dacryocytes	Normal	Negative	Normal	TP1: 395 UI/g Hb (1361-2769) PK: Normal
10	12.8	6	192	0.8	No alterations	Normal	Negative	Normal	5-oxoprolin in urine: 46.4 mmol/mol creatinine Reduced Glutathion: 9.37 mg/dL RBC (54.7-82.7)

NP, not performed.