UNIVERSITY OF NAPOLI FEDERICO II

Doctorate School in Molecular Medicine

Doctorate Program in
Genetics and Molecular Medicine
Coordinator: Prof. Lucio Nitsch
XXVIII Cycle

"A new approach to Congenital Dyserythropoietic Anemias: toward a better definition of molecular mechanisms"

Roberta Russo

Napoli 2016

A new approach to Congenital Dyserythropoietic Anemias: toward a better definition of molecular mechanisms

Table of contents

List of publications related to the thesis (2013-present)

1. Gambale A, Iolascon A, Andolfo I, **Russo R**. Diagnosis and management of Congenital Dyserythropoietic Anemias. Expert Rev Hematol. 2016;9:283-96

2. Andolfo I, **Russo R**, Manna F, Shmukler BE, Gambale A, Vitiello G, De Rosa G, Brugnara C, Alper SL, Snyder LM, Iolascon A. Novel Gardos channel mutations linked to dehydrated hereditary stomatocytosis (Xerocytosis). Am J Hematol. 2015;90:921-6

3. Di Pierro E*, **Russo R***, Karakas Z, Brancaleoni V, Gambale A, Kurt I, Winter SS, Granata F, Czuchlewski DR, Langella C, Iolascon A, Cappellini MD. Congenital Erythropoietic Porphyria linked to GATA1-R216W mutation: challenges for diagnosis. Eur J Haematol. 2015;94:491-7

4. **Russo R**, Gambale A, Langella C, Andolfo I, Unal S, Iolascon A. Retrospective cohort study of 205 cases with congenital dyserythropoietic anemia type II: Definition of clinical and molecular spectrum and identification of new diagnostic scores. Am J Hematol. 2014;89:E169-75

5. Unal S, **Russo R**, Gumruk F, Kuskonmaz B, Cetin M, Sayli T, Tavil B, Langella C, Iolascon A, Uckan Cetinkaya D. Successful hematopoietic stem cell transplantation in a patient with congenital dyserythropoietic anemia type II. Pediatr Transplant. 2014;18:E130-3

6. Andolfo I, Alper SL, De Franceschi L, Auriemma C, **Russo R**, De Falco L, Vallefuoco F, Esposito MR, Vandorpe DH, Shmukler BE, Narayan R, Montanaro D, D'Armiento M, Vetro A, Limongelli I, Zuffardi O, Glader BE, Schrier SL, Brugnara C, Stewart GW, Delaunay J, Iolascon A. Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1. Blood. 2013;121:3925-35

7. **Russo R**, Langella C, Esposito MR, Gambale A, Vitiello F, Vallefuoco F, Ek T, Yang E, Iolascon A. Hypomorphic mutations of SEC23B gene account for mild phenotypes of congenital dyserythropoietic anemia type II. Blood Cells Mol Dis. 2013;51:17-21

8. Andolfo I, Alper SL, Delaunay J, Auriemma C, **Russo R**, Asci R, Esposito MR, Sharma AK, Shmukler BE, Brugnara C, De Franceschi L, Iolascon A. Missense mutations in the ABCB6 transporter cause dominant familialpseudohyperkalemia. Am J Hematol. 2013;88:66-72.

Manuscript in preparation or submitted

1. **Russo R**, Andolfo I, Manna F, Gambale G, Pignataro P, De Rosa G, Iolascon A. RedPlex: a targeted next generation sequencing-based diagnosis for patients with hereditary hemolytic anemias. (in preparation)

2. **Russo R**, Andolfo I, Gambale A, Vallefuoco F, Esposito MR, Asci R, Wandroo F, Iolascon A. Regulatory network GATA1-mediated on SEC23B gene in erythroid cells. (in preparation)

3. **Russo R**, Andolfo I, Manna F, De Rosa G, De Falco L, Gambale A, Bruno M, Mettè A, Girelli D, De Franceschi L, Iolascon A. ERFE-encoding *FAM132B* in patients with Congenital Dyserythropoietic Anemia type II. (in preparation)

4. Andolfo I, **Russo R**, Manna F, De Rosa G, Gambale A, Zouwail S, Detta N, Alper SL, Brugnara C, Sharma AK, De Franceschi L and Iolascon A. Functional characterization of novel ABCB6 mutations and their clinical implications in familial pseudohyperkalemia. (in revision Haematologica 2016)

List of Abbreviations

AGLT, acidified glycerol lysis test
BM, bone marrow
CDA, congenital dyserythropoietic anemia
CEP, congenital erythropoietic porphyria
CHC, cryohydrocytosis
COP, cytoplasmic coat protein
DBA, Diamond-Blackfan anemia
DHS, stomatocytosis
E:G, erythropoietic/granulopoietic ratio
EM, electron microscopy
EMA, eosin-5-maleimide
ER, endoplasmic reticulum
FA, Fanconi anemia
FP, familial pseudohyperkalemia
HA, hyporegenerative anemia
HAMD, hemolytic anemia due to red cell membrane defects
Hb, hemoglobin
HDW, Hb distribution width
HE, hereditary elliptocytosis
HHA, hereditary hemolytic anemia
HPP, hereditary pyropoikilocytosis
HS, hereditary spherocytosis
HST, hereditary stomatocytosis
IBMFS, inherited bone marrow failure syndrome
LDH, lactate dehydrogenase
MAF, minor allele frequency
MCHC, mean corpuscular Hb concentration
MCV, mean cell volume
NGS, Next generation sequencing
OF, osmotic fragility
OHS, overhydrated hereditary stomatocytosis
OMIM, Online Mendelian Inheritance in Man
PBMC, peripheral blood mononuclear cell
PLT, platelet
qRT-PCR, quantitative RT-PCR
RBC, red blood cell
RDW, RBC distribution width
RhAG, Rh-associated glycoprotein
ROI, regions of interest
SAO, Southeast Asian ovalocytosis
SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis
t-NGS, targeted-NGS
WB, western blotting
WES, whole-exome sequencing
XLTDA, X-linked thrombocytopenia with or without dyserythropoietic anemia
αLELY, Low Expression Lyon
sTfR, soluble transferrin receptor

Abstract

Hereditary hemolytic anemias (HHAs) embrace a highly heterogeneous group of chronic disorders with a highly variable clinical picture. HHA encompass (1) hyporegenerative anemias (HAs), as congenital dyserythropoietic anemias (CDAs); (2) hemolytic anemias due to red cell membrane defects (HAMDs), as hereditary spherocytosis (HS) and hereditary stomatocytosis (HST). Although the workflow to diagnose these conditions is a normal clinical practice, differential diagnosis, classification, and patient stratification among HHAs are often very difficult. Beyond achieving a definitive diagnosis, knowing the genetic basis of these patients can be valuable also for guiding treatment. Next generation sequencing (NGS) refers to non-Sanger-based high-throughput DNA sequencing technologies. This technology plays a major role either in disease gene discovery or in clinical use for establishing a genetic diagnosis. Particularly, the major current application of NGS in diagnostics is through design of disease specific panel, named targeted (t)-NGS, in which a selected fraction of genes is sequenced.

The primary aim of our study was the development of a fast, accurate, reliable and cost effective diagnostic/prognostic tool for HHAs based on t-NGS. In order to assess the reliability of this approach we created a t-NGS gene panel, named RedPlex, composed by 34 *loci* causative or candidates of HHAs. *In silico* design was performed by Agilent SureDesign web tool. For each locus, all coding regions, 5' and 3'UTRs, and 100 bp flanking splice junctions were included. Sequence length was set at 150×2 nucleotides, and the predicted target size amounted to 538 regions (239.764 kb). Targeted enrichment was performed on 32 patients from 27 unrelated families by HaloPlex Target Enrichment System. High-throughput sequencing was performed by Illumina NextSeq 500. SureCall software was used for bioinformatic and computational analyses. RedPlex panel showed high sensitivity and specificity. It was able to capture at least 99.4% of 538 target regions with high and uniform coverage. We were able to obtain a conclusive diagnosis in approximately 72% of cases. In addition, a lot of patients (39%) showed multiple disease-associated variants suggesting complex inheritance. Indeed, t-NGS approach also allows the identification of "polygenic" genotypes, which may account for the phenotypic variability among HHA patients. Thus, the secondary aim of this project was the study of the interaction between mutated genes in HHA patients. We particularly focused on the functional interaction between two CDA-related genes, *GATA1* and *SEC23B*. The demonstration of the direct interaction of GATA1 transcription factor on the *SEC23B* promoter provided also an explanation of the variability of phenotypes GATA1-related by means of the crosstalk of this gene with its target *SEC23B*.

1. Background

Hereditary hemolytic anemias (HHAs) embrace a highly heterogeneous group of disorders characterized by hemolytic anemia of variable degree and by complex and often unexplained genotype-phenotype correlations. HHA are genetic disorders caused by mutations in more than 30 genes controlling red blood cell (RBC) production and structure. Mutations in these genes can lead to alterations in hemoglobin (Hb) levels, RBC differentiation and proliferation, cell membrane structure, defective activity of erythrocyte enzymes. This large group of pathologies comprises hemoglobinopathies, thalassemias, hyporegenerative anemias (HAs), hemolytic anemias due to red cell membrane defects (HAMDs). We will focus on (1) HAs, as congenital dyserythropoietic anemias (CDAs), (2) HAMDs, as hereditary spherocytosis (HS) and hereditary stomatocytosis (HST).

1.1 Classification, diagnostic criteria and epidemiology of CDAs

The term dyserythropoiesis refers to a condition of abnormal erythropoiesis affecting the differentiation and proliferation pathways of the erythroid lineage with a consequent defective production of RBCs (Iolascon et al. 2011). Dyserythropoietic anemias can be divided into primary and secondary forms, and both inherited and acquired types can occur. Among these different conditions, the CDAs are hereditary diseases that embrace a highly heterogeneous set of rare or very rare anemias that result from various kinds of abnormalities during late stages of erythropoiesis. They are counted as subtypes of inherited bone marrow failure syndromes (IBMFS), characterized by morphological abnormalities of erythroblasts in the bone marrow (BM) and ineffective erythropoiesis as predominant mechanism of anemia, accompanied by a hemolytic component (Iolascon et al. 2013).

CDAs can be suspected in the presence of anemia and hemolytic signs, accompanied with reticulocytosis inadequate to the degree of anemia. In particular, the following criteria are required: (1) evidence of anemia, jaundice, splenomegaly; (2) evidence of ineffective erythropoiesis; (3) occurrence of typical morphological features of erythroblasts at BM examination; and (4) exclusion of other congenital anemias that fulfill criteria 1 and 2, such as thalassemia syndromes and other IBMFS (Gambale et al. 2016). The BM of CDA patients is always hypercellular, due to an exclusive and pronounced increase of erythroblasts, with erythropoietic/granulopoietic ratio (E:G) of 4 to 10 (normal reference values 0.3 - 1.0). The expansion of the erythropoietic tissue leads to elevated serum concentration of the soluble transferrin receptor (sTfR), if iron deficiency is excluded. Dyserythropoiesis appears to be a morphological feature common to several conditions, and this could account for the difficulties in diagnosis of CDAs. However, the specific morphological

alterations of the erythroid precursors justify the heterogeneity of these disorders. Indeed, the three classical types of CDAs (types I, II and III) are defined on the basis of BM morphology (Iolascon et al. 2013). Inclusions of additional CDAs, the so called CDA variants, despite remarkable morphological studies, gradually led to overlapping entities and imposed a limitation on classification. In spite of these difficulties, morphological classification is still widely used in clinical practice (Table 1-1).

Table 1-1. Morphological classification of CDAs

Disease symbol	Phenotype	Bone marrow biopsy	
		Optical microscopy	*Electron microscopy*
CDA I	Congenital dyserythropoietic anemia type I	Binucleated erythroblasts (3-7%); thin chromatin bridges between nuclei of erythroblasts	"Swiss cheese appearance" of the erythroblasts heterochromatin
CDA II	Congenital dyserythropoietic anemia type II	Binucleated (10-30%); rare multinucleate erythroblasts	Double plasma membrane of the erythroblasts
CDA III	Congenital dyserythropoietic anemia type III	Giant multinucleate (up to 12 nuclei) erythroblasts	Clefts within heterochromatin, autophagic vacuoles, iron-laden mitochondria, myelin figures in the cytoplasm
CDA IV	Congenital dyserythropoietic anemia type IV	Tri- and multi-nucleate erythroblasts	Invagination of nuclear membrane, intra-nuclear precipitated and nuclear blebbing
XLTDA	Thrombocytopenia X-linked with or without dyserythropoietic anemia	Erythroblasts: megaloblastic features, nuclear irregularities, bi- and multi-nucleation Megakaryocytes: Small, dysplastic with signs of incomplete maturation	Reduced numbers of platelet alpha granules and dysplastic features in megakaryocytes and platelets

Nevertheless, the identification of the causative genes of the most common forms among CDAs in the last two decades represented an evident advantage for reclassifying these disorders, as well as in understanding their pathogenesis. Moreover, uncovering the molecular basis of CDAs helped to unravel novel aspects of the molecular biology of erythropoiesis. From the genetic standpoint, six different types of CDAs are included in the Online Mendelian Inheritance in Man (OMIM) compendium of human genes and genetic phenotypes so far (Table 1-2). However, this list is bound to extend. Indeed, the discovery of new causative genes is a continuous evolving process thanks to the development of high-throughput technologies, such as next generation sequencing (NGS). The identification of genetic variations underlying hereditary disorders marked the opening of a new era for genetic and clinical research. Indeed, beyond, obtaining definitive diagnosis and planning patient management, knowledge of the genetic basis of these disorders is crucial in

estimating their prevalence and geographical distribution (Gambale et al. 2016).

Table 1-2. Classification of CDAs by OMIM database

Disease symbol	Phenotype	Phenotype MIM number	Gene *location*	Inheritance	N cases[§]
CDA Ia	Congenital dyserythropoietic anemia type Ia	224120	CDAN1 *15q15.2*	AR	< 100
CDA Ib	Congenital dyserythropoietic anemia type Ib	615631	C15orf41 *15q14*	AR	< 10
CDA II	Congenital dyserythropoietic anemia type II	224100	SEC23B *20p11.23*	AR	> 200
CDA III	Congenital dyserythropoietic anemia type III	105600	KIF23 *15q21*	AD	< 20
CDA IV	Congenital dyserythropoietic anemia type IV	613673	KLF1 *19p13.2*	AD	< 10
XLTDA	Thrombocytopenia X-linked with or without dyserythropoietic anemia	300367	GATA1 *Xp11.23*	XLR	< 10

[§] Number of cases with positive molecular analysis
AD, Autosomal dominant; AR, Autosomal recessive; XLR, X-linked recessive

In the majority of CDAs, inheritance is autosomal recessive, and due to the small number of offspring in most European families, single cases in one family are the rule rather than the exception. Together with the rarity of the disorder and the need to obtain bone marrow specimens for diagnosis, this explains why correct diagnosis is often delayed (Heimpel and Iolascon 2009), particularly in mild cases, even when anemia and/or hyperbilirubinemia have been evident for many years (Figure 1-1).

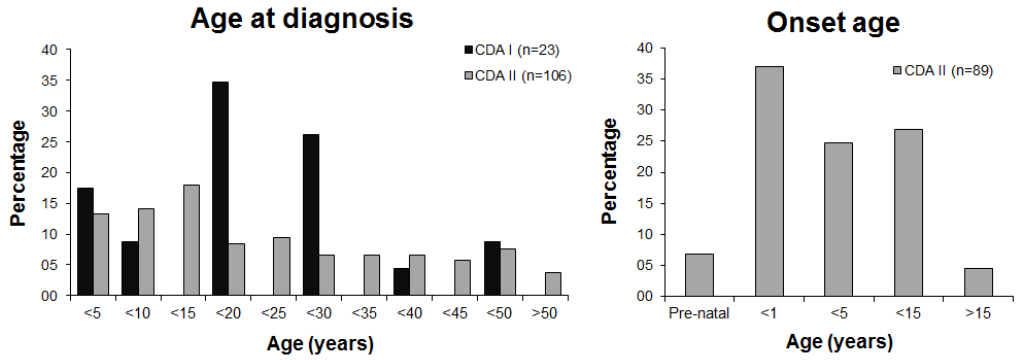

Figure 1-1. Age at diagnosis in CDA I and CDA II and onset age in CDA II. CDA I and II are the most common types among CDAs. Adapted from Heimpel and Iolascon 2009.

According to the estimation by Heimpel and colleagues in 2010, the prevalence of CDAs varies widely among European regions, with minimal values of 0.04 cases/million in North Europe and the highest in Mediterranean countries, particularly in Italy (2.49/million). This is mainly true for CDA II, which is more frequent than CDA I with an overall ratio of approximately 3.0 (Heimpel et al. 2010b). The studies on molecular epidemiology of CDA I and II highlighted the elevated allelic heterogeneity of both conditions as most of the causative variations are inherited as private mutations (Iolascon et al. 2012). However, recurrent mutations have been found in both CDA I and CDA II patients. In particular, the Arg1042Trp mutation in the CDA I causative gene *CDAN1* is a founder mutation in the Bedouin population (Tamary et al. 2008), whereas the amino acid substitutions Glu109Lys and Arg14Trp in the CDA II causative gene *SEC23B* show the highest frequency in Mediterranean area (Morocco, Israel and Italy), where a founder effect for both variants has been described (Russo et al. 2011) (Figure 1-2).

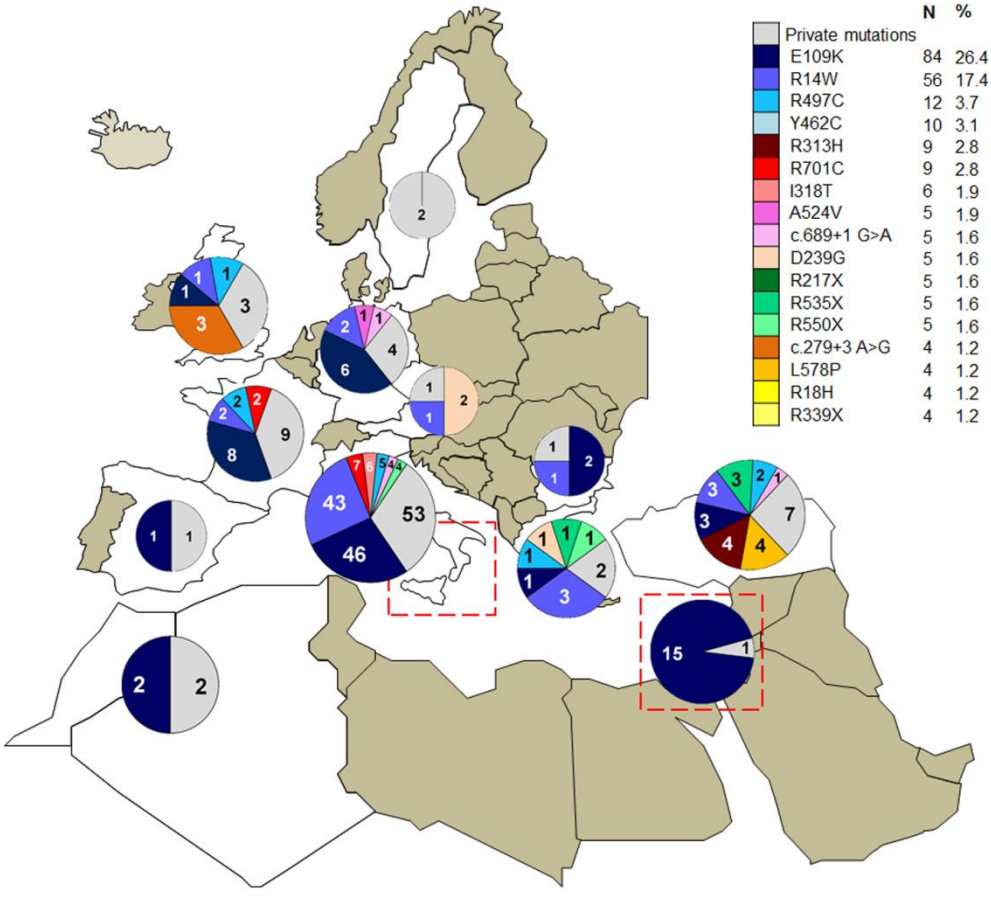

Figure 1-2. Molecular geocode of *SEC23B* alleles in Europe. In white are represented the countries in which have been reported CDA II cases so far. Mutations found in more than three independent pedigrees were shown with different colors. For each variant the total number of alleles and the relative frequency were indicated. Private mutations (i.e., the mutations present

in less than 3 unrelated patients) were colored in gray. Pie charts show the number of alleles for each mutation. Dotted red line highlights the region in which a founder effect has been demonstrated. Adapted from Russo et al. 2014.

However, molecular geocode of CDA II pathogenic alleles suggested the presence of multiple founder effects in different geographical areas of the world (Russo et al. 2014) (Figure 1-3).

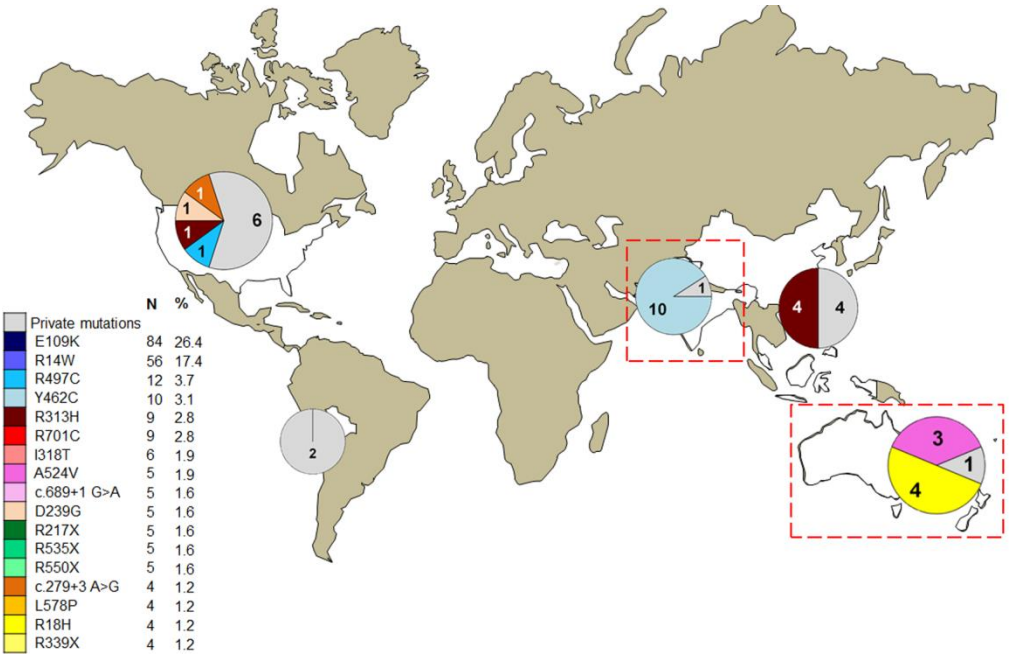

Figure 1-3. Molecular geocode of *SEC23B* alleles in the world. In white are represented the countries in which have been reported CDA II cases so far. Mutations found in more than three independent pedigrees were shown with different colors. For each variant the total number of alleles and the relative frequency were indicated. Private mutations (i.e., the mutations present in less than three unrelated patients) were colored in gray. Pie charts show the number of alleles for each mutation. Dotted red line highlights the region in which a founder effect has been suggested (Southeast Asia and Oceania). Adapted from Russo et al. 2014.

The pathomechanisms of CDAs involve the regulation of both DNA replication and cell division. Particularly, CDA II, CDA III and probably CDA Ib could be due to deregulation of mechanisms involved in cell division. Indeed, SEC23B is a component of the midbody, an essential structure in the telophase; KIF23 mutant causes the furrow regression, thus inhibiting the cytokinesis; *C15orf41* might encode a protein with homology to the Holliday junction resolvases, which are involved in chromosome segregation. Otherwise, the pathogenic mechanisms of transcription factor-related CDAs could be due to impairment of mechanisms involved in DNA synthesis and chromatin assembly. Indeed, during cell division, the transcription factors GATA1 and KLF1 could be

retained focally within mitotic chromatin to facilitate the rapid reactivation of the transcription of tissue-specific gens upon entry into G1 phase. Similarly, CDA Ia-Ib could be due to impaired mechanisms involved in chromatin assembly. Indeed, codanin-1 interacts with the cytosolic Asf1-H3-H4-importin-4 complex, involved in nucleosome assembly and disassembly, while C15orf41 interacts with Asf1b (Figure 1-4).

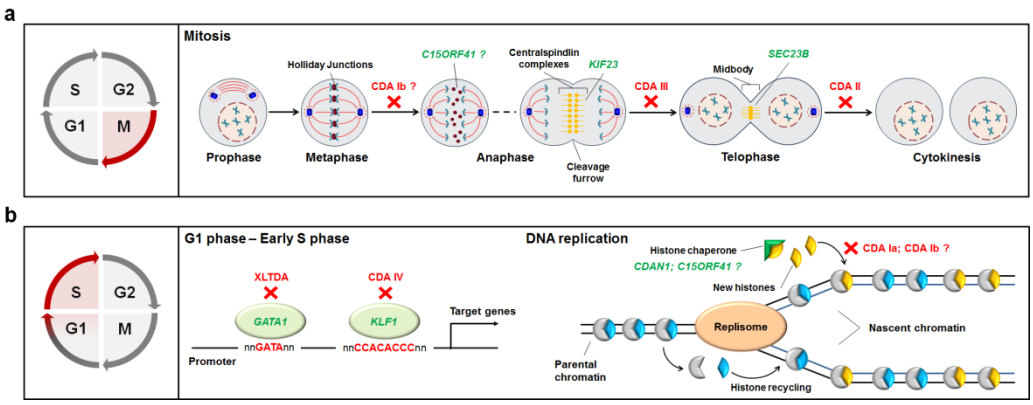

Figure 1-4. Pathogenetic mechanisms of CDAs involve the regulation of DNA replication and cell division. (a) The pathogenic mechanisms of CDA II, CDA III and probably CDA Ib could be due to deregulation of mechanisms involved in cell division. (b) The pathogenic mechanisms of transcription factor-related CDAs, as well as of CDA Ia–Ib could be due to impairment of mechanisms involved in DNA synthesis and chromatin assembly. Adapted from Gambale et al. 2016.

CDA type I

Most of the CDA I patients exhibit lifelong macrocytic anemia with variable values of Hb and mean cell volume (MCV) ranging between 100 and 120 fL; however, it can be normocytic in childhood. Relative reticulocytopenia is typically observed accompanied to increased Hb turnover as attested by indirect hyperbilirubinemia, high lactate dehydrogenase (LDH) value and low or absent plasma haptoglobin. All patients develop splenomegaly in adolescence or adulthood, while about 20% of cases show congenital anomalies, particularly syndactyly in hands or feet, absence of nails or supernumerary toes, pigeon chest deformity and short stature (Iolascon et al. 2012; Gambale et al. 2016). BM smear shows hypercellularity and erythroid hyperplasia (E:G of 4 and 8 times the normal). Approximately 30-60% of polychromatic erythroblasts show abnormalities of nuclear and chromatin structure. Indeed, the morphological pathognomonic feature of CDA I is the presence of thin chromatin bridges between the nuclei pairs of erythroblasts. A minority of erythroblasts shows bi- or multinuclearity, but in contrast to CDA II, the nuclei of binucleated cells are of different size and shape. At electron microscopy (EM), heterochromatin is denser than normal and forms

demarcated clumps with small translucent vacuoles, giving rise to the metaphor of "Swiss cheese appearance" (Heimpel et al. 2010a). CDA I is inherited in an autosomal-recessive manner. The first causative gene in which pathogenic variants were identified has been *CDAN1* (chr15q15.2) (Dgany et al. 2002) that encodes a ubiquitously expressed protein, codanin-1 (Table 1-2). More than 100 patients (CDA Ia) and 30 unique disease-causing mutations have been described so far. Codanin-1 is part of the cytosolic Asf1-H3-H4-importin-4 complex, which is implicated in nucleosome assembly and disassembly (Figure 1-4). Cdan1 knockout mice die in utero before the onset of erythropoiesis, suggesting a critical role of codanin-1 in developmental processes beyond erythropoiesis (Renella et al. 2011). Of note, no homozygous patients for null mutations have been described so far. Homozygous or compound heterozygous in *CDAN1* gene cover approximately 50% of CDA I patients, while in 30% of cases only a single mutant allele can be identified. Recently, the second causative gene of CDA I has been identified. In particular, two different mutations in *C15orf41* gene (chr 15q14) were found in three unrelated Pakistani families, classified as affected by CDA Ib (Table 1-2) (Babbs et al. 2013). C15orf41 is predicted to encode a divalent metal ion-dependent restriction endonuclease, with a yet unknown function. In cultured erythroblasts, C15orf41 produces a spliced transcript encoding a protein with homology to the Holliday junction resolvases (Figure 1-4). However, it has been shown that C15orf41 interacts with Asf1b, supporting the hypothesis that the primary defect in CDA Ib is in DNA replication and chromatin assembly.

CDA type II

The main clinical finding to diagnose CDA II is the presence of normocytic anemia of variable degree, with normal or only slightly increased reticulocyte count, but not adequate to the degree of anemia (ineffective erythropoiesis); it is often accompanied with jaundice and splenomegaly due to the hemolytic component. As described in a recent survey on 205 patients, CDA II generally presents mild anemia (mean Hb 9.6 ± 0.2 g/dL), but a wide spectrum of clinical presentations can occur, from asymptomatic to severe (Hb range 3.6-16.4 g/dL). Indeed, approximately 10% of cases result symptomless, whereas 20% of patients undergo a regimen of transfusion dependence (Russo et al. 2014). The term "congenital" could sound inappropriate in some cases since this disorder is often diagnosed during adulthood. Of note, the mean age of onset symptoms is approximately 3-4 years, but six patients were reported to have a prenatal onset. However, the age of diagnosis is delayed, 22.2 ± 1.7 years. This could be explained either by the occurrence of mild symptoms or by the misdiagnosis of CDA II with HS. Similarly to CDA I, the BM is hypercellular with distinct erythroid hyperplasia and subsequent increased E:G. The most specific finding of CDA II marrow is the presence of more than 10% mature binucleated erythroblasts with equal size of two nuclei. Upon EM examination, mature erythroblasts show discontinuous double membrane, which is due to the presence of vesicles loaded with proteins of endoplasmic reticulum (ER) that

appear to be running beneath the plasma membrane. Analysis of RBC membrane proteins by sodium dodecyl sulfate polyacrylamide (SDS-PAGE) gel electrophoresis reveals a narrower band size and faster migration of band 3 in most of the CDA II patients (95%) (Figure 1-5). Thus, this biochemical feature represents a specific diagnostic hallmark of the disease. Moreover, the hypoglycosylation of band 3 has been associated to the occurrence of the hemolytic component observed in CDA II patients by means of increased clusterization of this protein on RBC surface, which in turn leads to IgG binding and phagocytosis of RBCs (De Franceschi et al. 1998).

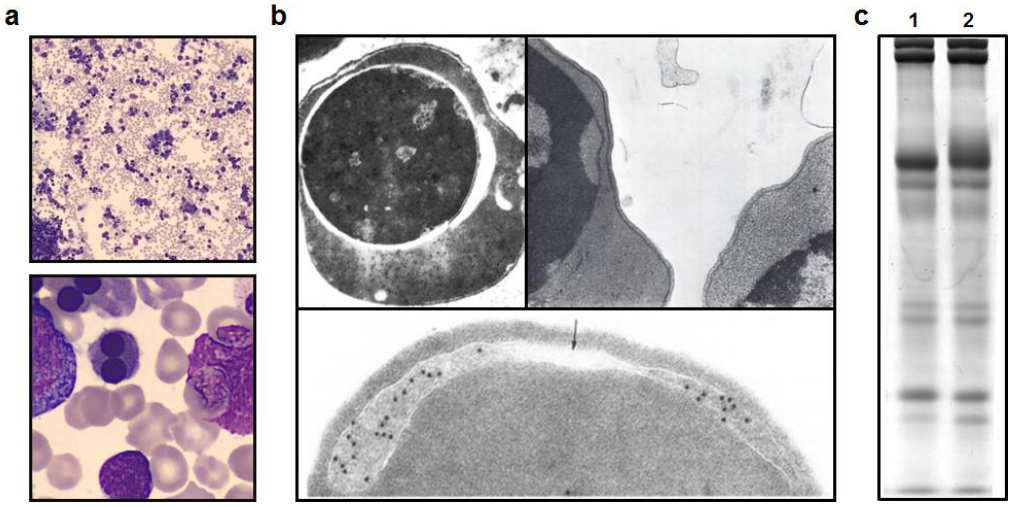

Figure 1-5. Morphological and biochemical features of CDA II erythroblasts. (a) CDA II bone marrow at light microscopy highlights erythroid hyperplasia with bi- or multinucleated late erythroid precursors. (b) CDA II bone marrow at EM highlights morphological abnormalities of CDA II erythroblasts. (c) Biochemical analysis of RBC membrane proteins from CDA II patient (lane 1) shows the typical hypoglycosylation of Band 3, with an increased anodic mobility on SDS-PAGE compared to healthy control (lane 2).

Both morphological features of the BM and biochemical alterations of RBC membrane proteins can be explained by the mutations in the causative gene *SEC23B* (chr 20p11.23) (Table 1-2) (Schwarz et al. 2009). CDA II is an autosomal recessive disease that belongs to the group of cytoplasmic coat protein (COP) II-related human genetic disorders (Russo et al. 2013a). *SEC23B* gene encodes the homonymous member of COPII complex, which is involved in the secretory pathway of eukaryotic cells. This multi-subunit complex mediates anterograde transport of correctly folded cargo from the ER toward the Golgi apparatus. Most of the cases (86%) show biallelic mutations in the *SEC23B* gene, although a subset of patients with an incomplete pattern of inheritance (14.0%) has been identified. More than 80 different mutations in SEC23B have been described so far, even if recurrent variants have been also described (Figure 1-6).

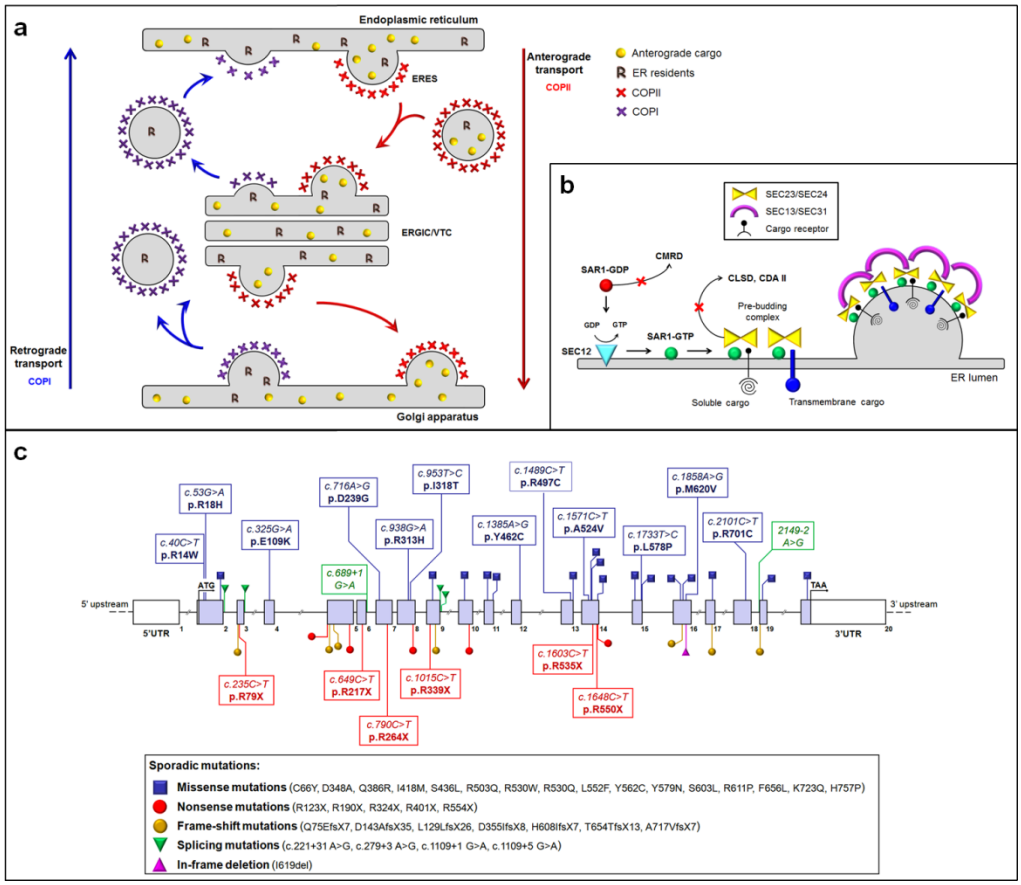

Figure 1-6. SEC23B function and mutations. (a) Schematic representation of ER-Golgi transport. After translation, folded nascent proteins are exported from the ER in COPII anterograde transport vesicles. In mammalian cells, COPII vesicles generate a structure known as ER-Golgi intermediate compartment (ERGIC). The ERGIC is a site for concentrating retrograde cargo into COPI vesicles, which bud from pre-Golgi and Golgi compartments to recycle vesicle components and retrieve resident proteins that have escaped the ER. Adapted from Russo et al. 2013a. (b) Model for COPII vesicle assembly. In yeast, COPII-coated vesicles form by the sequential binding of Sar1-GTP, the inner complex proteins Sec23- Sec24 and the outer complex components Sec13-Sec31 on ER. The transport of both integral membrane cargo and soluble secretory cargo is shown. Adapted from Russo et al. 2013a. (c) Localization of mutations in SEC23B gene. Exons and introns from a reference sequence (Ensembl transcript ID: ENST00000377475) are shown, distances to scale. Above the physical map, missense and splicing mutations are shown. Below the map, nonsense mutations (direct stop codons and frameshift mutations) and the only one in frame deletion are shown. The common mutations (found in at least two alleles) are shown in the boxes, whereas those that are sporadic (only one allele mutated) are represented by symbols, as indicated in the key. Adapted from Iolascon et al. 2011

Despite the elevated number of described cases and identified mutations, the pathomechanism of CDA II is not yet well understood. Although no direct evidence exists, SEC23B could play an active role in assembly or

deconstruction of the midbody, where it was identified in a proteomic screen (Skop et al. 2004). The presence of SEC23B in this subcellular compartment could account for an explanation of the impaired cytokinesis observed in CDA II erythroblasts (Figure 1-4). Otherwise, the multinuclear phenotype could be secondary to the aberrant glycosylation of specific proteins required for cell division, leading to defects in this process. However, it remains to unravel how alterations in a ubiquitous gene can result in clinical manifestations restricted to the erythropoietic tissue. The specificity of the CDA II phenotype seems to be due to the tissue-specific expression of SEC23B during erythroid differentiation (Schwarz et al. 2009). Alternatively, it could be explained by the presence of erythroid-specific cargoes (such as band 3), which might require high levels and full function of a specific COPII component to be correctly transported (Russo et al. 2013a). The pathophysiology study is difficult mainly due to the absence of a reliable animal model. Different models of SEC23B-deficient mice have been generated without reproducing CDA II phenotype. Indeed, SEC23B deficiency results in different phenotypes in humans and mice (Tao et al. 2012; Khoriaty et al. 2014). In particular, the absence of phenotype in SEC23B-deficient mice seems to be related to the different SEC23B/SEC23A expression ratio in murine and human tissues. Indeed, this ratio is higher in mouse pancreas compared to BM, whereas it is higher in human BM relative to pancreas. Of note, SEC23A and SEC23B are paralogous components of the COPII complex. This observation is in agreement with the compensatory expression of SEC23A that seems to ameliorate the effect of low SEC23B expression alleles in CDA II patients (Russo et al. 2013b). So far, the only reliable *in vitro* model for CDA II is the SEC23B-silencing in K562 cells, which recapitulates the cytokinesis defect, with a significant increase of percentage of binuclearity and an increased size of nuclei in SEC23B-silenced cells (Schwarz et al. 2009).

CDA type III

CDA III is the rarest form among classical CDAs. Patients show absent to moderate anemia with normal or slightly elevated MCV, and normal or faintly low relative number of reticulocytes. Twenty percent of patients received blood transfusions. Common symptoms are weakness, fatigue and headache; jaundice and biliary symptoms are also reported. Unlike from other CDAs, none of the described patients show enlarged liver or spleen. Hemolysis is also present as attested by low or absent haptoglobin and increased LDH. No significant changes in serum iron, transferrin or ferritin concentrations are observed. In some cases, serum electrophoresis analysis showed an M-component, IgG-k type: in particular, one patient with myeloma and four with monoclonal gammopathy were described. At optical microscopy, BM highlights erythroid hyperplasia, with the characteristic giant multinucleate erythroblasts; at EM, clefts within heterochromatin, autophagic vacuoles, iron-laden mitochondria and myelin figures in the cytoplasm were described. In 2013, the causative gene of this condition was identified by target sequencing: *KIF23* gene (chr

15q21) encodes a kinesin-superfamily protein MKLP1 that is a component of centralspindlin, a subcellular structure required for proper formation of the central spindle and the midbody and thus essential for cytokinesis (Liljeholm et al. 2013). MKLP1 mutant affects the function of this protein during cytokinesis, leading to the formation of the large multinucleated erythroblasts found in BM of the patients (Table 1-2; Figure 1-4).

Transcription factor-related CDAs

CDA IV and X-linked thrombocytopenia with or without dyserythropoietic anemia (XLTDA) belong to this subgroup of CDAs (Figure 1-4).

To date, four patients with CDA IV have been reported. All of them are characterized by the occurrence of normocytic anemia, generally severe, with Hb 5-9.5 g/dL. Conversely to classical CDAs, reticulocyte count is normal or slightly increased with respect to the degree of anemia; elevated values of HbF (>30%) are also observed (Arnaud et al. 2010; Jaffray et al. 2013). At BM, light microscopy erythroid hyperplasia, dyserythropoiesis signs, as basophilic stippling of polychromatic erythroblasts and erythrocyte, and internuclear bridging are observed. EM shows immature erythroid progenitors with atypical cytoplasmic inclusions, enlarged nuclear pores, invagination of nuclear membrane and marked heterochromatin. All four patients exhibit the same autosomal-dominant mutation (p.Glu325Lys), in heterozygous state, in *KLF1* gene (chr 19p13.2) (Table 1-2). *KFL1* encodes the homonymous protein, which is an essential erythroid-specific transcription factor, member of the Krüppel-like factor family. KLF1 is a well-known transcriptional activator in erythropoiesis, but it also exerts transcriptional repression in megakaryopoiesis. It plays a critical role in regulating the switch between fetal and adult Hb expression and is required in terminal erythroid differentiation for the cell-cycle progression.

XLTDA is a CDA variant characterized by anemia of variable degree, ranging from hydrops fetalis and transfusion-dependency to dyserythropoiesis without anemia, macro-thrombocytopenia with hypo-granulated platelets (PLT) and bleeding tendency. BM features are dyserythropoiesis, megakaryocytes decreased in number with cytoplasmic vacuoles and absence of platelet membrane demarcation. This is an X-linked recessive disease (Table 1-2), due to mutations in the X chromosomal gene *GATA1* gene (chr Xp11.23), encoding for the zinc finger DNA binding protein GATA1. This latter belongs to the GATA family of transcription factors, involved in the regulation of hematopoiesis. In particular, GATA1 plays an essential role in the development and maintenance of both erythroid and megakaryocytic lineages. GATA1 has two zinc finger domains: the C-terminal is necessary for DNA binding, while the N-terminal mediates interaction with FOG1 (friend of GATA1), a cofactor of GATA1. Of note, the most likely pathogenic mechanism in these disorders involves the GATA1-FOG1 interaction.

The role of transcriptional regulator of different genes and pathways may explain how different mutations in the same gene can result in disparate phenotypes. Indeed, beyond CDA IV, *KLF1* mutations have been associated to in(Lu)blood type and hereditary persistence of fetal hemoglobin. Similarly, there are different syndromic conditions related to *GATA1* mutations, in which thrombocytopenia can be associated to thalassemia, congenital erythropoietic porphyria or Diamond–Blackfan anemia (DBA)-like disease. Moreover, the co-inheritance of GATA1 and other CDA-gene mutations could explain the occurrence of more severe phenotypes (Di Pierro et al. 2015).

1.2 Classification, diagnostic criteria and epidemiology of HAMDs

Red cell membrane disorders are inherited diseases due to mutations in genes encoding for several membrane or cytoskeletal proteins of RBCs, resulting in decreased red cell deformability, reduced half-life and premature removal of the erythrocytes from bloodstream. Extensive studies on RBC membrane have allowed the comprehension of both structure and function of this subcellular compartment (Figure 1-7). Thus, the molecular bases of the overwhelming majority of cases of RBC membrane alteration have been completely defined.

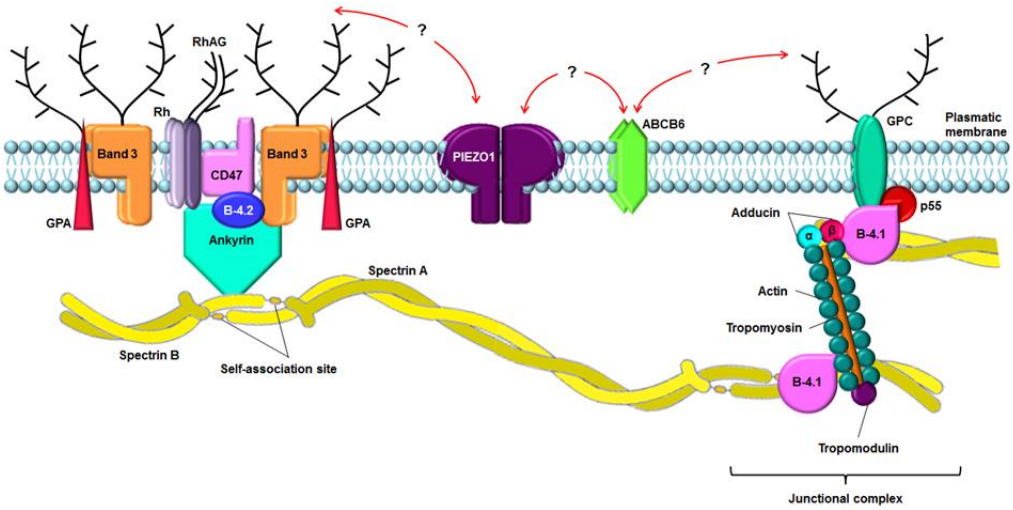

Figure 1-7. A simplified cross-section of the erythrocyte membrane. The RBC membrane is composed of integral membrane proteins incorporated into a phospholipid bilayer. The network of cytoskeletal proteins is anchored to the membrane via the transmembrane proteins. GPA, glycophorin A; Rh, Rhesus polypeptide; B-4.1, protein band 4.1; B-4.2, protein band 4.2; GPC, glycophorin C; RhAG, Rh-associated glycoprotein.

The red cell membrane disorders include both (i) structural defects and (ii) altered permeability of RBC membrane. The first subgroup comprises: HS, hereditary elliptocytosis (HE), hereditary pyropoikilocytosis (HPP), Southeast

Asian ovalocytosis (SAO); the second one: dehydrated hereditary stomatocytosis (DHS), overhydrated hereditary stomatocytosis (OHS), familial pseudohyperkalemia (FP), and cryohydrocytosis (CHC) (Table 1-3).

Table 1-3. Classification of HAMDs by OMIM database

Disease symbol	Phenotype	Phenotype MIM number	Gene	Location	Inheritance
HS1	Hereditary spherocytosis type 1	182900	ANK1	8p11.21	AD
HS2	Hereditary spherocytosis type 2	616649	SPTB	14q23.3	AD
HS3	Hereditary spherocytosis type 3	270970	SPTA1	1q23.1	AR
HS4	Hereditary spherocytosis type 4	612653	SLC4A1	17q21.31	AD
HS5	Hereditary spherocytosis type 5	612690	EPB42	15q15.2	AR
HE1	Hereditary elliptocytosis 1	611804	EPB41	1p35.3	AD
HE2	Hereditary elliptocytosis 2	130600	SPTA1	1q23.1	AD
HE3	Hereditary elliptocytosis 3	-	SPTB	14q23.3	AD
HPP	Hereditary Pyropoikilocytosis	266140	SPTA1	1q23.1	AR
SAO	Ovalocytosis Southeast Asian type	166900	SLC4A1	17q21.31	AD
OHS	Overhydrated hereditary stomatocytosis	185000	RHAG	6p12.3	AD
DHS1	Dehydrated hereditary stomatocytosis with or without pseudohyperkalemia and/or perinatal edema	194380	PIEZO1	16q24.3	AD
DHS2	Dehydrated hereditary stomatocytosis 2	616689	KCNN4	19q13.31	AD
FP	Familial pseudohyperkalemia 2	609153	ABCB6	2q35-q36	AD
CHC	Cryohydrocytosis	185020	SLC4A1	17q21.31	AD

AD, Autosomal dominant; AR, Autosomal recessive

HAMDs due to RBC structural defects

HS, also known as the Minkowski Chauffard disease, is the most common inherited red cell membrane disorder with an estimated prevalence of 1:2000-5000 in the Caucasian population. However, this value is probably even higher, due to under-diagnosed moderate forms. HS refers to a group of heterogeneous inherited anemias showing a broad spectrum of clinical severity, ranging from asymptomatic condition to severe form, characterized by transfusion-dependent life-threatening anemia and rarely by hydrops fetalis or fetal death. The main clinical findings of HS are hemolytic anemia of variable degree, hyperbilirubinemia, jaundice, splenomegaly, cholelithiasis. Laboratory findings are heterogeneous too: increased reticulocyte count (range 6% - >10%, until to 35% in severe cases); increased mean corpuscular Hb concentration (MCHC, ≥ 34.5g/dl) in > 50% of patients; increased RBC distribution width (RDW, >14) in most of patients; increased number of spherical-shaped erythrocytes (spherocytes) in peripheral blood smear. The first tests for HS diagnosis are:

NaCl osmotic fragility (OF), on both fresh blood and blood incubated at 37°C after 24h), standard and acidified glycerol lysis test (AGLT), pink test and flow cytometric analysis of eosin-5-maleimide-labeled erythrocytes (EMA binding). Unfortunately, neither EMA binding nor OF testing are able to detect 100% of HS patients, especially mild HS, since OF has low sensitivity. The quantitation of major erythrocyte membrane proteins via SDS-PAGE allows the identification of different subsets of HS patients; however, some subjects remain unclassified by this technique. Finally, ektacytometry is a highly sensitive test of membrane deformability (Gallagher et al. 2013). The phenotype variability is linked to different molecular defects associated to extent of loss of membrane surface area relative to intracellular volume, which leads to spherically shaped erythrocytes with decreased deformability (Table 1-3). The increased membrane fragility is caused by highly heterogeneous molecular defects due to deficiency and/or dysfunction in erythrocyte membrane proteins, particularly ankyrin (ANK1), α- and β-spectrin (SPTA and STPB), band 3 (SLC4A1), and protein 4.2 (EPB42) (Figure 1-7). Approximately 75% of HS cases exhibit an autosomal dominant pattern of inheritance, associated with mutations in *ANK1*, *SPTB* and *SLC4A1* genes. In the remaining 25% of patients, autosomal recessive *de novo* mutations in *SPTA* and *EPB42* genes were observed (Table 1-4).

Table 1-4. Genetic and biochemical features of HS, HE and HPP

Disease	Inheritance	Genes	Biochemical features of RBC membrane proteins[§]
HS	75% AD 25% *de novo*/AR	ANK1 (50% of cases) SPTB (20% of cases) SPTA1 (rare) SLC4A1 (20% of cases) EPB42 (rare)	-Mutations in ANK1: ↓ spectrin and band 4.2 -Mutations in SPTB: ↓ spectrin -Mutations in SLC4A1: ↓ band 3 and 4.2 -Mutations in EPB42: absence of band 4.2
HE	AD	EPB41 (severe forms) SPTB SPTA1	-Mutations in EPB41: absence of band 4.1 and ↓ of glycophorin C and D -Mutations in SPTA1 and SPTB: ↑ dimer/tetramer ratio on non-denaturing gel
HPP	AR	SPTA1 mutations associated with αLELY variant	

AD, Autosomal dominant; AR, Autosomal recessive
[§]Biochemical features at SDS-PAGE analysis

HE belongs to a heterogeneous group of disorders characterized by the presence of elliptical-shaped erythrocytes (elliptocytes) at peripheral blood smear with variable clinical manifestations, from asymptomatic carrier state to severe, transfusion-dependent hemolytic anemia. The worldwide incidence of HE is 1:2000-4000 individuals, but it results higher in some African regions (1:100). A subtype of HE is HPP, a rare severe hemolytic anemia characterized by poikilocytosis, fragmented erythrocytes, resulting in low MCV (50-60 fL), microspherocytes. The patients show marked splenomegaly, and splenectomy is therefore usually recommended. There is a strong association between HE and HPP, with a third of family members of HPP patients exhibiting typical

HE. Moreover, many HPP patients suffer from severe hemolytic anemia in infancy that gradually improves, evolving toward typical HE during adulthood. The main defect in HE erythrocytes is mechanical weakness or fragility of the erythrocyte membrane skeleton due to defective horizontal connections of cytoskeletal proteins such as dimer-dimer interactions of spectrin and spectrin-actin-protein 4.1 of the junctional complex (Figure 1-7). HE is inherited in an autosomal dominant pattern with rare cases of de novo mutations and the exception of recessive autosomal inheritance of HPP. HE can be due to mutations in *EPB41*, *SPTA1* and *SPTB* genes that lead to serious damage in the association of spectrin dimers (Iolascon et al. 2003). The HE shows high inter- and intra-familial phenotypic variability, due to the modifier alleles. One example is the αLELY (Low Expression Lyon) in *SPTA1* gene, a hypomorphic haplotype composed by two variants, the missense Leu1857Val and the splicing variant in the intron 45. This hypomorphic haplotype alone causes a minimum damage in both heterozygous and homozygous state since the spectrin α chains are produced in excess (3 to 4 fold compared to spectrin β chain); otherwise, when it is associated with a HE mutation in *SPTA1*, the resulting phenotype is severe, i.e. the HPP (Table 1-4).

SAO is a very common condition in the aboriginal peoples from Papua New Guinea, Indonesia, Malaysia, Philippines and southern Thailand, in areas where malaria is endemic, with prevalence varying between 5 and 25%. Indeed, this condition offers protection against all forms of malaria. Despite the reduced *in vitro* deformability of SAO erythrocytes, patients are asymptomatic and the diagnosis is made accidentally as a result of the examination of the peripheral blood smear, showing the characteristic rounded elliptocytes (ovalocytes). However, in newborns it can manifest as hemolytic anemia and may require phototherapy. SAO is inherited as an autosomal dominant trait and is caused by a deletion of 27 nucleotides in *SLC4A1* leading to the loss of the amino acids 400-408 of band 3 protein. The deletion is in linkage disequilibrium with the Memphis polymorphism (p.Lys56Glu) in *SLC4A1* gene. SAO erythrocytes show a small loss of monovalent cations when exposed to low temperatures, with a reduction of anions flux. Thus, SAO could be classified as genetic disease affecting the permeability of RBC membrane. Despite the frequency of heterozygotes, homozygotes are not viable. Indeed, the homozygosity could create an alteration of the ion flux of RBCs and a severe distal renal acidosis due to the loss of band 3 (also expressed in the kidney), that could be lethal (Iolascon et al. 2003).

<u>HAMDs due to altered permeability of RBC membrane</u>

HST includes both DHS and OHS, which show alteration of the permeability of the RBC membrane to monovalent cations Na+ and K+ with consequent alteration of the intracellular cationic content and alterations of cell volume (Delaunay 2007). DHS is the most represented among the HST, with an incidence of approximately 1:50000 births. It is 10-20 times less frequent than HS, with which it can be, however, confused. The phenotype ranges from

asymptomatic forms to severe, with massive hemolysis. Generally, DHS patients show hemolytic well-compensated anemia, with high reticulocyte count and tendency to macrocytosis, mild jaundice. The main characteristic of RBCs is the cell dehydration, because of the loss of the cation content, with a consequent increase of MCHC (>36 g/dL). At blood smear the stomatocytes, erythrocytes with the characteristic central spot mouth-shaped, are quite rare, making diagnosis often difficult. In addition, it may be difficult when the clinical picture is associated with pseudohyperkalemia and/or perinatal edema, in the so-called form pleiotropic syndrome (Delaunay 2004). For these reasons, the condition may be overlooked for years or decades before reaching a conclusive diagnosis. DHS is inherited as an autosomal dominant trait. The candidate gene locus was firstly localized at 16q23-24 (Carella et al. 1998). Subsequently, *PIEZO1* gene was identified as causative of both isolated and syndromic forms of DHS by exome sequencing (Zarychanski et al. 2012; Andolfo et al. 2013). *PIEZO1* encodes a mechanoreceptor, an ion channel activated by pressure (Figure 1-7). This protein has been identified in the RBC membrane and in mice it has been shown to form a tetramer of about 1.2 million daltons; it is therefore the largest ion channel identified to date. Recently, a novel DHS causative gene, *KCNN4*, has been identified in four different DHS2 families (Table 1-3). *KCNN4* gene encodes the Gardos channel, a widely expressed Ca^{2+}-dependent K^+ channel of intermediate conductance that mediates the major K^+ conductance of erythrocytes (Rapetti-Mauss et al. 2015; Andolfo et al. 2015).

OHS is a very rare subtype among HST, overall 20 cases reported worldwide. Contrary to DHS, RBCs are hydrated due to an increase, from 20 to 40 times, in the loss of cations (Stewart 2004). OHS is associated with more severe phenotypes compared to DHS. In addition to reticulocytosis, it is characterized by a sharp increase in MCV (> 110 fL) and decreased MCHC (24-30 g/dl). The number of stomatocytes is usually much higher than that observed in DHS. The causative gene of this condition is *RHAG,* encoding the Rh-associated glycoprotein (RhAG) which acts as an ammonia channel (Genetet et al. 2012) (Figure 1-7). The stomatin protein has been found at low or absent levels in OHS patients, but no mutations have been found in the encoding gene so far.

FP and CHC are additional forms of stomatocytosis. FP is not associated with hemolytic anemia and stomatocytes are only rarely observable at peripheral smear. Conversely, CHC patients show hemolytic anemia of variable degree. RBCs from FP patients exhibit a loss of K^+ at low temperatures (<37°C, mostly 8-10°C), but not at 37°C. In CHC the main feature is the temperature dependence of the loss of cations: instead of being around 8-10°C, the minimum is around 23°C (Delaunay 2007). The gene responsible for FP was mapped at 2q35-q36 (Carella et al. 2004), and then identified in the *ABCB6* gene (Andolfo et al. 2013), encoding the homonymous protein, ABCB6. It belongs to the family of ABC transporters with binding cassette for ATP, one of the most abundant families of integral membrane proteins. ABCB6 is

expressed on the membrane of the erythroid precursors throughout the differentiation and on the mature RBC membrane. Instead, CHC is due to mutations in *SLC4A1* gene; these are gain of function mutations, since they are able to transform the band 3 anion exchanger to a cation transporter (Table 1-3).

1.3 Differential diagnosis of CDAs, HAMDs and related hereditary anemias

Although the workflow to diagnose HHA is a normal clinical practice, differential diagnosis, classification, and patient stratification among CDAs, HAMDs and clinically related anemias are often very difficult. Indeed, the variety of unspecific and overlapping phenotypes often hampers a correct clinical management of the patients.

Thalassemias and hemoglobinopathies are the first conditions to be excluded. The presence of microcytic anemia, pathological Hb electrophoresis and positive familial anamnesis can direct the diagnosis toward these disorders. In particular, the differential diagnosis with thalassemia is important in suspected cases of either CDA IV or CDA variants GATA1-related. Among IBMFS, DBA and FA are disorders that most frequently undergo differential diagnosis with CDAs. Unlike the CDAs, FA generally presents reduction to absent trilinear hematopoiesis, acute myelogenous leukemia or solid tumors; moreover, it can also present developmental abnormalities more frequently compared to CDAs, particularly CDA I. The positivity to the diepoxybutane (DEB) test is a very sensitive and specific tool for guiding FA diagnosis (Chirnomas and Kupfer 2013). Similarly to CDAs, DBA presents as isolated inherited red cell production failure. However, conversely to CDAs, DBA BM exhibits reduced proliferation and survival of erythroid progenitors. Moreover, growth retardation, congenital malformations and increased HbF levels are more frequent features of DBA compared to CDAs. Increased activity of erythrocyte adenosine deaminase is a good II level test for establishing the diagnosis of DBA (sensitivity 84%, specificity 95%, positive and negative predictive values 91%) (Fargo et al. 2013) (Figure 1-8).

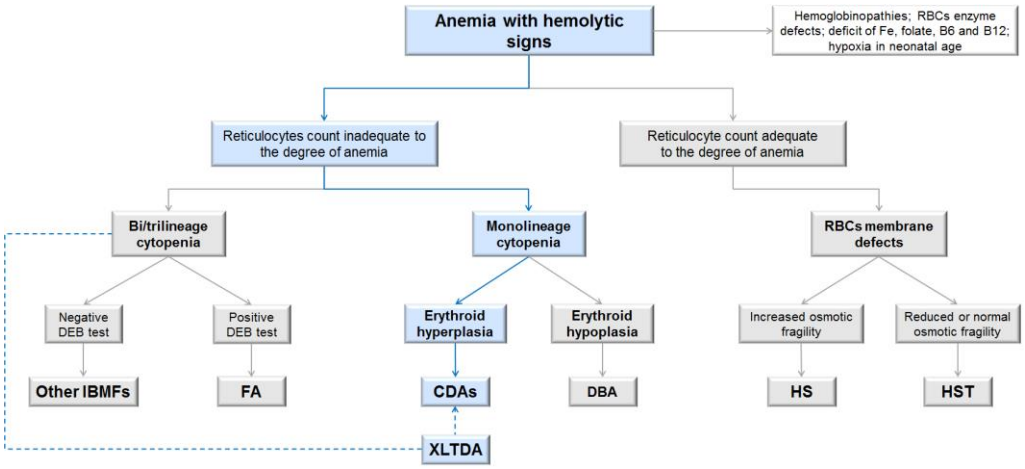

Figure 1-8. Flow diagram for the differential diagnosis of CDAs and related hereditary anemias. The flow diagram shows the main steps (light blue) for guiding the clinical suspicion toward the diagnosis of a CDA. Adapted from Gambale et al. 2016.

CDAs can be also misdiagnosed with hereditary hemolytic anemias. For example, CDA II shares several clinical findings with hemolytic anemias due to red cell membrane defects, such as HS (King et al. 2015). Of note, CDA II patients are often erroneously diagnosed as HS, and consequently, they undergo unnecessary splenectomy. The lack of substantial improvement after intervention leads to a re-examination of the case, allowing the correct diagnosis of CDA II. The most useful pointer to correctly establish the diagnosis of CDA II is the inadequate reticulocyte count for the degree of anemia. Indeed, the marrow stress is higher in CDA II compared to HS for the same Hb level as attested by the increased sTfR levels observed in CDA II patients (Russo et al. 2014). Other parameters could be also used for distinguishing both conditions. For example, the RDW is characteristically increased in CDA II, while the Hb distribution width (HDW) is increased in HS, resulting in an RDW/HDW ratio that is significantly greater in CDA II than in HS (Heimpel and Iolascon 2009). Recently, a new clinical index, named BM responsiveness index, has been developed to discriminate a hemolytic anemia from ineffective erythropoiesis one. This index resulted to be a high sensitive parameter (90.4%) to achieve a clinical diagnosis of CDA II (Russo et al. 2014). In addition, other RBC membrane defects as DHS1-2 or OHS should be evaluated in the differential diagnosis of CDA I. Both CDA I and HST present macrocytosis associated with hemolytic signs. Moreover, it has been described as a novel variant of HST due to a de novo band 3 mutation, transmitted in a dominant fashion, characterized by conversion of band 3 from an anion exchanger to a cation transporter, associated with a dyserythropoietic phenotype (Iolascon et al. 2009).

As effect of chronic anemia and hemolytic component, several complications are associated to CDAs. The main are: iron loading, gallstones,

hyperbilirubinemia, hypersplenism. The variety of unspecific and overlapping phenotypes observed in CDA patients, even in those sharing the same genetic pathogenesis, often hampers a correct clinical management of affected individuals. Beyond achieving a definitive diagnosis, knowing the genetic basis of these patients can be valuable also for guiding treatment.

1.4 Second-generation DNA sequencing

Second-generation sequencing technologies, commonly referred to as NGS, are based on massive parallel sequencing of millions of DNA templates through cycles of enzymatic treatment and image-based data acquisition. The major current application of NGS in diagnostics is through design of disease specific panel, named targeted-NGS (t-NGS), which targets a group of selected genes. T-NGS is faster and cheaper than other NGS technologies, such as exome sequencing, especially for the analysis of certain distinct disease phenotypes or for the differential diagnosis of overlapping clinical conditions.

For the past 25 years, dideoxy DNA Sanger sequencing and fluorescence based electrophoresis technologies, known as the first-generation technology, were the gold standard for somatic and germline genetic studies, as well as for diagnosis of genetic diseases. However, these methods are expensive and have low throughput owing to implementation on single amplicons. Because of the limited adaptability of automated Sanger sequencing, there was a need for new and improved technologies for sequencing large numbers of human genomes at the same time by using various strategies that rely on a combination of template preparation, sequencing and imaging, genome alignment, and assembly methods. After almost three decades, NGS was developed. This technology of high-throughput sequencing has provided a large impetus for *de novo* sequencing, resequencing, exome sequencing, transcriptome profiling, methylation profiling, and metagenomics studies (Desai and Jere 2012). NGS is a very versatile technology, applicable to various questions either in basic research or in clinical research. Indeed, the broadest application of NGS is the resequencing of human genomes to enhance the understanding of how genetic differences affect health and disease. For this purpose, NGS has reached an adequate level of evolution as much to be considered as robust enough for clinical applications, mainly for discovering new genes underlying monogenic and/or multifactorial diseases. Indeed, over the past decade, NGS has led to an exponential increase in our understanding of the genetic basis of Mendelian diseases. More recently, NGS has been successfully deployed in the clinics, with a reported diagnostic yield of ~25 % (Jamuar and Tan 2015).

Sequencing technologies include a number of methods that are gathered broadly as template preparation, sequencing and imaging, and data analysis. The unique combination of specific protocols distinguishes one technology from another and determines the type of data produced from each platform. The flowcharts of library preparation for NGS sequencing are quite similar:

1. Templates preparation: this step requires a fragmentation of genomic DNA into smaller size and the creation the DNA libraries, followed by *in vitro* ligation of templates to common adaptor sequences on a solid surface or support.

2. Clonally amplified templates: the generation of clonally clustered amplicons can be achieved by several approaches. After the successful amplification and enrichment, beads, representing millions of PCR, can be immobilized with various strategies in a polyacrylamide gel on a standard microscope slide or solid surface.

3. Sequencing and imaging: the sequencing reaction consists of alternating cycles of enzyme-driven biochemistry and imaging-based data acquisitions. The methods for sequencing provide several protocols, such as cyclic reversible termination, sequencing by ligation, pyrosequencing, and real-time sequencing

4. Genome alignment and assembly: after generation of NGS reads, the alignment to a known reference sequence or assembled de novo is performed (Shendure and Ji 2008).

The major advance offered by NGS is the ability to produce an enormous volume of data cheaply. While automated Sanger sequencing produced a maximum of approximately six Megabase (Mb) of DNA sequences per day (low throughput), the amount of NGS works around 450-50000 Mb of DNA sequences per day, based on platform used. The lower cost per base of NGS and the ability to sequence millions of reads in parallel, allowing for simultaneous analysis of a large number of genes, give to this technology other benefits. In the end, with NGS technologies the amount of DNA to be sequenced is no longer a barrier to launching a new or expanded clinical test. The advantages of second-generation DNA sequencing are currently offset by several disadvantages. First, the most prominent of these include read-length: for all of the new platforms, read-lengths are currently much shorter than conventional sequencing (35-400 nucleotides compared to 600-1000 in length of automated Sanger sequencing). Second, raw accuracy, on average, of base-calls generated by the new platforms are at least tenfold less accurate than base-calls generated by Sanger sequencing (1/100-10000 nucleotides compared to 1/10000-100000 of conventional sequencing). Despite their weaknesses, improvements in these technologies promise to meet the technical requirements and strict quality standards for clinical diagnostic practises including analytical sensitivity, reproducibility and cost effectiveness. Next-generation sequencing includes several high-throughput sequencing approaches:

- Targeted sequencing and gene panel sequencing involve sequencing of selected parts of the genome and/or of selected known/candidate disease-causing genes. This approach is especially suitable for the diagnosis of genetically heterogeneous disorders.

- Whole-exome sequencing (WES) targets sequencing of the exome, which constitutes 1-2% of the human genome. The application of whole-exome sequencing in diagnostics and research has enabled the discovery of novel disease genes.

- Whole-genome sequencing is the sequencing of the entire genome, including non-coding, regulatory DNA regions.

- RNA sequencing, also called whole transcriptome shotgun sequencing, uses NGS for both mapping and quantifying RNA in a biological sample

1.5 Targeted-next generation sequencing

The major current application of NGS in diagnostics is through disease-targeted tests for which multiple causal genes are known. The combination of high-throughput and relatively small DNA target selection allows for many genes and samples to be processed simultaneously, making it a reliable solution for the processing of large sample numbers in a diagnostic laboratory. Several t-NGS panels have been already launched and actually used in diagnostic practice for different disorders (Table 1-5).

Table 1-5. Examples of targeted available designs for clinical practice

Disease area	Disease type	Number of genes
Cancer	Hereditary cancers (breast, colon, ovarian)	10–50
	Cardiomyopathies	50–70
Cardiac disorders	Arrhythmias (long QT syndrome)	10–30
	Aortopathies (Marfan's syndrome)	10
Immune disorders	Severe combined immunodeficiency syndrome	18
	Periodic fever	7
Neurological, neuromuscular and metabolic disorders	Ataxia	40
	Congenital disorders of glycosylation	23-28
	Dementia (Parkinson's and Alzheimer's diseases)	32
	Developmental delay, autism, intellectual disability	30-150
	Epilepsy	53-130
	Hereditary neuropathy	34
	Microcephaly	11
	Mitochondrial disorders	37-450
	Muscular dystrophy	12-45
Sensory disorders	Eye disease (retinitis pigmentosa)	66–140
	Hearing loss and related syndromes	23-72
Others	Rasopathies (Noonan's syndrome)	10
	Pulmonary disorders (cystic fibrosis)	12-40
	Short stature	12

The available disease-targeted panels comprise from ten to hundred genes. Some panels contain many genes that can give rise to indistinguishable presentations, while other panels have been developed with the aim to diagnose diseases with overlapping phenotypic presentations.

The general principle of t-NGS provides for specific probes carefully designed to target regions of interest relevant to genetically heterogeneous disease

phenotypes, for which essential is the differential diagnosis. T-NGS is faster and cheaper than other NGS technologies, leads to a higher sequencing coverage, and, therefore, highly accurate DNA variant calling for the region of interest. Consequently, considerable effort has been dedicated to develop 'target-enrichment' methods, in which genomic regions are selectively captured from a DNA sample before sequencing. The development of targeted technology offers various methodological enrichment approaches to construct and enrich a DNA library and use both PCR and/or hybridization as mode of target selection. First, the classic PCR approach to library generation requires amplification of target regions, pooling, concatenation, shearing and ligation of adaptors and sequencing primers. Secondary one, droplet based microfluidic technologies have evolved to facilitate high-throughput PCR in picoliter droplets. With hybridization-based methods, libraries are constructed by shearing total gDNA followed by adaptor ligation and hybridization to oligonucleotides that are complementary to the desired target. Hybridization can be performed either on a solid surface array, on a filter, or by hybridization in solution. A third general approach to target selection uses molecular inversion probes, which consist of two primers linked together by a backbone and, similar to PCR, bind to specific target DNA. Gap filling, ligation, and enrichment steps follow this (Mamanova et al. 2010). There are several parameters by which the performance of each enrichment can be monitored, which vary from one approach to another:

- sensitivity, or the percentage of the target bases that are represented by one or more sequence reads;

- specificity, or the percentage of sequences that map to the selected targets;

- uniformity, or the variability in sequence coverage across target regions;

- reproducibility;

- cost and ease of use;

- amount of DNA required per experiment, or per Mb of target.

Finally, different types of platforms, known as "bench top" sequencers, have also been developed specifically for high-throughput sequencing of targeted technologies: Ion Torrent PGM (Life Technologies Ltd, Paisley, UK), 454 GS Roche Junior (Roche Applied Science, Indianapolis, IN), and the Illumina MiSeq (Illumina, San Diego, CA).

2. Aims of the study

The variety of unspecific and overlapping phenotypes observed in HHA patients, even in those sharing the same genetic pathogenesis, often hampers a correct clinical management of affected individuals. The primary aim of our study was the development of a fast, accurate, reliable and cost effective diagnostic/prognostic tool for HHA based on t-NGS. Indeed, HHA refers to a highly heterogeneous group of inherited rare anemias in which differential diagnosis, classification, and patient stratification are often very difficult. Beyond achieving a definitive diagnosis, knowing the genetic basis of these patients can be valuable also for guiding treatment. In order to evaluate the reliability of this approach, we firstly performed a t-NGS pilot study including 10 known causative genes of two HHA sub-groups, CDAs and HS. Subsequently, we created a second t-NGS gene panel, named RedPlex, composed by 34 *loci* causative or candidates of CDAs, HS and HST.

The t-NGS approach also allowed the identification of "polygenic" conditions, i.e. patients in which the phenotypic variability could be explained by the presence of modifier variants associated to causative mutations. Thus, the secondary aim of this project was the study of the interaction between mutated genes in HHA patients. We particularly focused on the regulatory network GATA1-mediated on *SEC23B* gene in CDAII patients with peculiar clinical phenotypes or those with incomplete mutation pattern in SEC23B.

3. Materials and Methods

3.1 Patients and genomic DNA preparation

Diagnosis was based on history, clinical findings, laboratory data, morphological analysis of peripheral blood and/or aspirated bone marrow. For CDAII patients the hypoglycosylation of band 3 was evaluated by SDS-PAGE, whenever possible. For HS/HST patients, diagnosis was also based on indirect tests (OF, AGLT50, EMA binding) as well as on ektacytometry. Whenever possible, RBC membrane proteins were analysed by SDS-PAGE. We enrolled a first set of 13 HHA patients from 12 unrelated families for the pilot study, six out them with a well-defined phenotype and genotype annotation. Particularly, molecular diagnosis in two HS and four CDA patients was already established. The second set was composed of 32 HHA patients from 27 unrelated families. Samples were obtained after informed consent for the studies, according to the Declaration of Helsinki. Relatives were also enrolled. Local university ethical committees approved collection of patient's data from Medical Genetics Ambulatory in Naples (University Federico II, DAIMedLab).

Genomic DNA preparation, mutational search, oligonucleotide primers design and direct sequencing were performed as previously described (Russo et al. 2010). Sequence primers are available on request (russor@ceinge.unina.it). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of ATG translation initiation codon in the reference sequence (Ensembl transcript ID), according to the nomenclature for the description of sequence variants of Human Genome Variation Society (www.hgvs.org/mutnomen). The initiation codon is codon 1. In order to evaluate the quality of the extracted gDNA before the fragmentation, samples were quantified by NanoDrop 2000 (Thermo Scientific, Italy). Then, gDNA was loaded on 0.8% DNA agarose gel electrophoresis.

3.2 Custom target enrichment

We created two t-NGS gene panels: (1) the first, the pilot study design, including only 10 genes causative of HS and CDAs; (2) the second, named RedPlex, composed by 34 *loci* causative or candidates of HHA.

Pilot study and RedPlex designs

For the first gene panel 10 causative genes of both CDAs and HS were selected (Table 3-1). For the probe design, coding regions, 5'UTR, 3'UTR, 100 bp flanking splice junctions were selected as regions of interest (ROI). The sequences corresponding to the genomic regions of ROI were uploaded to the web-based tool SureDesign (https://earray.chem.agilent.com/suredesign.htm, Agilent Technologies, USA). Sequence length was set at 100×2 nucleotides, and the predicted target size amounted to 236 target regions (87594 bp). For

the second gene panel 34 causative/candidate genes of CDAs, HS and HST were included (Table 3-1). Similarly to the first panel, all coding regions, 5' and 3'UTRs, and 100 bp flanking splice junctions were included as ROI in the electronic design. Sequence length was set at 150×2 nucleotides, and the predicted target size amounted to 538 regions (239.764 kb). The coordinates of all the sequence data were determined using NCBI build 37 (GRCh37/hg19).

Table 3-1. Target genes in HaloPlex Design

Target ID	ENST - GRCh37	Regions	Coverage (%)	High Coverage[§] (>= 90%)	Low Coverage[§] (< 90%)
ABCB6	ENST00000265316	9	100	9	0
ABCG5	ENST00000260645	11	100	11	0
ABCG8	ENST00000272286	10	100	10	0
ADD2	ENST00000264436	16	100	16	0
ANK1	ENST00000347528	41	100	41	0
ATP11C	ENST00000327569	30	100	30	0
ATP2A2	ENST00000308664	18	99.88	18	0
C15ORF41	ENST00000566621	12	99.81	12	0
CDAN1	ENST00000356231	17	99.4	17	0
EPB41	ENST00000373800	19	100	19	0
EPB42	ENST00000300215	13	100	13	0
EPB49	ENST00000265800	11	100	11	0
FOG1	ENST00000319555	9	99.89	9	0
FOXL1	ENST00000320241	1	100	1	0
GATA1	ENST00000376670	5	100	5	0
GATA2	ENST00000487848	6	100	6	0
GFI1B	ENST00000339463	11	100	11	0
KIF23	ENST00000260363	19	100	19	0
KLF1	ENST00000264834	3	100	3	0
MPP1	ENST00000369534	1	100	1	0
PIEZO1	ENST00000301015	33	100	33	0
RHAG	ENST00000371175	10	100	10	0
SAR1A	ENST00000373241	8	99.8	8	0
SEC23B	ENST00000377475	20	100	20	0
SEC31A	ENST00000395310	26	100	26	0
SLC12A4	ENST00000316341	15	100	15	0
SLC12A6	ENST00000354181	25	99.97	25	0
SLC12A7	ENST00000264930	23	99.97	23	0
SLC2A1	ENST00000426263	7	99.3	7	0
SLC4A1	ENST00000262418	14	99.86	14	0
SPTA1	ENST00000368147	48	99.95	48	0
SPTB	ENST00000389721	31	99.84	31	0
STOM	ENST00000286713	7	100	7	0
TMOD1	ENST00000259365	10	100	10	0

[§]Number of regions with High or Low coverage
In bold are highlighted the genes included in the first panel (pilot study design)

Sample preparation, capturing and enrichment

Sample preparation was performed following the instruction's manufacturer for HaloPlex Target Enrichment kit for Illumina Sequencing - Custom Design from 1-500kb, p/n G9901A (protocol version D - Agilent Technologies). All procedures require the use of a control gDNA provided to kit (Enrichment Control DNA, ECD). Briefly, each DNA sample (225 ng), included the ECD DNA, is fragmented in two specific double-digestion reactions defining the ROI at 37°C for 30 minutes. Validation of the digestion is performed on the ECD reactions using 2100 Bioanalyzer system analysis with 2100 Expert Software (version B.02.07, Agilent Technologies) and High Sensitivity DNA Assay Kit (p/n 5067-4626, Agilent Technologies). Because of amplicon redundancy, target region is covered many times. Then, target fragments are mixed with a primer cassette and custom HaloPlex biotinylated probes, which are specific for target regions and contain Illumina adapters and index sequences. During the hybridization (at 54°C for 3 hours), the probes circularize the targeted DNA fragments. Correctly hybridized fragments form a nicked double strand DNA structure with the HaloPlex probe. Thus, circularized target DNA-HaloPlex probe hybrids are captured by streptavidin beads (HaloPlex Magnetic Beads, Agilent Technologies) and then washed. Subsequently, circularized HaloPlex probe-target DNA hybrids are ligated with DNA ligase to close nicks between the target fragment and the primer cassette, at 55°C for 10 minutes. Then, the circularized hybrid is amplified by PCR with Herculase II Fusion Enzyme (Agilent Technologies), using an exclusive primer composed by two regions, one complementary to the target and the other useful for sequencing. Finally, each amplicon contains one target insert surrounded by the Illumina paired-end sequencing elements, the sample index and the library bridge PCR primers. The amplified target DNA is purified using Agencourt AMPure XP beads (Beckman Coulter Genomics, USA) and 10 mM Tris-HCl buffer (pH 8.0). After target enrichment, validation and quantification of PCR library are performed by 2100 Bioanalyzer system analysis using High Sensitivity DNA Assay Kit (Agilent Technologies) before sequencing. The amplicon size ranges between 175 and 625 bp, with the majority of products between 225-525 bp.

3.3 Sequencing and data analysis

High-throughput sequencing was performed by Illumina Hiseq2000 platform for the pilot gene panel and Illumina NextSeq 500 for the RedPlex gene panel. Agilent SureCall software (v 3.0.3.1, Agilent Technologies) was used for bioinformatic and computational analyses. Analysis in SureCall starts from raw reads; after removal of the adaptor sequences, the reads are aligned to GRCh37/hg19 human reference sequence with Burrows - Wheeler aligner (BWA) included in SureCall software. SAMTools are used to recalibrate the base call quality scores, perform local realignment, and index the reads for

improved performance. SAMTools are also used to identify mutations and to assess the significance of the mutations. Each of these was evaluated by SureCall software based on its location, amino acid change, effect on protein function (SIFT), and impact on structure and function of the protein (PolyPhen-2). Moreover, mutations information were aggregated from various public sources, including NCBI, COSMIC (Catalog of Somatic Mutations in Cancer), PubMed, and Locus Specific Databases (HGVS). The significance of the mutations were evaluated by the mutational classifier following customized guidelines. The custom categorization was based on the criteria shown in the Table 3-2. The SureCall output for each patient is a genome viewer and a mutation table, from which originates a report file.

Table 3-2. SureCall custom categorization

Category I *Coding variants*	Any missense mutation
	Nonsense or a frame shift mutation
	Introduction of a stop codon
	Results in codon change
	Results in codon insertion
	Results in codon change and codon deletion
	Results in codon change and codon insertion
	Results in codon deletion
	Missense mutation of the normal stop codon
	Mutates in the initiation codon (ATG)
	Deletes nucleotide(s) that lead(s) to a shift of reading frame
	Deletes exon which results in shift of reading frame
	Is non-synonymous coding variant in start
	Is non-synonymous coding variant in stop
	Is non-synonymous coding variant
	Is synonymous coding variant in stop
	Is synonymous coding variant in start
	Is synonymous coding variant
	In-frame amino acid insertion/deletion
	In-frame exon deletion
Category II *Splicing variants*	Located within a splice consensus sequence
	Is splice site donor
	Is splice site acceptor
	Alters the sequence at a splice junction
	Likely to produce a cryptic splice site
Category III *Regulatory variants*	Modifies UTR 3'
	Modifies UTR 5'
	Deletes UTR 3'
	Deletes UTR 5'
	Likely to affect transcription
Category IV *Intronic variants*	Sequence changes that occur in the intron
	Is inter-genic
	Is intronic variant

The mutations were divided in four categories, each one composed by several sub-types. The first category includes coding variants, such as non-synonymous coding variants, nonsense and/or frame shift mutations. The second category includes splicing variant mutations located in consensus, donor or acceptor sites, and mutations that generate a novel cryptic splice site. The third category comprises the regulatory variants, modifying and/or deleting both 3' and 5' UTRs, and the mutations affecting the transcription levels. Finally, the last category embraces the intronic variants.

Annotated variants were firstly filtered on the basis of their presence in variation databases. Indeed, variants with a minor allele frequency (MAF) > 0.01, as reported by NCBI-dbSNP (http://www.ncbi.nlm.nih.gov/SNP/), Exome Variant Server (EVS, http://evs.gs.washington.edu/EVS/), and 1000 Genomes (http://browser.1000genomes.org/index.html), were excluded. Rare and low frequency variants (MAF < 0.01 and 0.05, respectively) were selected. Subsequently, we applied a strand bias filtering and we choose a strand bias threshold of 0.90. Thus variants with a strand bias > 0.90 were also excluded. Finally, the remaining variants were prioritized by using the scores provided by the prediction tools PolyPhen2 and SIFT for coding variants. Prediction analysis for splice sites mutations was also performed by web server tool Human Splicing Finder (http://www.umd.be/HSF/), as previously described (Russo et al. 2013b).

All focused variants were confirmed by Sanger sequencing and by the analysis of inheritance pattern. The validations were performed using 50 ng of genomic DNA. Custom primers were designed by Primer3 program (Primer3 v. 0.4.0, freeware online). Direct sequencing was performed using the Dye Terminator Cycle Sequencing FS Ready Reaction Kit (Applied Biosystems) and a 373A DNA sequencer (Applied Biosystems).

3.4 Gene and protein expression analysis

RNA isolation and reverse transcription

Total RNA was extracted from peripheral blood mononuclear cells (PBMCs) and cell lines using Trizol reagent (Life Technologies). Synthesis of cDNA from total RNA (2 μg) was performed using cDNA synthesis kit (Applied Biosystems, Milan, Italy).

Quantitative real-time PCR analysis

Quantitative RT-PCR (qRT-PCR) using Power SYBR Green PCR Master Mix (Applied Biosystems) was performed to evaluate the gene expression of *GATA1* and *SEC23B* genes. Samples were amplified on Applied Biosystems 7900HT Sequence Detection System using standard cycling conditions. The primers were designed by the Primer Express 2.1 program (Applied Biosystems). β-actin was used as internal control. Relative gene expression was calculated by using the $2^{-\Delta Ct}$ method, while the mean fold change = 2^{-}

$^{(\text{average } \Delta\Delta Ct)}$ was assessed using the mean difference in the ΔCt between the gene and the internal control (Russo et al. 2013b). Primer sequences are available upon request (russor@ceinge.unina.it).

Protein isolation and western blotting analysis

Protein extraction from PBMCs and western blotting (WB) were performed as previously described (Andolfo et al. 2010; Russo et al. 2013b). Particularly, 30 µg of total extract proteins was loaded into each lane and was separated by 10% SDS PAGE bisacrylamide gel, followed by transfer to PVDF membranes (Biorad, Milan, Italy). A specific rabbit anti-SEC23B antibody (1:500) (BioLegend, San Diego, CA) and anti-GATA1 antibody (1:500) (ab11963 - Abcam, Cambridge, UK) were used. Mouse anti-β-actin antibody (1:5000) (Sigma-Aldrich, Milan, Italy) was used as the control for equal loading. Semi-quantitative analysis of protein expression was performed as previously described (Persico et al. 2008). The bands were quantified by densitometry to obtain an integral optical density value, which then was normalized with respect to the β-actin value.

3.5 Vector cloning

SEC23B promoter characterization

The in silico analysis of human *SEC23B* (HuSEC23B) upstream region (chr20:18487188-18491479: 4292 bp from ATG) has been performed by ENCODE web tool implemented in the UCSC Genome Browser (http://genome.ucsc.edu/). Prediction analyses for CpG island were performed by CpG Islands Track (UCSC Genome Browser) and by EMBOSS Cpgplot, at EMBL-EBI website (http://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/). The hypothetical promoter region of SEC23B gene (3450 bp from ATG) has been divided into 10 overlapping fragments (HuSEC23B/3.44, HuSEC23B/2.52, HuSEC23B/2.14, HuSEC23B/1.75, HuSEC23B/1.28, HuSEC23B/0.99, HuSEC23B/0.81, HuSEC23B/0.61, HuSEC23B/0.45, HuSEC23B/0.21), amplified from genomic DNA (figure 3). Each construct was purified with QIAquick PCR Purification Kit (Qiagen, Milan, Italy), cloned upstream the luciferase gene into PGL3 vector (Life Technologies) in the Hind III and Xho I sites. Two-primer site-direct PCR mutagenesis was used to mutagenize the GATA1 binding sites (GATA1bs) in the two fragments HuSEC23B-2475 and HuSEC23B-463. The amplified region and ligation boundaries were verified by direct sequencing, as previously described.

Cloning of GATA1 cDNA product into expression vector

GATA 1 cDNA product (BC009797_Clone ID 4048082, Invitrogen) was amplified from pDNR-LIB vector, purified with QIAquick PCR Purification Kit (Qiagen) and cloned into pcDNA3.1 vector (Invitrogen, Milan, Italy) in the EcoRV and XhoI sites. All plasmids were transformed into the DH5α bacterial strain (Invitrogen) and subsequently purified with the QIAprep Miniprep Kit

(Qiagen). Direct sequencing was performed as previously described. Two-primer site-direct PCR mutagenesis was used to introduce GATA1-associated missense mutations G208R, R216W, D218G, and V205M. The amplified region and ligation boundaries were verified by direct sequencing.

3.6 Cell cultures and transfection

HEK-293 and K562 cells (ATCC, Manassas, VA, USA) were maintained according to the manufacturer's instructions. HuSEC23B plasmids were transfected (2 µg/well) using the X-tremeGENE HP DNA Transfection Reagent (Roche, Milan, Italy) according to the manufacturer's procedures. The cells were collected 48h after the transfection to luciferase assays. Co-transfection of GATA1bs/HuSEC23B-2475 and -463 mutants (2.5 µg/well) and GATA1 WT (2.5 µg/well) was performed using the X-tremeGENE HP DNA Transfection Reagent (Roche) according to the manufacturer's procedures. The cells were collected 48h after the transfection to luciferase assays. Similarly, co-transfection of GATA1 mutants (GATA1/G208R, GATA1/R216W, GATA1/D218G, and GATA1/V205M) (2.5 µg/well) and the HuSEC23B/3.44 (2.5 µg/well) in HEK-293 cell line was performed using the X-tremeGENE HP DNA Transfection Reagent (Roche) according to the manufacturer's procedures. The cells were collected 48h after the transfection to luciferase assays.

Erythroid differentiation of K562 and Hel cells

Fifty µM hemin (Sigma-Aldrich) was added to the culture medium of the wild-type K562 and Hel cells ($2x10^5$/mL). Samples were collected at days 6 after hemin addition. Erythroid differentiation was assessed by FACS analysis for transferrin receptor 1 (CD71) and glycophorin A (CD235A) as previously described (Andolfo et al. 2010).

3.7 Chromatin immunoprecipitation assay

The hypothetical GATA1 binding sites in the promoter region of the SEC23B gene has been predicted by the software tool MatInspector, implemented in Genomatix web server (http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html) and by PROMO at the ALGGEN server (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). Chromatin immunoprecipitation (ChIP) studies were performed by Immunoprecipitation Kit - Dynabeads, following the manufacturer's instruction (Life Technologies). Isolated chromatin from K562 and HEL cells at 6 days of erythroid differentiation was immunoprecipitated using control immunoglobulin G or 5 µg rabbit polyclonal GATA1 antibody (ab11963 - Abcam). Ten µl of immunoprecipitated DNA were used for the qRT-PCRs, to detect the presence of specific DNA segments. We used ETO (Ajore R et al. 2010) as a positive

control and a not-binder GATA1-sequence of the HuSEC23B promoter as a negative control.

3.8 Promoter assay

Luciferase assays with HuSEC23B deletion mutant constructs were performed in the HEK-293 and K562 cell lines. Cells were grown according to the manufacturer's protocol. For the promoter assay, HEK-293 and K562 cells were plated in 6-well plates and transfected with the HuSEC23B deletion mutants using Extreme gene (Roche). Transfected cells were harvested after 48 hrs. The PRL-CMV vector (100 ng) was used for normalization. Luciferase activities were analysed using a Dual Luciferase Reporter Assay system (Promega).

4. Results

4.1 Pilot study

Thirteen HHA patients from 12 unrelated families were included in the pilot study. Overall clinical features and diagnostic suspicion for each subjects was summarized in Table 4-1.

Table 4-1. Clinical data of 15 HHA patients included in the pilot study

Patient ID	Gender (M)ale/(F)emale	Ethnicity	Consanguinity	Diagnostic suspicion
HHA1	F	France	No	CHC
HHA2	F	United Kingdom	No	HPP
HHA3	M	Italy	No	HS
HHA4	F	Italy	No	HS
HHA5	M	Egypt	No	HS
HHA6	M	Italy	Yes[1]	HS
HHA7	M	France	No	HS
HHA8	M	Italy	No	CDA II
HHA9	F	Italy	No	CDA II
HHA10	M	Turkey	Yes[2]	CDA I
HHA11	F	Czech Republic	No	CDA IV
HHA12	M	Italy	No	Healthy subject
HHA13	F	Italy	No	HS
HHA14	M	Italy	-	CDA II/HS
HHA15	F	Italy	-	Healthy subject

[1]Parents are distant relatives; [2]Parents are first cousins
In bold are highlighted the patients with known molecular diagnosis
In italics are indicated subjects from the same family

Among 13 probands, six showed a clinical suspect of HS. Among these, HHA14 was firstly diagnosed as CDA II because of increased value of sTfR (11.10 mg/L) and the presence of ineffective erythropoiesis (Hb: 4.8 g/dL with inadequate reticulocyte absolute count, 114500/µL). Because of uncertain hematologic data and in presence of a positive family history (HHA14's mother, HHA13, was suspected as HS), HS indirect tests were performed. OF, AGLT50 and EMA binding tests resulted suggestive of HS in HHA14 and in HHA13, except for EMA binding that was normal. Among the remaining seven patients, five were already diagnosed at both clinical and molecular level. Particularly, HHA1 and HHA2 were diagnosed as CHC and HPP, respectively, while HHA9, HHA10, and HHA11 were diagnosed as CDAs. HHA9 and HHA10 were firstly misdiagnosed as HS, and they underwent

splenectomy without any improvements. Moreover, BM examination in HHA9 highlighted erythroid hyperplasia with >10% of mature binuclear erythroblasts, while SDS-PAGE analysis showed the hypoglycosylation of band 3. On the other side, HHA10 exhibited internuclear chromatin bridges between erythroblasts at BM light microscopy examination. Almost all CDAs patients received few transfusions during their lifespan, but HHA11 resulted transfusion-dependent.

Based on HaloPlex Sure Design, the design for the pilot study on 10 genes encompassed 211 target regions within 87.383 Kb. After the submission of the design, the resulting amplicons were 5131 with total sequenceable size regions of 148.94 Kb. The mean sequencing coverage of target regions was 99.34% ± 0.48%. The mean percentage of all reads in covered regions was 93.42% ± 3.21%. The mean percentages of analyzable target region bases with at least 20 and 100 reads were 99.79% ± 0.09% and 99.40% ± 0.45%, respectively. The insufficiently covered regions were found mostly outside coding regions. For each patient target gene variants were in the range 55-105; among these 5-13 were coding variants (Table 4-2).

Table 4-2. Total and selected variants

Category	No. of variants[†]
Selected variants	1-5
Focused variants	
Coding mutations related to clinical phenotype	0−2
Coding mutations modifying clinical phenotype	0-3
Coding mutations unrelated to the clinical phenotype	0−3
Intronic and regulatory gene variants	0-2
Total variants[§]	
Total gene variants	62−122
Off-target gene variants	0−2
Target gene variants	55-105
Intronic and regulatory gene variants	48-92
Coding gene variants	5-13

[†]Number of variants refers to the range observed (from lowest to highest number of variants) in the 15 cases with MAF<0.01 or without MAF
[§]Number of all variants refers to the range observed (from lowest to highest number of variants) in the 15 cases without any filters

Beyond to establish a definitive diagnosis for HHA patients, this approach highlighted the presence highlighted the presence of phenotype modifier mutations in those cases with overlapping phenotypes (Figure 4-1).

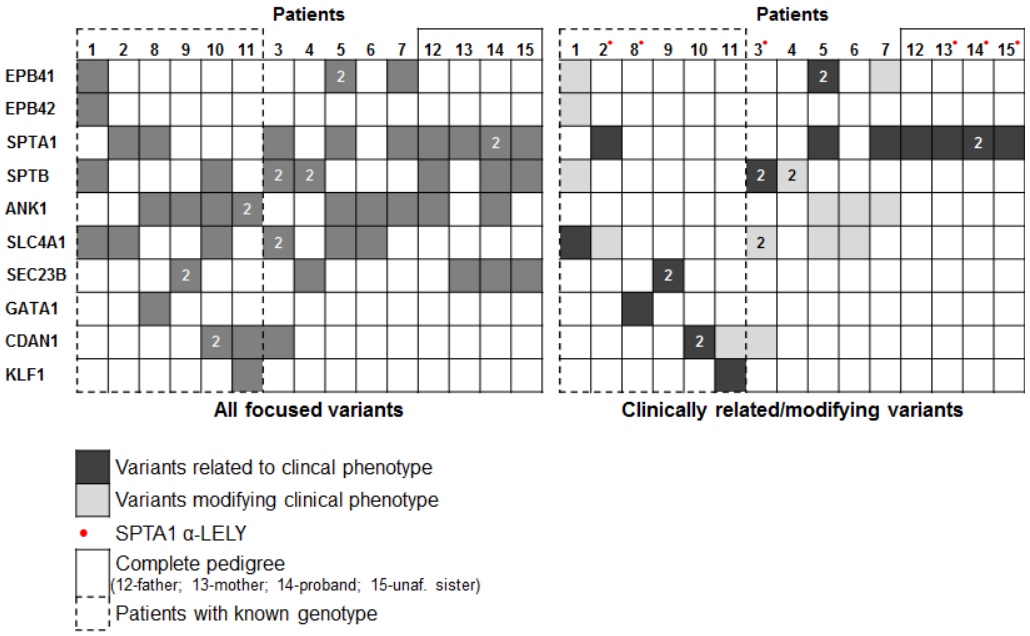

Figure 4-1. T-NGS data from pilot study on 13 patients with HHA. The charts on the left and on the right represent in the column the patients analyzed by t-NGS and in row the genes analyzed. On the left are indicated all the focused variants (variants filtered on the basis of MAF and strand bias). The grey square indicates the presence of at least one variant in the gene, while the number 2 indicates the presence of two variants in the same gene. On the right are shown the clinically related/modifier variants selected among the focused one on the basis of clinical phenotype and predictions by bioinformatic tools.

HHA1, HHA2, HHA9, HHA10 and HHA11 were previously diagnosed at molecular level. HaloPlex approach was able to confirm the molecular diagnosis in all of these patients. Particularly, HHA1 showed the variant p.Gly796Arg in *SLC4A1* gene in heterozygous state, accordingly to the clinical suspect of CHC. HHA2 patient revealed the presence of αLELY polymorphisms (rs3737515 and rs28525570 in compound heterozygous state) in *SPTA1*; however, a previously described large deletion spanning from two exons in *SPTA1* gene was not identified (Iolascon et al. 2011). In HHA9 we confirmed the presence of two causative *SEC23B* mutations, p.Arg14Trp (rs121918222) and p.Glu109Lys (rs121918221). Accordingly to the clinical suspicion of CDA I, we found two mutations in *CDAN1* gene in HHA10 patient: the missense p.Pro86Ser in heterozygous state and a new intronic variant c.774-89G>T, which is predicted to alter the splice site of the nearby exon. Of note, in this case HaloPlex approach was able to identify a non-coding variant far from splice site. Finally, HHA11 patient presented one missense mutation in *KLF1* p.Glu325Lys (rs267607201), in agreement with the clinical suspicion (Table 4-1).

Among the remaining seven patients with unknown diagnosis, we were able to define the causative genotype in five out them. Two patients, HHA4 and

HHA6, resulted undiagnosed. Two probable causative mutations in *SPTB* were identified in HHA3 patient, who had a clinical suspect of HE: p.Arg1737Trp (rs149727354) and a novel missense variation, p.Arg52Gln, both in heterozygous state. In addition, we reported the αLELY variant in homozygous state. In the absence of the parental DNA, we did not define the inheritance pattern of these variants. HHA4 patient presented only a synonymous variant p.Ser603= (rs139882548) in *SEC23B* gene, which resulted unrelated to the phenotype. Accordingly to the dominant transmission in both HHA5 and HHA7 patients, we found a causative mutation, p.Arg568Pro, and a missense annotated variant, p.Glu11Asp (rs41273533), in *SPTA1* that could explain the HS phenotype in both cases. Finally, in the patient HHA8 we confirmed the diagnostic suspicion of CDA II/XLTDA; indeed, a missense mutation p.Gly208Arg in *GATA1* gene was identified, as explained in the next sections of the results (Figure 4-1).

4.2 RedPlex study

Thirty-two HHA patients from 27 unrelated families were included in the RedPlex study. Overall clinical features and diagnostic suspicion of HHA patients included in this study was summarized in Table 4-3.

Table 4-3. Clinical data of 44 HHA patients included in the RedPlex

Patient ID	Family ID	Gender (M)ale/(F)emale	Ethnicity	Consanguinity	Diagnostic suspicion
RED1	F1	F	Sweden	Yes[2]	HST
RED4	F2	M	Italy	No	HST/CDAs
RED5	F2	F	Italy	No	HST/CDAs
RED6	F3	M	Oman	Yes[1]	HS
RED9	F3	F	Oman	Yes[1]	HS
RED10	F4	F	Italy	No	OHS
RED11	F5	F	Italy	No	HS/Undefined
RED12	F6	M	Italy	No	HS/Undefined
RED14	F7	M	Italy	No	Undefined
RED17	F8	F	Italy	No	HS
RED18	F8	M	Italy	No	HS
RED20	F9	F	Italy	No	HST/CDAs
RED21	F9	M	Italy	No	HST/CDAs
RED23	F10	M	Italy	No	CDA I
RED26	F11	M	Australia	No	CDAs
RED29	F12	M	Turkey	Yes[2]	CDA I
RED30	F13	M	United Kingdom	No	CDA II

RED33	F14	M	USA	Yes[1]	CDA II
RED34	F15	M	Turkey	Yes[2]	CDA II
RED35	F16	M	Turkey	Yes[2]	CDA II
RED36	F17	F	Turkey	No	CDA II
RED37	F18	M	Turkey	Yes[2]	CDA II
RED38	F19	M	Turkey	No	CDA II
RED39	F20	M	Turkey	Yes[1]	CDA II
RED40	F21	F	Turkey	Yes[2]	CDA I
RED41	F22	F	France	No	DHS
RED42	F23	F	France	No	DHS
RED43	F24	F	Italy	No	FP
RED45	F25	F	Germany	No	DHS
RED46	F26	M	Italy	No	DHS
RED47	F26	F	Italy	No	DHS
RED48	F27	M	United Kingdom	No	DHS

[1]Parents are distant relatives; [2]Parents are first cousins
In bold are highlighted the undiagnosed patients

Among the 32 probands, four showed a clinical suspect of HS, nine of HST, 12 of CDAs and seven showed overlapping phenotypes or undefined condition. Based on HaloPlex Sure Design, the design for the RedPlex study on 34 genes encompassed 538 target regions within 239.764 Kb. After the submission of the design, the resulting amplicons were 8874 with total sequenceable size regions of 419.47 Kb. RedPlex panel showed high sensitivity and specificity. All regions resulted at high coverage (>= 90%), with a total target coverage of 99.93 %. Moreover, approximately 97.0% and 90.0% of ROIs were covered by at least 100 and 200 reads, respectively. The insufficiently covered regions were found mostly outside coding regions. Several non-synonymous variants of unknown significance were identified in the largest genes. For the majority of patients the molecular findings obtained by RedPlex analysis confirmed the clinical suspect, while for other patients the multi-genic approach has modified the initial suspicion (Figure 4-2). Indeed, we were able to obtain a conclusive diagnosis in 23 out 32 patients (approximately 72%); instead, nine patients (RED 10, 14, 17, 18, 20, 21, 23, 26, 35) remained undiagnosed (Table 4-3). We also identified some patients showing multiple disease-associated variants suggesting complex inheritance, as described in the next section of the results (Figure 4-2).

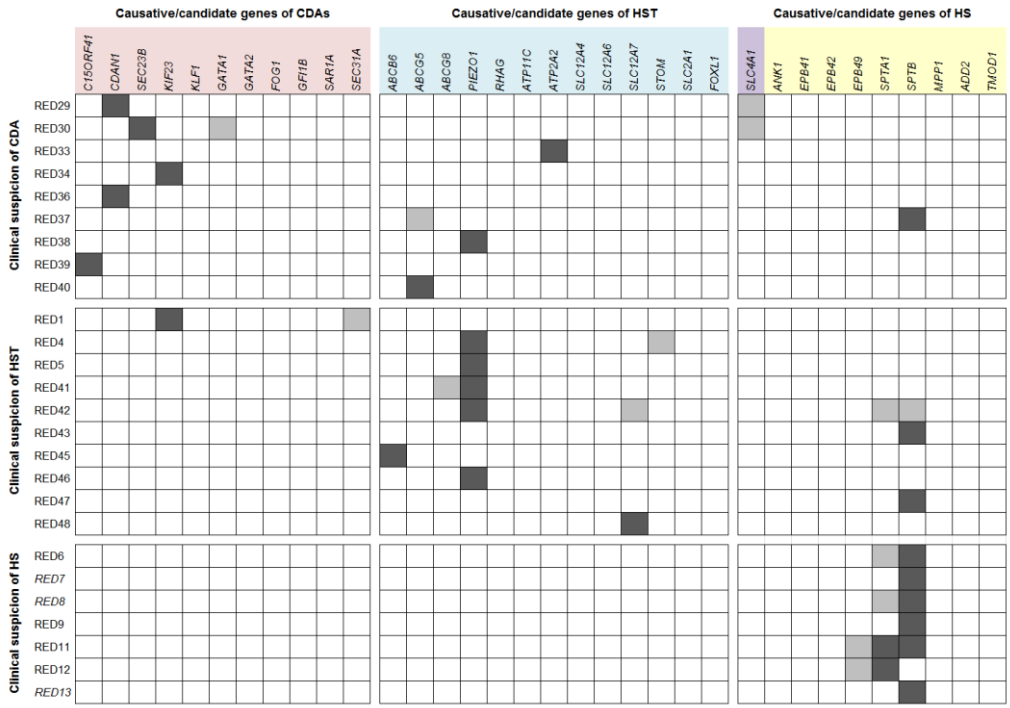

Figure 4-2. T-NGS data from RedPlex study on 23 HHA patients. In the column the causative/candidate genes analyzed by RedPlex divided according to different HHA phenotypes; in the row the investigated patients divided according to their clinical suspicion. Dark gray squares highlight variants related to the phenotype or those prioritized as damaging by prediction tools. Light gray squares indicate clinically related/modifier variants. In italics are indicate relatives of the probands (see Table 4-3).

4.3 Clinical features of HHA8 and RED30 patients

We functionally investigated both HHA8 and RED30 patients. Both showed clinical and biochemical data compatible with the diagnosis of CDA II. Clinical features of the two proband are described below, and hematologic status of each proband at diagnosis is summarized in Table 4-4.

Table 4-4. Clinical data of HHA8 and RED30 patients

	HHA8	**RED30**
Age at diagnosis (years)	22	48
Onset symptoms (years)	neonatal	40
Complete blood count		
RBC (10^6/µl)	2.2	
Hb (gr/dl)	8.0	10.0
Ht (%)	22.2	
MCV (fl)	89	86
MCH (pg)	29.0	
MCHC (gr/dl)	29.0	
Retics %	4.6	

Absolute retics count (10^6/L)	17520	102000
PLT (10^3/μL)	32.0	198.0
Biochemical data		
Bc+u/Bu (mg/dl)	2.6/0.7	7.5/0.7
LDH (U/L)	1121	171
Hpt (mg/dl)	9.3	11.0
Ferritin (ng/dl)	313	1157
IS transferrin (%)	25	85

Bc, conjugated bilirubin; Bu, unconjugated bilirubin
Hpt, Haptoglobin

Case1 (HHA8). Twenty-two years old male, third son of healthy non-consanguineous parents of Italian origin. At birth respiratory distress, hemorrhagic manifestations and hypospadias were observed. Family history is not indicative for anemia, both brothers were monorchid. At four months old, anemia with thrombocytopenia was observed (Table 4-4). During childhood two transfusions of packed RBCs and several PLT transfusions were performed. Peripheral blood smear shows anisopoikolocytosis and hypochromic red cells; some erythroblasts were also present. BM analysis revealed dyserythropoiesis with bi- and multi-nucleated erythroblasts. EM analysis revealed the presence of double membranes in erythroblasts, a typical feature of CDA II, as well as ultrastructural abnormalities in megakaryocytes and platelets. Particularly, lack of PLT demarcation membranes, presence of cytoplasmic vacuoles, net decreased dense granules were described. Thrombocytopenia with some large platelets with fairly normal granulation pattern was also observed. At present, the patient does not require transfusions. He shows macrocytic anemia (Hb 9.8 g/dl; MCV 100.3fl), with a reticulocyte count not adequate to the degree of anemia (150000/μl). The PLT count is very low (25000/μl) and petechiae are observed at physical examination.

Case2 (RED30). Male, 48 years old, one of four children of healthy parents of Punjabi origin. Family history is not indicative for anemia or jaundice. At 40 years he shows jaundice at physical examination. He was previously fit and well. Ultrasound scan of the abdomen showed multiple gallstones. A subsequent magnetic resonance cholangiopancreatography showed a low signal density liver consistent with iron overload. Indeed, ferritin level was 1157 ng/dl and transferrin saturation 85% (Table 4-4); however no mutations in *HFE* gene were detected. Mild normocytic anemia was observed (Hb 10.0 g/dl, MCV 86.1 fl) with a reticulocyte count not adequate to the degree of anemia (102000/μl). Platelet count was 198000/μl. Hemolysis was confirmed by low level of haptoglobin. No mutations in *HBA* and *HBB* genes were found. Deficit of pyruvate kinase and *G6PD*, porphyria, sideroblastic anemia and hepatitis viruses (HAV, HBV and HCV) were also ruled out. Blood smear showed polychromasia, nucleated red blood cells, some irregular contracted cells and spherocytes. BM examination showed hypercellularity with severe dyserythropoiesis with multinucleated forms, mild internuclear bridges. No changes in leukocytes and megakaryocytes were observed. Iron content in

macrophages was high with many sideroblasts but no evidence of ring sideroblasts. EM of BM showed binucleated erythroblasts with double membrane and SDS-PAGE highlighted the presence of hypoglycosylated band 3, a hallmark of CDA II. At 43 years the proband started on regular venesections for the high levels of ferritin; Hb level was maintained and he did not require blood transfusions. Ferritin level reduced slightly this period. Screening for complications of hemochromatosis was also performed. He had a liver biopsy which revealed early liver cirrhosis with iron overload. He developed diabetes and started on Insulin Lantus. He complained of impotence which may also be a complication of his underlying hemochromatosis (FSH and LH levels were normal).

4.4 Identification of mutations in GATA1 and SEC23B genes

We analysed both patients by t-NGS (Figures 4-1, 4-2). The first proband HHA8 was hemizygous for the c.722G>A transition in GATA1 gene, which results in the aminoacidic substitution p.Gly208Arg; this missense mutation has been already described as causative of XLTDA (Del Vecchio et al. 2005; Kratz et al. 2008). Accordingly to X-linked inheritance pattern, the mutation was inherited from the mother. No additional mutations have been found in *SEC23B* gene (Figure 4-3a).

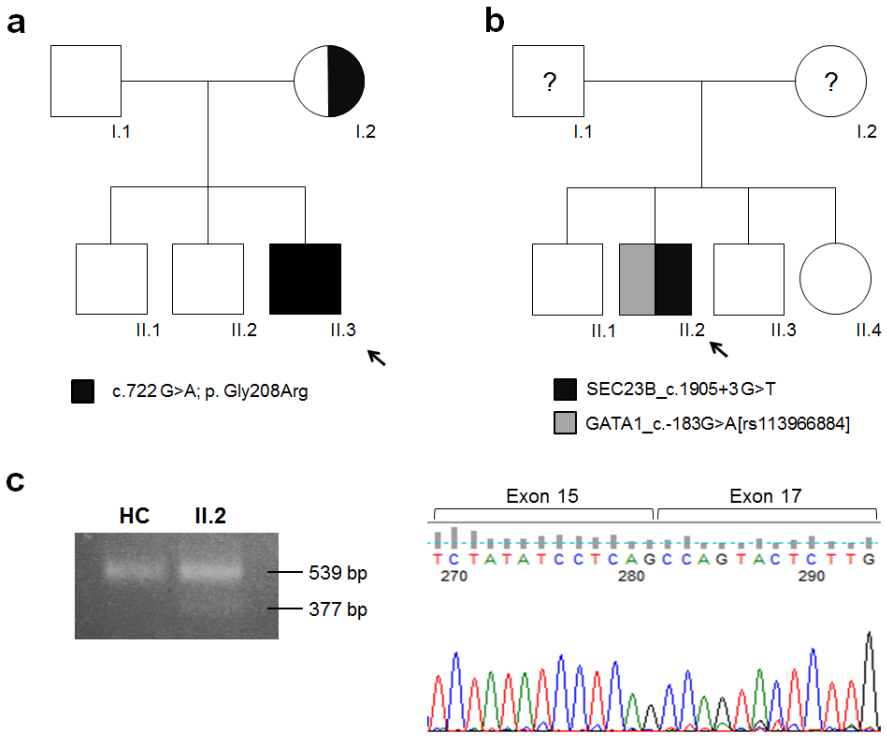

Figure 4-3. GATA1 and SEC23B mutations. (a) Inheritance pattern of *GATA1* mutation c.722G>A, Gly208Arg, was shown. (b) Proband II.2 showed a digenic inheritance for *SEC23B*

intronic mutation, c.1905+3G>T, and *GATA1* c.-71-92G>A rare variant. DNA from relatives was not available. (c) Amplification by PCR of the exon region 17-20 of *SEC23B* cDNA from patient II.2 and a control subject (HC). After the PCR, the products were separated on 1.5% agarose gel. The top band measures 539 bp, while the bottom band in lane 2 measures approximately 377 bp (exon-16 skipped transcript). On the left, sequencing analysis of the 377-bp PCR product is shown.

RedPlex analysis in the second proband highlighted the presence of a novel intronic mutation, c.1905+3G>T, in heterozygous state in *SEC23B* gene (Figure 4-3b). *In silico* analysis predicted for this nucleotide replacement a slight reduction of the score between wild type and mutated donor site sequence (ref score=95.15, mut score=91.28, variation=-4.06%). Accordingly, amplification of the specific exon regions, encompassing the mutation, of SEC23B cDNA from whole blood mRNA of the patient highlighted the presence of two bands on agarose gel: one corresponding to the expected size fragment (539 bp) and an additional 162-bp shorter transcript, due to the skipping of exon 16, as confirmed by sequencing analysis of both cDNA products (Figure 4-3c). However, we did not find any other mutation in *SEC23B* gene, so the mutational pattern in this patient resulted incomplete. Moreover, we excluded large deletions within the coding region of the gene. Interestingly, RedPlex analysis highlighted the presence of a rare nucleotide transition c.-183G>A in the 5'upstream region of *GATA1* gene. This single nucleotide variant (SNV rs113966884) is present in 1000 Genomes project (http://www.1000genomes.org/), with a minor allele frequency 0.03, although some variations between different populations are observed (Figure 4-4a). We evaluated *GATA1* expression in healthy subjects stratified according to GATA1-rs113966884 genotype: it was shown that allele A is associated to a significant reduced expression of GATA1 compared to allele G (p = 0.01). A reduced SEC23B expression in control subjects with GATA1-rs113966884 allele A compared to allele G was observed, albeit not significant (p = 0.15) (Figure 4-4b).

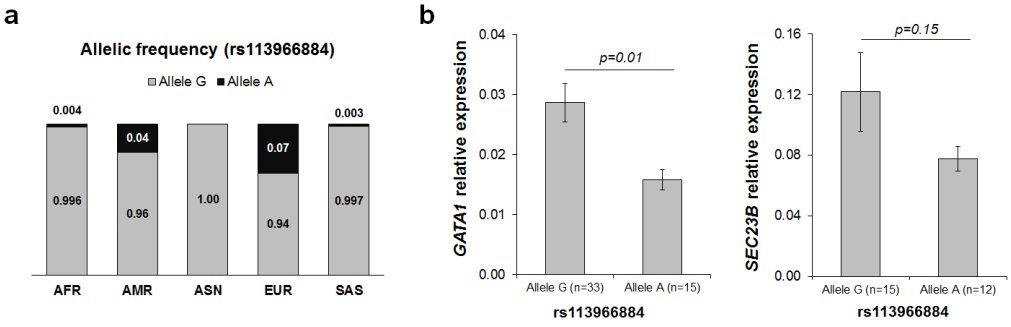

Figure 4-4. Genetic and functional description of GATA1-rs113966884. (a) Absolute allele frequencies of GATA1-rs113966884 SNV in control populations from 1000 Genomes project (AFR, African; AMR, Ad Mixed American; ASN, East Asian; EUR, European). (b) Gene relative expression of *GATA1* and *SEC23B* genes respect to the reference gene, β-actin, in 48

and 27 healthy subjects, respectively, stratified according to the genotype of GATA1-rs113966884 SNV is shown. Subjects with genotypes AG/AA/A0 showed a reduced expression of both genes when compared with genotypes GG/G0 (GATA1: allele G, 0.03±0.003; allele A, 0.02±0.002 – SEC23B: allele G, 0.12±0.03; allele A, 0.08±0.008). Data are presented as mean ± SE. P value has been calculated by Student t test.

4.5 The mutations c.722G>A and c.-183G>A account for a reduced expression of GATA1 and SEC23B

In order to establish the effect of both *GATA1* mutations, c.722 G>A and c.-71-92G>A, on GATA1 expression we firstly performed a qRT-PCR analysis of the gene in the two probands. We observed a marked downregulation (approximately 90%) of GATA1-c.722G>A mRNA expression in the first proband II.3 compared to healthy subjects (p=0.04). Accordingly, his mother I.2, heterozygous for the mutation, showed a reduced GATA1 expression of approximately 50%. Although sequencing analysis did not reveal any additional mutations in the coding sequence of SEC23B gene, expression profiling of this gene showed the same reduced trend in the proband II.3 (p=0.005) and in his mother I.2 (Figure 4-5a). WB analyses revealed comparable results for both mutations (Figure 4-5c).

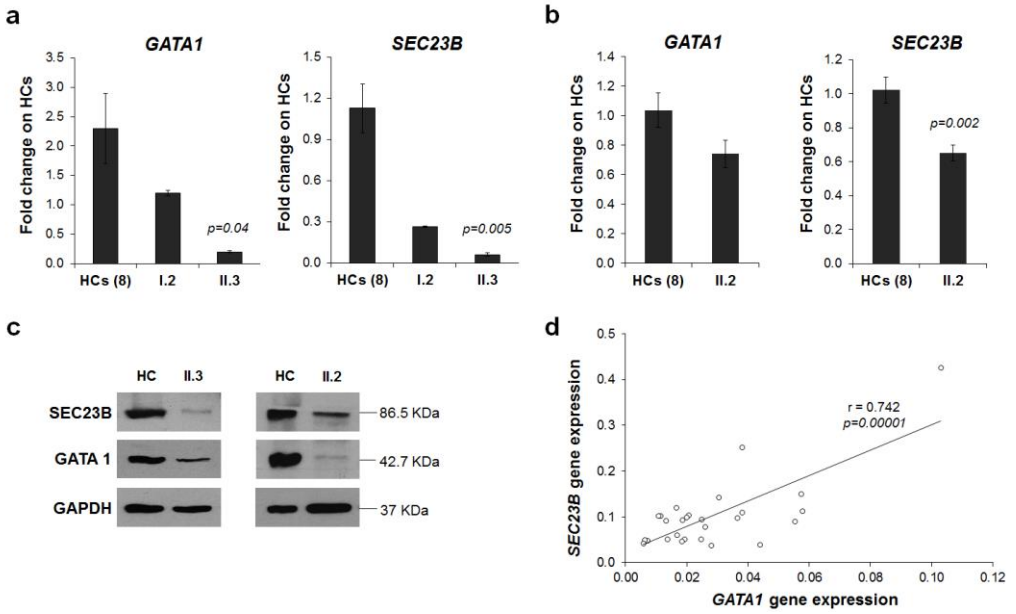

Figure 4-5. Analysis of the effect of *GATA1* mutations on GATA1 and SEC23B expression. (a) *GATA1* and *SEC23B* relative expression respect to HCs (n = 8) is shown. GATA1-G208R patient II.3 and in his mother I.2 showed a reduced expression of both genes *GATA1* (HCs, 2.29±0.60; I.2, 1.20±0.04; II.3, 0.20±0.02) and *SEC23B* (HCs, 1.13±0.18; I.2, 0.26±0.004; II.3, 0.06±0.01). Data are presented as mean ± SE. P value has been calculated by Student t test. (b) *GATA1* and *SEC23B* relative expression of patient II.2 respect to HCs (n = 8) is shown. Patient II.2 showed a slight reduction of *GATA1* gene expression (HCs, 1.04±0.11; II.2, 0.74±0.09) and a marked reduction of *SEC23B* expression (HCs, 1.02±0.08; II.2,

0.65±0.05). Data are presented as mean ± SE. P value has been calculated by Student t test. (c) WB analysis highlighted a marked reduced expression of GATA1 and SEC23B in both patients compared to HC. (d) Correlation analysis between *SEC23B* and *GATA1* gene expression, performed in 27 HCs, showed a direct correlation between the two genes (Pearson correlation r = 0.742, p = 0.00001).

In this second case, we also suspected an impaired GATA1 expression due to the variation c.-71-92G>A in the 5'upstream region of the gene. Accordingly, we observed a reduced gene expression in the proband II.2 compared to healthy controls, even if not significant (p=0.08) (Figure 4-5b). Since this patient was a compound heterozygous for a splice site mutation in *SEC23B* gene, a strong hypo-expression of the gene was observed when compared to healthy controls (p=0.002) (Figure 4-5b). WB analyses revealed comparable results for both mutations (Figure 4-5c).

Finally, correlation analysis performed on PBMCs from 27 healthy subjects revealed a significant direct correlation between *GATA1* and *SEC23B* expression (p=0.00001) (Figure 4-5d).

4.6 Functional characterization of GATA1-G208R mutation

In order to confirm the *ex-vivo* data on PBMCs from patient II.3, we cloned GATA1 wild type (WT) and G208R mutated sequence in the expressing vector pCDNA3.1. Then we transfected both constructs in HEK-293 cells. We observed a reduced GATA1 gene expression, albeit not significant, in cells transfected with pCDNA3.1-GATA1-G208R construct when compared to those detected in cells transfected with wild type sequence (fold = 0.72, p = 0.08) (Figure 4-6a). Accordingly to previous suggestion, a reduced endogenous expression of SEC23B was observed in mutated GATA1 clone when compared to wild type (fold = 0.66, p = 0.002) (Figure 4-6a). WB analyses revealed comparable results for both proteins (Figure 4-6b).

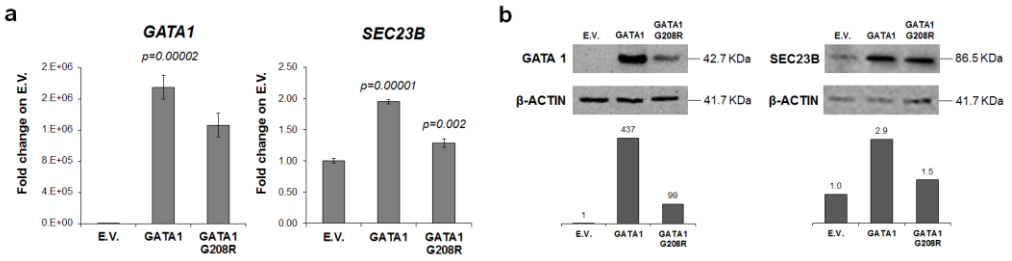

Figure 4-6. *In vitro* characterization of GATA1-G208R mutation. (a) *GATA1* and *SEC23B* relative expression of HEK-293 cells overexpressing pCDNA3.1-GATA1 and pCDNA3.1-GATA1-G208R compared to those transfected with empty vector (E.V.) were shown. GATA1-G208R construct showed a reduced expression compared to wild type sequence of both genes GATA1 (pCDNA3.1-GATA1, 1749833±150723; pCDNA3.1-GATA1-G208R, 1264421±151656) and SEC23B (pCDNA3.1-GATA1, 1.95±0.03; pCDNA3.1-GATA1-

G208R, 1.29±0.07). Data are presented as mean ± SE. P value has been calculated by Student t test. (b) WB analysis showed a reduced expression of GATA1 and SEC23B in HEK-293 cells transfected with pCDNA3.1-GATA1-G208R of comparable extent to that observed by qRT-PCR. The histograms show the densitometric quantification. Sizes (in kDa) are on the right.

4.7 Characterization of the human promoter region of SEC23B gene

In order to characterize HuSEC23B promoter we examined 4292 bp upstream the ATG of the gene. The evolutionary conservation of this nucleotide sequence was analysed by ENCODE web tool: it showed a highly conserved region of approximately 500 bp, HuSEC23B/3.29-2.78, which starts 3300 bp upstream ATG of the gene. This genomic sequence exhibited an enrichment of the H3K27Ac histone mark in K562 cell line, suggesting a high accessibility of the chromatin to transcription in this locus (Figure 4-7, top). Of note, HuSEC23B/3.29-2.78 region is predicted to contain a 720 bp CpG island by both *in silico* tools, CpG Islands Track and EMBOSS Cpgplot. However, the analysis of the methylation status of specific CpG dinucleotides in several cell types showed that this CpG island is unmethylated in K562 cell line, suggesting again an active transcription of this upstream region of HuSEC23B gene. Accordingly to *in silico* data, HuSEC23B/3.44 fragment showed a very high luciferase activity compared to pGL3 empty vector (fold = 933.8) in K562 cells (Figure 4-7, bottom), while HuSEC23B/2.52 exhibited a very low activity over empty vector (fold =7.1). Of note, this fragment does not include the active transcription region HuSEC23B/3.29-2.78. Although at low levels, the smallest fragments (HuSEC23B/0.45, HuSEC23B/0.21) upstream *SEC23B* gene still activated luciferase transcription. As expected, this region is predicted by MatInspector to contain the vertebrate TATA motif. Comparable results were obtained in HEK-293 cell line (data not shown).

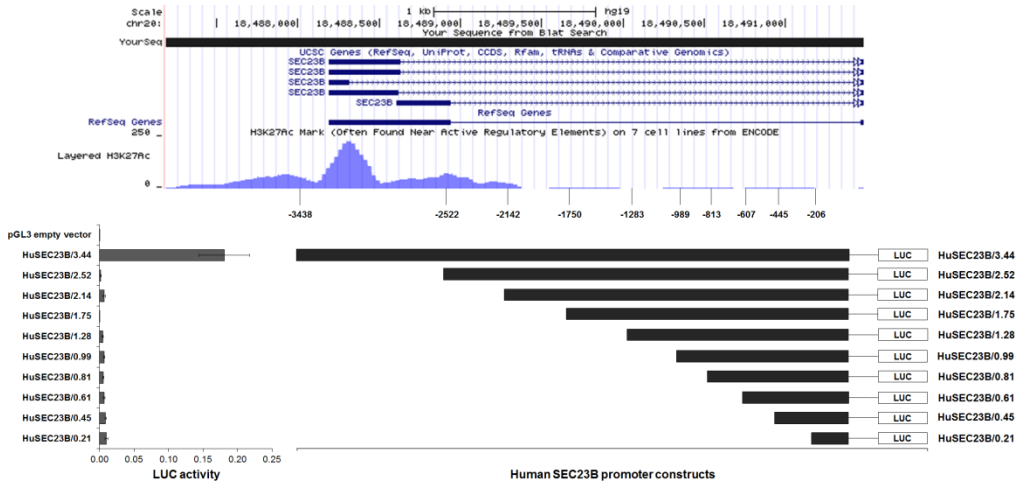

Figure 4-7. *In silico* and functional analyses of HuSEC23B promoter. (a) *In silico* analysis of HuSEC23B upstream region is shown. The H3K27Ac histone mark is the acetylation of lysine 27 of the H3 histone protein, and it is thought to enhance transcription. The blue track

shows the levels of enrichment of the H3K27Ac histone mark across the genome as determined by a ChIP-seq assay in K562 cell line from UCSC Genome Browser (ENCODE web tool). (b) A schematic representation of HuSEC23B promoter constructs is shown (on the right). The HuSEC23B promoter constructs containing the luciferase reporter gene have been transiently transfected into K562 cells. Co-transfection with the Renilla luciferase plasmid (pRL-TK) has been performed to normalize transfection efficiency. Luciferase normalized activity for each construct is shown on the left.

4.8 GATA1 protein binds to GATA binding sites located in the first intron of HuSEC23B

Ten putative binding sites for GATA1 transcription factor (GATA1bs) were predicted by MatInspector web server tool in the promoter sequence of *SEC23B*, in both strands, sense and antisense. Similarly, PROMO predicted the presence of 12 GATA1bs in the same genomic sequence. We selected the sense binding sites common to both prediction results. The sites which did not reflect the consensus sequence W(A/T)GATAR(A/G) were excluded for further analyses (Figure 4-8a). We demonstrated the direct binding of GATA1 to the GATA1bs/HuSEC23B-463 by ChIP assays in both K562 and HEL cells at 6 days of erythroid differentiation, while the site GATA1bs/HuSEC23B-2475 resulted directly bound by GATA1 only in HEL cells at 6 days of erythroid differentiation. Of note, the antisense GATA1bs/HuSEC23B-3735 region is not bound by the transcription factor (Figure 4-8b-c).

In order to validate the effect of the direct binding of GATA1 on HuSEC23B promoter, we co-transfected both GATA1bs/HuSEC23B-2475 and -463 mutants with GATA1 WT in HEK293 cells. Co-transfection of GATA1 WT with GATA1bs/HuSEC23B-2475 and -463 mutants induced a reduction of 30-40% luciferase activity respectively, compared to HuSEC23B/3.44 WT sequence (Figure 4-8d).

Finally, we analyzed the GATA1-mediated regulation of *SEC23B* expression by co-transfection of several GATA1 mutants (GATA1/G208R, GATA1/R216W, GATA1/D218G, and GATA1/V205M) and the HuSEC23B/3.44 in HEK-293 cell line. We demonstrated a reduction of HuSEC23B/3.44 luciferase activity induced by the mutants G208R (32%) and R216W (40%) compared to those observed in co-transfected cells with GATA1 WT (Figure 4-8e).

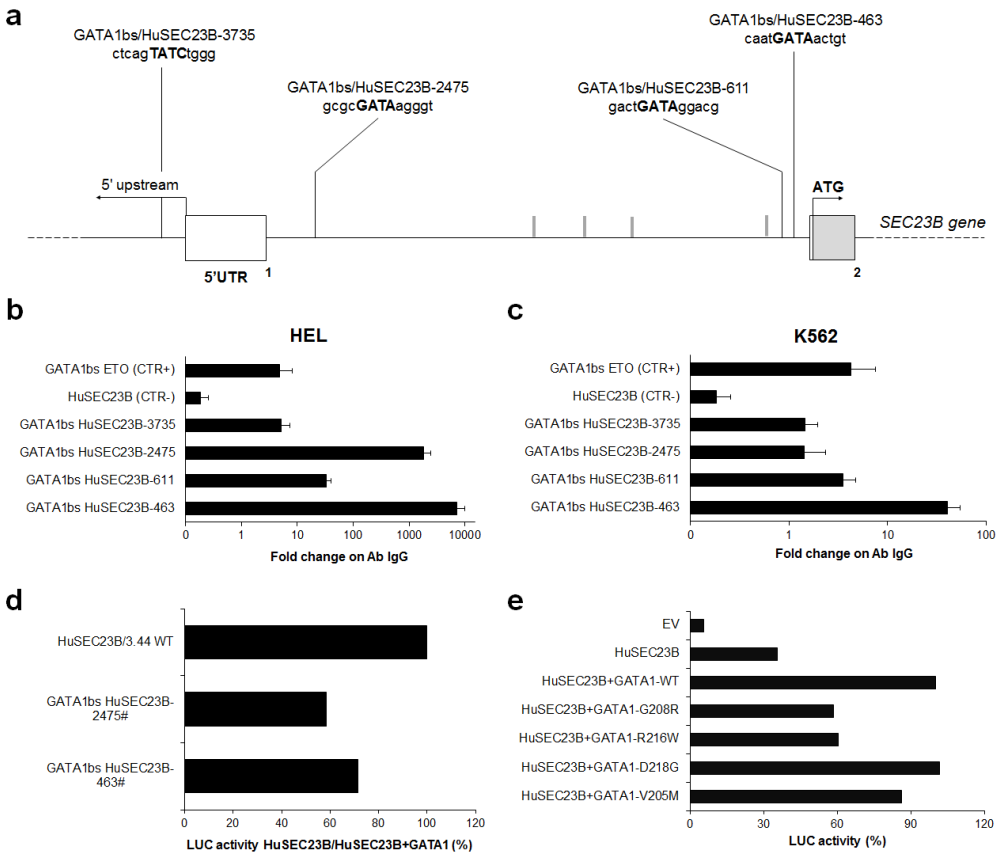

Figure 4-8. GATA1 directly binds to HuSEC23B promoter. (a) *In silico* analysis of HuSEC23B upstream region showed the presence of putative binding sites of the transcription factor GATA1. The binding sites predicted by both in silico tools (MatInspector and PROMO) were shown; light gray bar indicate the sense sites predicted by only one tool. (b-c) qRT-PCR analysis on GATA1 immunoprecipitated DNA in HEL and K562 cells at 6 days of differentiation by hemin. GATA1bs/ETO is the CTR+. HuSEC23B (CTR-) is a sequence flanking GATA1bs within HuSEC23B promoter. (d) Luciferase normalized activity of HEK-293 cells co-transfected with GATA1bs/HuSEC23B-2475 and -463 mutants and GATA1 WT. (e) Luciferase normalized activity of HEK-293 cells co-transfected GATA1 mutants (GATA1/G208R, GATA1/R216W, GATA1/D218G, and GATA1/V205M) and the HuSEC23B/3.44 in HEK-293 cell line.

5. Discussion

Mutations in more than 30 genes cause HHA, a highly heterogeneous group of rare anemias characterized by complex and often unexplained genotype-phenotype correlations. This group of disorders includes conditions with very different phenotypes, ranging from (1) hyporegenerative anemias, as CDAs to (2) hemolytic anemias due to RBC membrane defects, as HS and HST. Although the workflow to diagnose these conditions is a normal clinical practice, differential diagnosis, classification, and patient stratification among HHA are often very difficult. Indeed, the variety of unspecific and overlapping phenotypes often hampers a correct clinical management of the patients. In our experience, the clinical and molecular definition of some HHA patients can be frequently very hard to obtain. For some conditions, the great phenotypic variability is partially explained with the high genetic heterogeneity; this is the case of HS, in which five different causative genes have been described (Iolascon et al. 2003). However, it is sometimes complicated to distinguish one form to the others since the signs can be veiled in symptom-free carriers and in mild forms. Moreover, HS can be confused with other clinically-related hereditary hemolytic conditions such as CDA II. Indeed, because of hemolytic symptoms, frequently CDA II is first misdiagnosed as HS. CDA II patients erroneously diagnosed as HS often incur unnecessary interventions, such as splenectomy, which leads to a moderate increase in Hb concentration without reaching the normal range. The lack of substantial improvement after intervention frequently leads to a re-examination of the case, allowing the correct diagnosis of CDA II. The most useful indicator to correctly address the diagnosis of CDA II is the inadequate reticulocyte count for the degree of anemia. Recently, a new clinical index, named bone marrow responsiveness index (BMRI) has been created to discriminate a well-compensated hemolytic anemia from another characterized by ineffective erythropoiesis. Although this index showed a high sensitive parameter (90.4%) to achieve a targeted clinical diagnosis of CDA II, it exhibited a low specificity (64.9%) for discriminating HS and CDA II patients (Russo et al. 2014). Up to date the cost and time for diagnosis of HHA are high. This is largely due to the difficulties in establishing the correct diagnosis. This latter is generally based on either biochemical or molecular tests that require advanced technical skills that are not available in all countries. As a consequence, a long period of time passes before a patient can be conclusively diagnosed. This has major psychological impact on the patients and their families. Accurate genetic diagnosis of HHA is important not only for achieving a definitive diagnosis but also for determining appropriate genetic counseling, prognosis and treatment. This is particularly true for unspecific symptomatic treatment for hemolytic anemias, as splenectomy, that could result not only unuseful, as in CDA II, but also damaging for the patients. For example, splenectomy should be banned in DHS for it entails, with near certainty, thromboembolic accidents which may be very severe

(Stewart et al. 1996). Timely diagnosis could significantly reduce the distress for patients and their families, diagnostic costs, and it could have a high socio-economic impact. Indeed, delayed diagnosis is one of the major denominators of life expectancy and secondary effects like multi-organ damage due to iron overload, thrombosis or chronic transfusion regimen.

In the last decade, remarkable progresses have been made in discovering new disease genes involved in red blood disorders (Schwarz et al. 2009; Andolfo et al. 2012; Andolfo et al. 2013; Babbs et al. 2013; Liljeholm et al. 2013; Andolfo et al. 2015; Rapetti-Mauss et al. 2015). This increasing genetic heterogeneity underlines the problem of a very complex differential diagnosis (Gambale et al. 2016). Furthermore, genome sequencing studies suggest that the clinical genetic test may be incomplete not only when the causative mutation is missing, but also when the genotype/phenotype correlation appears weak. This is particularly true when the familial recurrence is unclear, with some relatives that only share minor affections. The absence of clear genotype/phenotype correlations is often problematic for both genetic counselling and forthcoming mutation-specific treatments (Savarese et al. 2014). It is proper in this context that the new genomic technologies take place. NGS refers to non-Sanger-based high-throughput DNA sequencing technologies. NGS techniques have significantly amplified the opportunities of genome analysis. Indeed, this technology plays now a major role either in disease gene discovery or in clinical use for establishing a genetic diagnosis, with several implications on the diagnostic flowchart: the NGS tests could become the first tier test, preceding biopsy and other invasive procedures. We can envision a future in which the functional integration among the next-generation technologies for genomic screening and genomic editing will allow reaching the always pursued goal of a targeted diagnosis and therapy (Figure 5-1).

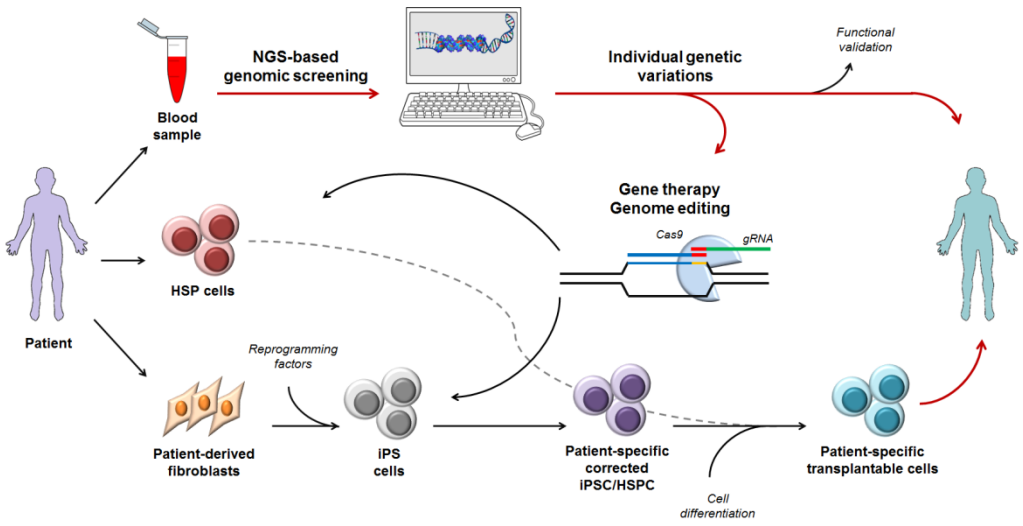

Figure 5-1. Integration between technological updates and clinical applications in diagnosis and therapy of RBCs diseases. Adult hematopoietic stem progenitor cells (HSPCs)

or induced pluripotent stem cells (iPSCs) can be used for gene-therapy approaches. DNA extracted from a peripheral blood sample can be used to identify genetic variations by NGS. The causative role of these variations can be validate by *in vitro/in vivo* functional studies and then the commonly used $CD34^+$ HSPCs may be corrected directly by gene therapy or genome editing by CRISPR/CAS9 technology. Alternatively, somatic cells can be isolated by fibroblasts of the patient and reprogrammed to pluripotency, with the resulting iPSCs then being corrected by gene therapy or genome editing and differentiated through erythroid lineage.

In the NGS era, the genetic testing is going to move from few candidate genes to wider panels of genes, namely t-NGS approaches. Recent studies have already demonstrated the usefulness of t-NGS as a comprehensive and invaluable diagnostic tool by means of achieving a correct diagnosis and proceeding with careful management of patients (Sikkema-Raddatz et al. 2013; Savarese et al. 2014). Indeed, t-NGS resulted to be cost-effective, faster and more specific than other NGS technologies, such as WES. Indeed, in WES applications typical coverage of exons is approximately 90-95%, with coverage much lower in some genes. By contrast, t-NGS can have a much higher or often almost complete coverage of all target regions. Of note, in our gene panel the predicted coverage has not dropped below 99.3% and the mean sequencing coverage of target regions was 99.3%. We used HaloPlex as target enrichment, which provides a high specificity and efficiency. Indeed, as described in a similar study less than 2% of reads generated by HaloPlex are off-target, in comparison with >12% obtained of WES (Savarese et al. 2014). Our data are in agreement with these observations. In the pilot study here described we were able to confirm the molecular diagnosis in all of the patients previously characterized at molecular level. Moreover, by means of pilot design we were also able to obtain a molecular characterization in five out seven (approximately 72%) HS patients, except for two patients, in which we did not find any causative variants. However, we were not able to identify in HHA2 patient a large deletion in *SPTA1* gene, previously detected (Iolascon et al. 2011). Indeed, large insertions, deletions, and other chromosomal abnormalities are often difficult to be identified by NGS approaches.

One of the most important aspects of the use of t-NGS gene panels in clinical practice is their ability to be easily upgradable in view of novel discoveries. Thus, after performing pilot study, we designed a broader gene panel, RedPlex, composed of 34 RBC-disease genes and assayed 32 HHA patients from 27 unrelated families. According to the pilot study data, we obtained a conclusive diagnosis in approximately 72% of cases. In addition, a lot of patients (39%) showed multiple mutations in different *loci*. The percentage of genetically complex patients could be higher, if we consider that many other important genes (even if not disease-causing) can also carry damaging alleles. Of note, in RedPlex panel we included not only causative genes but also candidate and modifier genes (12 out 34). We can easily predict that a broader NGS approach could strengthen this observation. For this reason, we have recently

implemented RedPlex including additional 37 genes. We are also designing a panel of 80-90 modifier genes of HHA. Indeed, monogenic diseases provide unique opportunities to understand the real contribution of causative single-gene mutations and to identify modifiers as they have uniform aetiology, detailed phenotyping of affected individuals, and familial clustering (Cutting 2010). Thus, beyond achieving a definitive diagnosis, knowing the genetic basis of these patients can be valuable also for their prognosis. NGS plays a pivotal role in this context. The implementation of NGS in clinical practice is already a matter of fact. It will result in increased knowledge of genetic and genomic differences among individuals that gradually will lead to shift the focus from population-based to patient-individualization of the clinical management (Iolascon et al. 2015). For the HHA diseases, the role of the modifier genes could assume a consistent significance because of the high intra-familial and inter-familial-phenotypic differences that may be explained with the combinations of multiple disease-causing alleles, or their combination with polymorphic variants.

In our cohort of examined patients, we found at least two interesting genetically complex patients. The first, identified as HHA14, was a patient with an overlapping phenotype HS/CDA. In this patient, the spherocytic component of the phenotype was due to mutations in HS-disease gene *SPTA1*, while the dyserythropoietic component was explainable with the presence of additional variants in CDA-related genes, *SEC23B* and *GATA1*. Similarly, in RED30 patient we observed a complex genotype *SEC23B-GATA1* which could explain the occurrence CDA II phenotype. Particularly, this patient showed a typical clinical and biochemical framework of CDA II; RedPlex analysis highlighted the presence of a monoallelic hypomorphic mutation in CDA II-disease gene *SEC23B*. Mutations in deep regulatory regions of the *SEC23B* gene as well as digenic inheritance have been hypothesized as additional pathogenetic mechanisms of CDA II (Russo et al. 2014). Of note, RED30 patients showed neither *SEC23B* regulatory mutations nor large deletions within the coding region of the gene. However, a low-expression GATA1 allele was found in the same patient. So, a digenic inheritance was hypothesized in this case. Thus, we further investigated the functional relationship between these two genes, demonstrating the direct interaction of GATA1 transcription factor on the promoter region of *SEC23B*. Of note, mutations in *GATA1* have been already to the CDA variant, XLTDA. *GATA1* gene encodes for the homonymous transcriptional factor. It regulates the expression of numerous erythroid-specific genes, such as erythropoietin receptor gene (Zon et al. 1991; Chiba et al. 1991), α- and β-globin genes (Whitelaw et al. 1990), *ALAS2* gene (Surinya et al. 1997; Kaneko et al. 2014) and *GATA1* gene itself (Kobayashi et al. 2001) during erythroid differentiation. Beyond XLTDA, other rare syndromes have been already associated to alterations of GATA1 gene expression or protein function. They are: X-linked thrombocytopenia, X-linked thrombocytopenia with thalassemia, congenital erythropoietic porphyria (CEP), transient myeloproliferative disorder, acute

megakaryoblastic leukemia associated with trisomy 21, and anemia associated with the production of GATA1s isoform (Ciovacco et al. 2008). GATA1 needs two zinc finger domains for normal function. The C-terminal finger is necessary for DNA binding, while the N-terminal mediates interaction with its cofactor FOG1. In 2000, Nichols described a family with XLTDA due to a substitution p.Val205Met (Nichols et al. 2000). It has been demonstrated that this highly conserved residue is necessary for interaction of the amino-terminal zinc finger domain of GATA1 with FOG1. This mutation most closely resembles those of mice with the knockdown Gata1 mutation. The authors showed that the Val205Met mutation abrogates the interaction between GATA1 and FOG1, thus inhibiting the ability of GATA1 to rescue erythroid differentiation in an erythroid cell line deficient for GATA1. Their findings underscore the importance of FOG1:GATA1 interaction in both megakaryocyte and erythroid development and suggest that other X-linked anemias or thrombocytopenias may be caused by defects in GATA1 (Nichols et al. 2000). In 2005, a novel single base mutation, leading to an amino acid substitution (Gly208Arg) within the highly conserved portion of the GATA1 N-terminal finger domain was described (Del Vecchio et al. 2005). Similarly to the previously one described, this mutation leads to a dyserythropoietic anemia and macrothrombocytopenia. In our cohort we identified an analogous case that has been erroneously diagnosed at first instance as CDA II. By t-NGS approach we were able to correctly modify the diagnosis of this patient towards a XLTDA; indeed, he showed the same causative mutation p.Gly208Arg previously described.

The study of regulatory network GATA1-mediated on *SEC23B* gene not only provided new insights into the molecular mechanisms of *SEC23B* regulation, but also suggested an explanation of the variability of phenotypes GATA1-related by means of the crosstalk of this gene with *SEC23B*. Indeed, we demonstrated a specific involvement of two GATA1 mutations in SEC23B expression, p.Gly208Arg and p.Arg216Trp, which are the causative variants related to XLTDA and CEP, respectively. This observation is also in agreement with a CEP case recently reported, in which the severity of the anemia has been explained by the co-inheritance of the CEP-related GATA1 mutation, p.Arg216Trp, with a rare SEC23B missense substitution p.Thr257Ile (SNV rs146917730, MAF 0.006) that was predicted damaging (Di Pierro et al. 2015). Finally, the identification of transcriptional regulatory elements in *SEC23B* promoter could allow the definitive diagnosis of CDAII patients with peculiar clinical phenotypes or those with incomplete mutation pattern.

6. Conclusions

HHA are chronic disorders with a highly variable clinical picture. Phenotypic variety does not only depend on the molecular defect but includes the mode of action of specific modifier genes. Knowing these modifier genes and the way they contribute to the clinical picture would be a major breakthrough in the clinical management of these patients. NGS plays a pivotal role in this context and its implementation in clinical practice is already a matter of fact. Indeed, NGS studies have identified a greater-than-expected number of genetic variations in the human genome. This suggested that existing clinical monogenic testing systematically can miss very relevant information, and that a lot of genetically complex patients need to be investigated.

Our analyses demonstrated that RedPlex represents a reliable diagnostic/prognostic tool for HHA patients. Indeed, it resulted to be a robust platform that overcomes for power, costs, speed, sensitivity and specificity the gene-by-gene strategy. Moreover, this approach also allowed the identification of "polygenic" conditions, i.e. patients in which the phenotypic variability could be explained by the presence of modifier variants associated to causative mutations. For the HHA diseases, the role of the modifier genes could assume a consistent significance because of the high intra-familial and inter-familial-phenotypic differences that may be explained with the combinations of multiple disease-causing alleles, or their combination with polymorphic variants. We investigated the functional interaction between two CDA-related genes, *GATA1* and *SEC23B*, thus providing not only new insights into the molecular mechanisms of *SEC23B* regulation, but also an explanation of the variability of phenotypes GATA1-related by means of the crosstalk of this gene with *SEC23B*.

7. Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Achille Iolascon for the continuous support of my PhD study and related research, for his valuable guidance, motivation, and immense knowledge. His supervision helped me in all the time of research.

I am particularly grateful to Dr. Mario Capasso for enlightening me the first glance of research.

I would like to acknowledge Dr. Immacolata Andolfo as the second player of this study, and I am gratefully indebted to her for her very valuable comments and supports on this thesis.

I would also like to thank the other collaborators who were involved in this research project: Francesco Manna, Gianluca De Rosa, Antonella Gambale, Alessandra Arillo. Without their passionate participation and input, their insightful comments and encouragement, this study could not have been successfully conducted.

My sincere thanks also go to Dr. Luigia De Falco and Dr. Flora Cimmino: without their precious support it would not be possible to conduct this research.

I thank my fellow labmates, Marianna, Lucia, Piero, Alessandro, Mariasole, Daniela, for the stimulating discussions and also for all the fun we have had in the last three years.

Last but not least; I would like to thank my family: my parents, my brothers, my sisters in law and my nephews for supporting me spiritually throughout this journey and my life in general.

8. References

Ajore R, Dhanda RS, Gullberg U, Olsson I. The leukemia associated ETO nuclear repressor gene is regulated by the GATA-1 transcription factor in erythroid/megakaryocytic cells. BMC Mol Biol. 2010;11:38.

Andolfo I, De Falco L, Asci R, Russo R, Colucci S, Gorrese M, Zollo M, Iolascon A. Regulation of divalent metal transporter 1 (DMT1) non-IRE isoform by the microRNA Let-7d in erythroid cells. Haematologica. 2010;95:1244-52.

Andolfo I, Alper SL, De Franceschi L, Auriemma C, Russo R, De Falco L, Vallefuoco F, Esposito MR, Vandorpe DH, Shmukler BE, Narayan R, Montanaro D, D'Armiento M, Vetro A, Limongelli I, Zuffardi O, Glader BE, Schrier SL, Brugnara C, Stewart GW, Delaunay J, Iolascon A. Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in PIEZO1. Blood. 2013;121:3925-35

Andolfo I, Alper SL, Delaunay J, Auriemma C, Russo R, Asci R, Esposito MR, Sharma AK, Shmukler BE, Brugnara C, De Franceschi L, Iolascon A. Missense mutations in the ABCB6 transporter cause dominant familial pseudohyperkalemia. Am J Hematol. 2013;88:66-72.

Andolfo I, Russo R, Manna F, Shmukler BE, Gambale A, Vitiello G, De Rosa G, Brugnara C, Alper SL, Snyder LM, Iolascon A. Novel Gardos channel mutations linked to dehydrated hereditary stomatocytosis (xerocytosis). Am J Hematol. 2015;90:921-6.

Arnaud L, Saison C, Helias V, Lucien N, Steschenko D, Giarratana MC, Prehu C, Foliguet B, Montout L, de Brevern AG, Francina A, Ripoche P, Fenneteau O, Da Costa L, Peyrard T, Coghlan G, Illum N, Birgens H, Tamary H, Iolascon A, Delaunay J, Tchernia G, Cartron JP. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. Am J Hum Genet. 2010;87:721-7.

Babbs C, Roberts NA, Sanchez-Pulido L, McGowan SJ, Ahmed MR, Brown JM, Sabry MA; WGS500 Consortium, Bentley DR, McVean GA, Donnelly P, Gileadi O, Ponting CP, Higgs DR, Buckle VJ. Homozygous mutations in a predicted endonuclease are a novel cause of congenital dyserythropoietic anemia type I. Haematologica. 2013;98:1383-7.

Carella M, Stewart G, Ajetunmobi JF, Perrotta S, Grootenboer S, Tchernia G, Delaunay J, Totaro A, Zelante L, Gasparini P, Iolascon A. Genomewide search for dehydrated hereditary stomatocytosis (hereditary xerocytosis): mapping of locus to chromosome 16 (16q23-qter). Am J Hum Genet. 1998;63:810-6.

Carella M, d'Adamo AP, Grootenboer-Mignot S, Vantyghem MC, Esposito L, D'Eustacchio A, Ficarella R, Stewart GW, Gasparini P, Delaunay J, Iolascon A. A second locus mapping to 2q35-36 for familial pseudohyperkalaemia. Eur J Hum Genet. 2004;12:1073-6.

Chiba T, Ikawa Y, Todokoro K. GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. Nucleic Acids Res. 1991;19:3843-8.

Chirnomas SD, Kupfer GM. The inherited bone marrow failure syndromes. Pediatr Clin North Am. 2013;60:1291-1310.

Ciovacco WA, Raskind WH, Kacena MA. Human phenotypes associated with GATA-1 mutations. Gene. 2008;427:1-6.

Cutting GR. Modifier genes in Mendelian disorders: the example of cystic fibrosis. Ann N Y Acad Sci. 2010;1214:57-69.

De Franceschi L, Turrini F, del Giudice EM, Perrotta S, Olivieri O, Corrocher R, Mannu F, Iolascon A. Decreased band 3 anion transport activity and band 3 clusterization in congenital dyserythropoietic anemia type II. Exp Hematol. 1998;26:869-73.

Del Vecchio GC, Giordani L, De Santis A, De Mattia D. Dyserythropoietic anemia and thrombocytopenia due to a novel mutation in GATA-1. Acta Haematol. 2005;114:113-6.

Delaunay J. The hereditary stomatocytoses: genetic disorders of the red cell membrane permeability to monovalent cations. Semin Hematol 2004;41:165-72.

Delaunay J. The molecular basis of hereditary red cell membrane disorders. Blood Rev 2007;21:1-20.

Desai AN, Jere A. Next-generation sequencing: ready for the clinics? Clin Genet. 2012;81:503-10.

Dgany O, Avidan N, Delaunay J, Krasnov T, Shalmon L, Shalev H, Eidelitz-Markus T, Kapelushnik J, Cattan D, Pariente A, Tulliez M, Crétien A, Schischmanoff PO, Iolascon A, Fibach E, Koren A, Rössler J, Le Merrer M, Yaniv I, Zaizov R, Ben-Asher E, Olender T, Lancet D, Beckmann JS, Tamary H. Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1. Am J Hum Genet. 2002;71:1467-74

Di Pierro E, Russo R, Karakas Z, Brancaleoni V, Gambale A, Kurt I, Winter SS, Granata F, Czuchlewski DR, Langella C, Iolascon A, Cappellini MD. Congenital erythropoietic porphyria linked to GATA1-R216W mutation: challenges for diagnosis. Eur J Haematol. 2015;94:491-7.

Fargo JH, Kratz CP, Giri N, Savage SA, Wong C, Backer K, Alter BP, Glader B. Erythrocyte adenosine deaminase: diagnostic value for Diamond-Blackfan anaemia. Br J Haematol. 2013;160:547-54.

Gallagher PG. Disorders of red cell volume regulation. Curr Opin Hematol. 2013;20:201-7.

Gambale A, Iolascon A, Andolfo I, Russo R. Diagnosis and management of congenital dyserythropoietic anemias. Expert Rev Hematol. 2016;9:283-96.

Genetet S, Ripoche P, Picot J, Bigot S, Delaunay J, Armari-Alla C, Colin Y, Mouro-Chanteloup I. Human RhAG ammonia channel is impaired by the Phe65Ser mutation in overhydrated stomatocytic red cells. Am J Physiol Cell Physiol. 2012;302:C419-28.

Heimpel H, Iolascon A. Congenital dyserythropoietic anaemias ESH Handbook on Disorders of Iron Metabolism (IRON2009) [Internet]. 2009. [cited 2009 Dec 4]. Available from: http://www.esh.org/esh-handbook-on-disorders-ofiron-metabolism-2009/.

Heimpel H, Kellermann K, Neuschwander N, Högel J, Schwarz K. The morphological diagnosis of congenital dyserythropoietic anemia: results of a quantitative analysis of peripheral blood and bone marrow cells. Haematologica. 2010a;95:1034-6.

Heimpel H, Matuschek A, Ahmed M, Bader-Meunier B, Colita A, Delaunay J, Garcon L, Gilsanz F, Goede J, Högel J, Kohne E, Leichtle R, Munoz J, Perrotta S, Piscopo C, Renella R, Schwarz K, Smolenska-Sym G, Wickramasinghe S, Zanella A, Iolascon A. Frequency of congenital dyserythropoietic anemias in Europe. Eur J Haematol. 2010b;85:20-5.

Iolascon A, Perrotta S, Stewart GW. Red blood cell membrane defects. Rev Clin Exp Hematol. 2003;7:22-56.

Iolascon A, De Falco L, Borgese F, Esposito MR, Avvisati RA, Izzo P, Piscopo C, Guizouarn H, Biondani A, Pantaleo A, De Franceschi L. A novel erythroid anion exchange variant (Gly796Arg) of hereditary stomatocytosis associated with dyserythropoiesis. Haematologica. 2009;94:1049-59.

Iolascon A, Russo R, Esposito MR, Piscopo C, Asci R, De Falco L, Di Noce F. Congenital dyserythropoietic anaemias: new acquisitions. Blood Transfus. 2011;9:278-80.

Iolascon A, King MJ, Robertson S, Avvisati RA, Vitiello F, Asci R, Scoppettuolo MN, Delaunay J. A genomic deletion causes truncation of α-spectrin and ellipto-poikilocytosis. Blood Cells Mol Dis. 2011;46:195-200.

Iolascon A, Esposito MR, Russo R. Clinical aspects and pathogenesis of congenital dyserythropoietic anemias: from morphology to molecular approach. Haematologica. 2012;97:1786-94.

Iolascon A, Heimpel H, Wahlin A, Tamary H. Congenital dyserythropoietic anemias: molecular insights and diagnostic approach. Blood. 2013;122:2162-6.

Iolascon A, Andolfo I, Russo R. Red cells in post-genomic era: impact of personalized medicine in the treatment of anemias. Haematologica. 2015;100:3-6.

Jaffray JA, Mitchell WB, Gnanapragasam MN, Seshan SV, Guo X, Westhoff CM, Bieker JJ, Manwani D. Erythroid transcription factor EKLF/KLF1 mutation causing congenital dyserythropoietic anemia type IV in a patient of Taiwanese origin: review of all reported cases and development of a clinical diagnostic paradigm. Blood Cells Mol Dis. 2013;51:71-5.

Jamuar SS, Tan EC. Clinical application of next-generation sequencing for Mendelian diseases. Hum Genomics. 2015;9:10.

Kaneko K, Furuyama K, Fujiwara T, Kobayashi R, Ishida H, Harigae H, Shibahara S. Identification of a novel erythroid-specific enhancer for the ALAS2 gene and its loss-of-function mutation which is associated with congenital sideroblastic anemia. Haematologica. 2014;99:252-61.

Khoriaty R, Vasievich MP, Jones M, Everett L, Chase J, Tao J, Siemieniak D, Zhang B, Maillard I, Ginsburg D. Absence of a red blood cell phenotype in mice with hematopoietic deficiency of SEC23B. Mol Cell Biol. 2014;34:3721-34.

King MJ, Garçon 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.

Kobayashi M, Nishikawa K, Yamamoto M. Hematopoietic regulatory domain of gata1 gene is positively regulated by GATA1 protein in zebrafish embryos. Development. 2001;128:2341-50.

Kratz CP, Niemeyer CM, Karow A, Volz-Fleckenstein M, Schmitt-Gräff A, Strahm B. Congenital transfusion-dependent anemia and thrombocytopenia with myelodysplasia due to a recurrent GATA1(G208R) germline mutation. Leukemia. 2008;22:432-4.

Liljeholm M, Irvine AF, Vikberg AL, Norberg A, Month S, Sandström H, Wahlin A, Mishima M, Golovleva I. Congenital dyserythropoietic anemia type III (CDA III) is caused by a mutation in kinesin family member, KIF23. Blood. 2013;121:4791-9.

Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. Nat Methods. 2010;7:111-8.

Nichols KE, Crispino JD, Poncz M, White JG, Orkin SH, Maris JM, Weiss MJ. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat Genet. 2000;24:266-70.

Persico M, Russo R, Persico E, Svelto M, Spano D, Andolfo I, La Mura V, Capasso M, Tiribelli C, Torella R, Iolascon A. SOCS3 and IRS-1 gene expression differs between genotype 1 and genotype 2 hepatitis C virus-infected HepG2 cells. Clin Chem Lab Med. 2009;47:1217-25.

Rapetti-Mauss R, Lacoste C, Picard V, Guitton C, Lombard E, Loosveld M, Nivaggioni V, Dasilva N, Salgado D, Desvignes JP, Béroud C, Viout P, Bernard M, Soriani O, Vinti H, Lacroze V, Feneant-Thibault M, Thuret I, Guizouarn H, Badens C. A mutation in the Gardos channel is associated with hereditary xerocytosis. Blood. 2015;126:1273-80.

Renella R, Roberts NA, Brown JM, De Gobbi M, Bird LE, Hassanali T, Sharpe JA, Sloane-Stanley J, Ferguson DJ, Cordell J, Buckle VJ, Higgs DR, Wood WG. Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1{alpha} localization in erythroblasts. Blood. 2011;117:6928-38.

Russo R, Esposito MR, Asci R, Gambale A, Perrotta S, Ramenghi U, Forni GL, Uygun V, Delaunay J, Iolascon A. Mutational spectrum in congenital dyserythropoietic anemia type II: identification of 19 novel variants in SEC23B gene. Am J Hematol. 2010;85:915-20.

Russo R, Gambale A, Esposito MR, Serra ML, Troiano A, De Maggio I, Capasso M, Luzzatto L, Delaunay J, Tamary H, Iolascon A. Two founder mutations in the SEC23B gene account for the relatively high frequency of CDA II in the Italian population. Am J Hematol. 2011;86:727-32.

Russo R, Esposito MR, Iolascon A. Inherited hematological disorders due to defects in coat protein (COP)II complex. Am J Hematol. 2013a;88:135-40.

Russo R, Langella C, Esposito MR, Gambale A, Vitiello F, Vallefuoco F, Ek T, Yang E, Iolascon A. Hypomorphic mutations of SEC23B gene account for mild phenotypes of congenital dyserythropoietic anemia type II. Blood Cells Mol Dis. 2013b;51:17-21.

Russo R, Gambale A, Langella C, Andolfo I, Unal S, Iolascon A. Retrospective cohort study of 205 cases with congenital dyserythropoietic anemia type II: definition of clinical and molecular spectrum and identification of new diagnostic scores. Am J Hematol. 2014;89:E169-75.

Savarese M, Di Fruscio G, Mutarelli M, Torella A, Magri F, Santorelli FM, Comi GP, Bruno C, Nigro V. MotorPlex provides accurate variant detection across large muscle genes both in single myopathic patients and in pools of DNA samples. Acta Neuropathol Commun. 2014;2:100.

Schwarz K, Iolascon A, Verissimo F, Trede NS, Horsley W, Chen W, Paw BH, Hopfner KP, Holzmann K, Russo R, Esposito MR, Spano D, De Falco L, Heinrich K, Joggerst B, Rojewski MT, Perrotta S, Denecke J, Pannicke U, Delaunay J, Pepperkok R, Heimpel H. Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat Genet. 2009;41:936-40.

Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol. 2008;26:1135-45.

Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, van Spaendonck-Zwarts KY, van Tintelen JP, Sijmons RH, Jongbloed JD, Sinke RJ. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. Hum Mutat. 2013;34:1035-42.

Skop AR, Liu H, Yates J 3rd, Meyer BJ, Heald R. Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. Science. 2004;305:61-6.

Stewart GW, Amess JA, Eber SW, Kingswood C, Lane PA, Smith BD, Mentzer WC. Thrombo-embolic disease after splenectomy for hereditary stomatocytosis. Br J Haematol. 1996;93:303-10.

Stewart GW. Hemolytic disease due to membrane ion channel disorders. Curr Opin Hematol 2004;11:244-50.

Surinya KH, Cox TC, May BK. Transcriptional regulation of the human erythroid 5-aminolevulinate synthase gene. Identification of promoter elements and role of regulatory proteins. J Biol Chem. 1997;272:26585-94.

Tamary H, Offret H, Dgany O, Foliguet B, Wickramasinghe SN, Krasnov T, Rumilly F, Goujard C, Fénéant-Thibault M, Cynober T, Delaunay J. Congenital dyserythropoietic anaemia, type I, in a Caucasian patient with retinal angioid streaks (homozygous Arg1042Trp mutation in codanin-1). Eur J Haematol. 2008;80:271-4.

Tao J, Zhu M, Wang H, Afelik S, Vasievich MP, Chen XW, Zhu G, Jensen J, Ginsburg D, Zhang B. SEC23B is required for the maintenance of murine professional secretory tissues. Proc Natl Acad Sci U S A. 2012;109:E2001-9.

Zarychanski R, Schulz VP, Houston BL, Maksimova Y, Houston DS, Smith B, Rinehart J, Gallagher PG. Mutations in the mechanotransduction protein PIEZO1 are associated with hereditary xerocytosis. Blood. 2012;120:1908-15.

Zon LI, Youssoufian H, Mather C, Lodish HF, Orkin SH. Activation of the erythropoietin receptor promoter by transcription factor GATA-1. Proc Natl Acad Sci U S 32A. 1991;88:10638-41.

EXPERT REVIEW OF HEMATOLOGY, 2016
http://dx.doi.org/10.1586/17474086.2016.1131608

Taylor & Francis
Taylor & Francis Group

REVIEW

Diagnosis and management of congenital dyserythropoietic anemias

Antonella Gambale[a,b], Achille Iolascon[a,b], Immacolata Andolfo[a,b] and Roberta Russo[a,b]

[a]Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli, Italy; [b]CEINGE Biotecnologie Avanzate, Napoli, Italy

ABSTRACT

Congenital dyserythropoietic anemias (CDAs) are inherited disorders hallmarked by chronic hyporegenerative anemia, relative reticulocytopenia, hemolytic component and iron overload. They represent a subtype of the inherited bone marrow failure syndromes, characterized by impaired differentiation and proliferation of the erythroid lineage. Three classical types were defined by marrow morphology, even if the most recent classification recognized six different genetic types. The pathomechanisms of CDAs are different, but all seem to involve the regulation of DNA replication and cell division. CDAs are often misdiagnosed, since either morphological abnormalities or clinical features can be commonly identified in other clinically-related anemias. However, differential diagnosis is essential for guiding both follow up and management of the patients.

ARTICLE HISTORY
Received 20 October 2015
Accepted 10 December 2015
Published online
6 January 2016

KEYWORDS
Inherited anemia; ineffective erythropoiesis; molecular genetics; differential diagnosis; follow up; patient management

Introduction

The term dyserythropoiesis refers to a condition of abnormal erythropoiesis affecting the differentiation and proliferation pathways of the erythroid lineage with a consequent defective production of red blood cells (RBCs).[1] Dyserythropoietic anemias can be divided into primary and secondary forms, and both inherited and acquired types can occur. Among these different conditions, the congenital dyserythropoietic anemias (CDAs) are hereditary diseases that embrace a highly heterogeneous set of rare or very rare anemias that result from various kinds of abnormalities during late stages of erythropoiesis. They are counted as subtypes of inherited bone marrow failure syndromes (IBMFS), characterized by morphological abnormalities of erythroblasts in the bone marrow (BM) and ineffective erythropoiesis as predominant mechanism of anemia, accompanied by a hemolytic component.[2]

Dyserythropoiesis appears to be a morphological feature common to several conditions, and this could account for the difficulties in diagnosis of CDAs. However, the specific morphological alterations of the erythroid precursors justify the heterogeneity of these disorders. Indeed, the three classical types of CDAs (types I, II and III) are defined on the basis of BM morphology.[2] Inclusions of additional CDAs, the so-called CDA variants, despite remarkable morphological studies, gradually led to overlapping entities and imposed a limitation on classification. In spite of these difficulties, morphological classification is still widely used in clinical practice. Nevertheless, the identification of the causative genes of the most common forms among CDAs in the last two decades represented an evident advantage for reclassifying these disorders, as well as in understanding their pathogenesis. Moreover, uncovering the molecular basis of CDAs helped to unravel novel aspects of the molecular biology of erythropoiesis. From the genetic standpoint, six different types of CDAs are included in the Online Mendelian Inheritance in Man (OMIM) compendium of human genes and genetic phenotypes so far (Table 1). However, this list is bound to extend. Indeed, as described later in this review, the discovery of new causative genes is a continuous evolving process thanks to the development of high-throughput technologies, such as next-generation sequencing (NGS). The identification of genetic variations underlying hereditary disorders marked the opening of a new era for genetic and clinical research. Indeed, beyond, obtaining definitive diagnosis and planning patient management, knowledge of the genetic basis of these disorders is crucial in estimating their prevalence and geographical distribution. According to the estimation by Heimpel and colleagues in 2010, the prevalence of CDAs varies widely among European regions, with minimal values of 0.04 cases/million in North Europe and the highest

CONTACT Achille Iolascon achille.iolascon@unina.it CEINGE – Biotecnologie Avanzate, Via Gaetano Salvatore 486, 80145 Naples, Italy

RESEARCH ARTICLE

Novel Gardos channel mutations linked to dehydrated hereditary stomatocytosis (xerocytosis)

Immacolata Andolfo,[1,2] Roberta Russo,[1,2] Francesco Manna,[1,2] Boris E. Shmukler,[3,4] Antonella Gambale,[1,2] Giuseppina Vitiello,[2,5] Gianluca De Rosa,[1,2] Carlo Brugnara,[6] Seth L. Alper,[3] L. Michael Snyder,[7,8] and Achille Iolascon[1,2]*

Dehydrated hereditary stomatocytosis (DHSt) is an autosomal dominant congenital hemolytic anemia with moderate splenomegaly and often compensated hemolysis. Affected red cells are characterized by a nonspecific cation leak of the red cell membrane, reflected in elevated sodium content, decreased potassium content, elevated MCHC and MCV, and decreased osmotic fragility. The majority of symptomatic DHSt cases reported to date have been associated with gain-of-function mutations in the mechanosensitive cation channel gene, *PIEZO1*. A recent study has identified two families with DHSt associated with a single mutation in the *KCNN4* gene encoding the Gardos channel (KCa3.1), the erythroid Ca^{2+}-sensitive K^+ channel of intermediate conductance, also expressed in many other cell types. We present here, in the second report of DHSt associated with *KCNN4* mutations, two previously undiagnosed DHSt families. Family NA exhibited the same *de novo* missense mutation as that recently described, suggesting a hot spot codon for DHSt mutations. Family WO carried a novel, inherited missense mutation in the ion transport domain of the channel. The patients' mild hemolytic anemia did not improve post-splenectomy, but splenectomy led to no serious thromboembolic events. We further characterized the expression of *KCNN4* in the mutated patients and during erythroid differentiation of CD34+ cells and K562 cells. We also analyzed KCNN4 expression during mouse embryonic development.

Am. J. Hematol. 90:921–926, 2015.

Introduction

Dehydrated hereditary stomatocytosis (DHSt), also known as hereditary xerocytosis (OMIM 194380), is an autosomal dominant congenital hemolytic anemia associated with a monovalent cation leak. DHSt consists of a usually compensated hemolysis, associated with moderate splenomegaly [1,2]. Blood smears show variable numbers of stomatocytes. The reticulocyte count is elevated, and red cell mean corpuscular volume (MCV) is increased. DHSt is associated with silent-to-mild hemolysis, pseudohyperkalemia (red cell K^+ loss upon storage at room temperature or in the cold), and perinatal edema [3]. DHSt red blood cells exhibit decreased K^+ content and increased Na^+ content, usually accompanied by increased mean corpuscular hemoglobin concentration (MCHC). The cation leak of DHSt red cells resembles that of control RBC in its temperature dependence, but is of greater magnitude at all temperatures [3]. The definitive diagnosis of DHSt is made by osmotic gradient ektacytometry, which shows a leftward shift of the bell-shaped curve [4,5]. Occasionally associated hepatosiderosis, beyond that expected from the mild hemolytic state, suggests a strong tendency to iron overload [6,7]. Unlike hereditary spherocytosis, in which splenectomy can be beneficial, splenectomy in DHSt is contraindicated due to increased risk of thromboembolic complications [8,9].

The causative gene of the condition was recently identified in the *FAM38A* gene encoding the mechanosensitive cation channel, PIEZO1 [10,11]. Several functional characterizations of identified PIEZO1 mutations in DHSt families have uniformly demonstrated gain-of-function properties consistent with the increased net ion fluxes leading to DHSt [7,10–13].

Rapetti-Mauss and colleagues recently identified in two DHSt families a single mutation in the *KCNN4* gene encoding the Gardos channel [14]. The Gardos channel/KCNN4 is a widely expressed Ca^{2+}-dependent K^+ channel of intermediate conductance that mediates the major K^+ conductance of erythrocytes [15]. Inhibition of the Gardos channel in sickle disease patients reduced RBC dehydration and hemolysis, and increased blood hemoglobin levels [16–19].

Additional Supporting Information may be found in the online version of this article.

[1]Department Of Molecular Medicine And Medical Biotechnologies, "Federico II" University Of Naples, Naples, Italy; [2]Biotecnologie Avanzate, CEINGE, Naples, Italy; [3]Renal Division And Vascular Biology Research Center, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; [4]Department Of Medicine, Harvard Medical School, Boston, Massachusetts; [5]Medical Genetics Unit, Policlinico Tor Vergata University Hospital, Viale Oxford, Rome, Italy; [6]Department Of Laboratory Medicine, Boston Children's Hospital And Department Of Pathology, Harvard Medical School, Boston, Massachusetts; [7]Dept Of Hospital Laboratories, University Of Massachusetts Medical Center, Worcester, MA; [8]Quest Diagnositics, LLC MA, Marlborough, Massachusetts

Conflict of interest: Nothing to report.
***Correspondence to:** Achille Iolascon, MD, PhD, CEINGE, Biotecnologie Avanzate, Via Gaetano Salvatore, Naples 486 80145, Italy. E-mail: achille.iolascon@unina.it
Contract grant sponsor: Italian Ministero dell'Università e della Ricerca [(MIUR) to A.I.]; Contract grant sponsor: Doris Duke Charitable Trust; Contract grant sponsor: NIH; Contract grant number: R01 HL077765 (to S.L.A.).
Contract grant sponsor: Dept. of Hospital Laboratories, UMass Memorial Medical Center, Worcester, MA (to L.M.S.).
Received for publication: 8 July 2015; **Accepted:** 8 July 2015
Am. J. Hematol. 90:921–926, 2015.
Published online: 15 July 2015 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/ajh.24117

European Journal of Haematology

European Journal of Haematology 94 (491–497)

ORIGINAL ARTICLE

Congenital erythropoietic porphyria linked to GATA1-R216W mutation: challenges for diagnosis

Elena Di Pierro[1,*], Roberta Russo[2,3,*], Zeynep Karakas[4,*], Valentina Brancaleoni[1], Antonella Gambale[2,3], Ismail Kurt[5], S. Stuart Winter[6], Francesca Granata[1], David Rodriguez Czuchlewski[7], Concetta Langella[2,3], Achille Iolascon[2,3], Maria Domenica Cappellini[1,8]

[1]Fondazione IRCCS "Cà-Granda" Ospedale Maggiore Policlinico, U.O. di Medicina Interna, Milan; [2]Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Naples; [3]CEINGE – Biotecnologie Avanzate, Naples, Italy; [4]Division of Hematology/Oncology, Department of Pediatrics, Istanbul Medical Faculty, Istanbul University, Istanbul; [5]Department of Biochemistry and Clinical Biochemistry - Laboratory of Porphiria, Gulhane Military Academy of Medicine, Ankara, Turkey; [6]Department of Pediatrics, University of New Mexico, Albuquerque, NM; [7]Department of Pathology, University of New Mexico, Albuquerque, NM, USA; [8]Dipartimento di Scienze Cliniche e di Comunità, Università degli Studi di Milano, Milan, Italy

Abstract

Congenital erythropoietic porphyria (CEP) is a rare genetic disease that is characterized by a severe cutaneous photosensitivity causing unrecoverable deformities, chronic hemolytic anemia requiring blood transfusion program, and by fatal systemic complications. A correct and early diagnosis is required to develop a management plan that is appropriate to the patient's needs. Recently only one case of X-linked CEP had been reported, describing the trans-acting GATA1-R216W mutation. Here, we have characterized two novel X-linked CEP patients, both with misleading hematological phenotypes that include dyserythropoietic anemia, thrombocytopenia, and hereditary persistence of fetal hemoglobin. We compare the previously reported case to ours and propose a diagnostic paradigm for this variant of CEP. Finally, a correlation between phenotype variability and the presence of modifier mutations in loci related to disease-causing gene is described.

Key words congenital erythropoietic porphyria; UROS, GATA1, and SEC23B mutations; neonatal hemolytic anemia

Correspondence Di Pierro Elena, PhD, Fondazione IRCCS "Cà-Granda" Ospedale Maggiore Policlinico-Via F. Sforza 35, 20122 Milan, Italy. Tel: +390255033363; Fax: +390250320296; e-mail: elena.dipierro@unimi.it

*These authors contributed equally to this paper.

Accepted for publication 17 September 2014

doi:10.1111/ejh.12452

Congenital erythropoietic porphyria (CEP) or Gunther's disease is a rare disorder characterized by markedly decreased uroporphyrinogen III synthase (UROS) activity in erythrocytes and significantly increased levels of urinary and plasmatic uroporphyrin I and coproporphyrin I isomers. The clinical manifestations of CEP are heterogeneous, above all depending on the amount of residual UROS activity and ranging from non-immune hydrops fetalis to milder late-onset forms of dyserythropoiesis (1). Deficit of UROS, the fourth enzyme in the heme biosynthetic pathway, causes an excess of uroporphyrin accumulation primarily in bone marrow and consequently in peripheral blood leading to dyserythropoiesis, hemolytic anemia, splenomegaly. The accumulation of uroporphyrin in other tissues gives rise to skin photosensitivity, ulcers, erosions, erythrodontia, osteodystrophia, combining osteolysis and osteoporosis. Secondary infections of cutaneous lesions may induce to scarring, deformities, and disfigurement of the light-exposed parts of the body such as hands, ears, nose, and eyelids (2).

The diagnosis is confirmed most commonly by the presence of biallelic mutations in the *UROS* gene, inherited in an autosomal recessive manner, which results in strikingly deficient activity of the enzyme. To date, 38 mutations have been identified in unrelated patients with CEP. The most common mutation is the missense C73R, occurring in approximately one-third of the CEP alleles, whereas the

RESEARCH ARTICLE

Retrospective cohort study of 205 cases with congenital dyserythropoietic anemia type II: definition of clinical and molecular spectrum and identification of new diagnostic scores

Roberta Russo,[1,2] Antonella Gambale,[1,2] Concetta Langella,[1,2] Immacolata Andolfo,[1,2] Sule Unal,[3] and Achille Iolascon[1,2]*

Congenital Dyserythropoietic Anemia II (CDA II) is a rare hyporegenerative anemia of variable degree, whose causative gene is *SEC23B*. More than 60 causative mutations in 142 independent pedigrees have been described so far. However, the prevalence of the CDA II is probably underestimated, since its clinical spectrum was not yet well-defined and thus it is often misdiagnosed with more frequent clinically-related anemias. This study represents the first meta-analysis on clinical and molecular spectrum of CDA II from the largest cohort of cases ever described. We characterized 41 new cases and 18 mutations not yet associated to CDA II, thus expanding the global series to 205 cases (172 unrelated) and the total number of causative variants to 84. The 68.3% of patients are included in our International Registry of CDA II (Napoli, Italy). A genotype–phenotype correlation in three genotypic groups of patients was assessed. To quantify the degree of severity in each patient, a method based on ranking score was performed. We introduced a clinical index to easily discriminate patients with a well-compensated hemolytic anemia from those with ineffective erythropoiesis. Finally, the worldwide geographical distribution of *SEC23B* alleles highlighted the presence of multiple founder effects in different areas of the world.

Am. J. Hematol. 00:000–000, 2014.

Introduction

Congenital Dyserythropoietic Anemia type II (CDA II) belongs to a subtype of bone marrow failure syndromes characterized by monolineage involvement and morphological abnormalities in erythroid precursor cells.

Congenital normocytic anemia of variable degree with a reticulocytosis not corresponding to the degree of anemia (ineffective erythropoiesis) is the main criteria to diagnose CDA II. Because of hemolytic symptoms, it is frequently accompanied with jaundice and splenomegaly [1].

The CDA II BM is hypercellular, due to an exclusive and pronounced increase of erythroblasts, with erythropoietic/granulopoietic ratios of 4.2 (1.7–10.0) (normal reference values 3.6, 1.7–5.4) [2]. Distinct morphological abnormalities of the erythroblasts are observable as a consequence of the impairment of differentiation and proliferation pathways of the erythroid lineage. Indeed, BM light microscopy reveals more than 10% mature binucleated erythroblasts, while on electron microscopy vesicles loaded with proteins of endoplasmic reticulum (ER) are found beneath the plasma membrane [2].

Analysis of red cell membrane proteins by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), identifying glycosylation abnormalities with fast moving band 3 (anion exchanger 1) and band 4.5 (glucose transporter 1), is a highly sensitive and specific diagnostic tool, even if it is limited by the availability of specialized laboratories.

CDA II belongs to COPII-related human genetic disorders, since it is due to mutations in *SEC23B* gene (chr 20p11.23) [3,4], which encodes the cytoplasmic coat protein (COP)II component SEC23B, involved in the secretory pathway of eukaryotic cells. This multisubunit complex mediates accumulation of secretory cargo, deformation of the membrane and anterograde transport of correctly folded cargo for budding from the ER towards the Golgi apparatus [5].

The clinical and biochemical suspect of CDA II is confirmed most commonly by the presence of biallelic mutations in the *SEC23B* gene, inherited in an autosomal recessive manner. The prevalence of CDA II in Europe shows substantial variations between different countries [6], but it exhibited the highest value in Italy, where a founder effect has been demonstrated [7]. To date, 66 different mutations in 142 unrelated cases have been reported [8–17].

Additional Supporting Information may be found in the online version of this article.

[1]Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Napoli, Italy; [2]CEINGE Biotecnologie Avanzate, Napoli, Italy; [3]Division of Pediatric Hematology, Hacettepe University, Ankara, Turkey

Conflicts of interest: Nothing to report
***Correspondence to:** Achille Iolascon, Biotecnologie Avanzate, Via Gaetano Salvatore 486, 80145 Naples, Italy, E-mail: achille.iolascon@unina.it
Contract grant sponsor: Italian Ministero dell'Università e della Ricerca; Contract grant number: MUR-PS 35–126/Ind.
Contract grant sponsor: Regione Campania; Contract grant number: DGRC2362/07.
Contract grant sponsor: EU Contract by Italian Telethon Foundation, Rome, Italy; Contract grant numbers: LSHM-CT-2006–037296; GGP 09044.
Received for publication: 2 July 2014; **Accepted:** 3 July 2014
Am. J. Hematol. 00:00–00, 2014.
Published online: 9 July 2014 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/ajh.23800

Pediatr Transplantation 2014: 18: E130–E133

Pediatric Transplantation
DOI: 10.1111/petr.12254

Successful hematopoietic stem cell transplantation in a patient with congenital dyserythropoietic anemia type II

Unal S, Russo R, Gumruk F, Kuskonmaz B, Cetin M, Sayli T, Tavil B, Langella C, Iolascon A, Cetinkaya DU. Successful hematopoietic stem cell transplantation in a patient with congenital dyserythropoietic anemia type II.

Abstract: CDA are a group of inherited, rare diseases that are characterized by dyserythropoiesis and ineffective erythropoiesis associated with transfusion dependency in approximately 10% of cases. For these latter patients, the only curative treatment is HSCT. There are very limited data on HSCT experience in this rare disease. Herein, we report a five-yr six-month-old girl with compound heterozygous mutations in *SEC23B* gene, who was diagnosed to have CDA type II and underwent successful HSCT from her matched sibling donor.

Sule Unal[1], Roberta Russo[2,3], Fatma Gumruk[1], Baris Kuskonmaz[1], Mualla Cetin[1], Tulin Sayli[4], Betul Tavil[1], Concetta Langella[2,3], Achille Iolascon[2,3] and Duygu Uckan Cetinkaya[1]

[1]Division of Pediatric Hematology, Hacettepe University, Ankara, Turkey, [2]Department of Molecular Medicine and Medical Biotechnologies, University Federico II, Naples, Italy, [3]CEINGE Advanced Biotechnologies, Naples, Italy, [4]Ankara Pediatric Hematology and Oncology Research Hospital, Ankara, Turkey

Key words: transplantation – congenital dyserythropoietic anemias – iron – *SEC23B*

Sule Unal, Division of Pediatric Hematology, Hacettepe University, Ankara 06100, Turkey
Tel.: +90 312 305 1170
Fax: +90 312 311 2398
E-mail: suleunal@hacettepe.edu.tr

Accepted for publication 24 February 2014

CDA are a group of rare heterogenous disorders characterized by dyserythropoiesis, ineffective erythropoiesis, iron overload, and specific light and electron microscopy findings of nucleated erythroid precursors (1–3). Patients usually present with anemia, jaundice, splenomegaly, low reticulocyte count despite erythroid hyperactivity (3). There are broadly three main types of CDA (CDA I, II, and III), due to mutations in *CDAN1, SEC23B,* and *KIF23,* respectively (3–7). However, there are additional CDA variants that do not fit to any three classical types, such as CDA patients with *KLF1* and *GATA1* mutations. To date, 157 cases from 137 different CDA II families with *SEC23B* mutations were molecularly analyzed (3, 8). Type II patients are characterized by erythroid hyperactivity in bone marrow with no megaloblastic changes, in addition to high numbers of binucleated normoblasts with occasional multinucleated erythroid precursors (1). The nucleated erythroid precursors of patients with CDA type II may exhibit double membrane appearance under electron microscopy (9). There are patients with CDA who have mild-to-moderate anemia who require no regular transfusions. On the other hand, treatment alternatives include erythrocyte transfusions for patients with severe anemia, iron chelation to decrease complications related to transfusional iron overload, interferon alpha for some of type I and splenectomy for some of the type II patients (10, 11). However, the only curative treatment for patients with CDA is HSCT (12, 13). There is scarce data in the reported literature on the use of HSCT as a therapeutic and curative option in patients with CDA (2, 12, 14–17). Herein, we report a five-yr six-month-old girl with CDA II who had compound heterozygous mutations in *SEC23B* gene and was successfully transplanted from a matched sibling donor.

Abbreviations: ATG, antithymocyte globulin; CDA, congenital dyserythropoietic anemias; CMV, cytomegalovirus; CyA, cyclosporin A; gDNA, genomic DNA; GVHD, graft-versus-host disease; HSCT, hematopoietic stem cell transplantation; LDH, lactate dehydrogenase; MCV, mean corpuscular volume; RDW, red cell distribution width; VOD, veno-occlusive disease.

RED CELLS, IRON, AND ERYTHROPOIESIS

Multiple clinical forms of dehydrated hereditary stomatocytosis arise from mutations in *PIEZO1*

Immacolata Andolfo,[1,2] Seth L. Alper,[3] Lucia De Franceschi,[4] Carla Auriemma,[1,2] Roberta Russo,[1,2] Luigia De Falco,[1,2] Fara Vallefuoco,[1,2] Maria Rosaria Esposito,[1,2] David H. Vandorpe,[3] Boris E. Shmukler,[3] Rupa Narayan,[5] Donatella Montanaro,[1] Maria D'Armiento,[6] Annalisa Vetro,[7] Ivan Limongelli,[7] Orsetta Zuffardi,[7] Bertil E. Glader,[5] Stanley L. Schrier,[8] Carlo Brugnara,[9] Gordon W. Stewart,[10] Jean Delaunay,[11] and Achille Iolascon[1,2]

[1]Department of Molecular Medicine and Medical Biotechnologies, Federico II University of Naples, Naples, Italy; [2]Centro di Ingegneria Genetica (CEINGE), Biotecnologie Avanzate, Naples, Italy; [3]Renal Division and Molecular and Vascular Medicine Division, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, Boston, MA; [4]Department of Medicine, University of Verona, Verona, Italy; [5]Division of Hematology-Oncology, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA; [6]Dipartimento di Scienze biomediche avanzate, Federico II University of Naples, Naples, Italy; [7]Department of Molecular Medicine, University of Pavia, Pavia, Italy; [8]Division of Hematology, Department of Medicine, Stanford University School of Medicine, Stanford, CA; [9]Department of Laboratory Medicine, Boston Children's Hospital and Department of Pathology, Harvard Medical School, Boston, MA; [10]Division of Medicine, University College London, London, United Kingdom; and [11]Institut National de la Sante et de la Recherche Médicale (INSERM), Faculté de Médecine Paris-Sud, Université Paris-Sud, Paris, France

Key Points

- Dehydrated hereditary stomatocytosis is characterized by abnormal RBC morphology but may involve pseudohyperkalemia and perinatal edema.
- This syndrome is associated with germline mutations in *PIEZO1*, encoding a transmembrane protein that induces mechanosensitive currents.

Autosomal dominant dehydrated hereditary stomatocytosis (DHSt) usually presents as a compensated hemolytic anemia with macrocytosis and abnormally shaped red blood cells (RBCs). DHSt is part of a pleiotropic syndrome that may also exhibit pseudohyperkalemia and perinatal edema. We identified *PIEZO1* as the disease gene for pleiotropic DHSt in a large kindred by exome sequencing analysis within the previously mapped 16q23-q24 interval. In 26 affected individuals among 7 multigenerational DHSt families with the pleiotropic syndrome, 11 heterozygous *PIEZO1* missense mutations cosegregated with disease. PIEZO1 is expressed in the plasma membranes of RBCs and its messenger RNA, and protein levels increase during in vitro erythroid differentiation of $CD34^+$ cells. PIEZO1 is also expressed in liver and bone marrow during human and mouse development. We suggest for the first time a correlation between a *PIEZO1* mutation and perinatal edema. DHSt patient red cells with the R2456H mutation exhibit increased ion-channel activity. Functional studies of *PIEZO1* mutant R2488Q expressed in *Xenopus* oocytes demonstrated changes in ion-channel activity consistent with the altered cation content of DHSt patient red cells. Our findings provide direct evidence that R2456H and R2488Q mutations in *PIEZO1* alter mechanosensitive channel regulation, leading to increased cation transport in erythroid cells. (*Blood*. 2013; 121(19):3925-3935)

Introduction

Dehydrated hereditary stomatocytosis (DHSt), also known as hereditary xerocytosis (Online Mendelian Inheritance in Man [OMIM] 194380), is an autosomal dominant congenital hemolytic anemia associated with a monovalent cation leak. DHSt consists of a usually compensated hemolysis, associated with moderate splenomegaly.[1,2] Blood smears show variable numbers of stomatocytes, sometimes rare and ill-formed, and likely to be overlooked. The reticulocyte count is elevated, and red cell mean corpuscular volume (MCV) is slightly increased. DHSt red blood cells (RBCs) exhibit decreased intraerythrocytic K^+ content and increased intraerythrocytic Na^+ content, usually accompanied by increased mean corpuscolar hemoglobin (Hb) concentration. The cation leak of DHSt red cells resembles that of control RBCs in its temperature dependence, but is of greater magnitude at all temperatures.[3] The definitive diagnosis of DHSt is ascertained by osmotic gradient ektacytometry, which shows a leftward shift of the bell-shaped curve.[4,5] Occasionally, associated hepatosiderosis beyond that expected from the mild hemolytic state suggests a strong tendency to iron overload.[6] Unlike hereditary spherocytosis, in which splenectomy can be beneficial, splenectomy in DHSt is contraindicated due to increased risk of thromboembolic complications.[7,8]

DHSt can present as an isolated erythroid phenotype or as associated with pseudohyperkalemia, with pre- and/or perinatal edema, or with both pseudohyperkalemia and effusions. The pre- and/or

Submitted February 1, 2013; accepted March 4, 2013. Prepublished online as *Blood* First Edition paper, March 11, 2013; DOI 10.1182/blood-2013-02-482489.

The online version of this article contains a data supplement.

There is an Inside *Blood* commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Blood Cells, Molecules and Diseases 51 (2013) 17–21

Contents lists available at SciVerse ScienceDirect

Blood Cells, Molecules and Diseases

journal homepage: www.elsevier.com/locate/bcmd

Hypomorphic mutations of *SEC23B* gene account for mild phenotypes of congenital dyserythropoietic anemia type II

Roberta Russo [a,b], Concetta Langella [a,b], Maria Rosaria Esposito [b], Antonella Gambale [b], Francesco Vitiello [a,b], Fara Vallefuoco [b], Torben Ek [c], Elizabeth Yang [d], Achille Iolascon [a,b,*]

[a] *Department of Molecular Medicine and Medical Biotechnologies, University Federico II of Naples, Naples, Italy*
[b] *CEINGE Biotecnologie Avanzate, Naples, Italy*
[c] *Department of Paediatrics, County Hospital, Halmstad, Sweden*
[d] *Center for Cancer and Blood Disorders of Northern Virginia, Children's National Medical Center, Falls Church, VA, USA*

ARTICLE INFO

Article history:
Submitted 31 January 2013
Available online 1 March 2013

(Communicated by P. Gallagher, M.D.,
1 February 2013)

Keywords:
CDA II
SEC23B
Hypomorphic mutations
Genotype–phenotype correlation

ABSTRACT

Congenital dyserythropoietic anemia type II, a recessive disorder of erythroid differentiation, is due to mutations in SEC23B, a component of the core trafficking machinery COPII. In no case homozygosity or compound heterozygosity for nonsense mutation(s) was found. This study represents the first description of molecular mechanisms underlying *SEC23B* hypomorphic genotypes by the analysis of five novel mutations. Our findings suggest that reduction of *SEC23B* gene expression is not associated with CDA II severe clinical presentation; conversely, the combination of a hypomorphic allele with one functionally altered results in more severe phenotypes. We propose a mechanism of compensation SEC23A-mediated which justifies these observations.

Introduction

Congenital dyserythropoietic anemia type II (CDA II, OMIM 224100) is a genetic hyporegenerative anemia characterized by ineffective erythropoiesis and distinct morphological abnormalities of the erythroblasts in the bone marrow (BM). Anemia of variable degree, jaundice and splenomegaly are common clinical findings [1]. This condition belongs to COPII-related human genetic disorders [2]. It is due to mutations in SEC23B (chr 20p11.23), a component of COPII complex, the core trafficking machinery of the endoplasmic reticulum-Golgi [3]. Approximately 60 different causative mutations have been described, localized along the entire coding sequence of the gene [1,4–6]. The most frequent are nucleotide substitutions (75% missense/nonsense), whereas frameshift and splicing mutations were observed in 15% and 10% respectively. The vast majority of patients have two mutations (in the homozygous or compound heterozygous state), according to the pattern of autosomal recessive inheritance. In no case homozygosity or compound heterozygosity for two nonsense mutations was found, a situation likely to be lethal. However, few cases with two hypomorphic mutations have been described so far [4,5].

Here we characterize three novel CDA II cases, two of them with fully hypomorphic genotype. We demonstrated a compensatory mechanism SEC23A-mediated of SEC23B hypo-expressed alleles.

Material and methods

Patients and mutational screening

Diagnosis of CDA II was based on history, clinical findings, laboratory data, morphological analysis of aspirated bone marrow and whenever possible on evidence of hypoglycosylated band 3 by SDS-PAGE. Samples were obtained after informed consent for the studies, according to the Declaration of Helsinki. Whenever possible, relatives were investigated. Genomic DNA and mutational screening were performed as previously described [4].

In silico and ex vivo analyses on mRNA

Prediction analyses for splice site mutations were performed by web server tools, splice site prediction by neural network (http://www.fruitfly.org/seq_tools/splice.html) and human splicing finder (http://www.umd.be/HSF/) (Table 2).

RNA isolation from peripheral blood mononuclear cells (PBMCs), cDNA preparation and quantitative real-time (qRT)-PCR were performed as described [7]. Relative gene expression was calculated by using the $2^{-\Delta Ct}$ method, while the mean fold change $= 2^{-(\text{average } \Delta\Delta Ct)}$ was

* Corresponding author at: CEINGE — Biotecnologie Avanzate, Via Gaetano Salvatore 486, 80145 Naples, Italy.
E-mail address: achille.iolascon@unina.it (A. Iolascon).

1079-9796/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bcmd.2013.02.003

Missense mutations in the ABCB6 transporter cause dominant familial pseudohyperkalemia

Immacolata Andolfo,[1,2] Seth L. Alper,[3,4,5] Jean Delaunay,[6] Carla Auriemma,[1,2] Roberta Russo,[1,2] Roberta Asci,[1] Maria Rosaria Esposito,[1] Alok K. Sharma,[3] Boris E. Shmukler,[3,4,5] Carlo Brugnara,[7] Lucia De Franceschi,[8] and Achille Iolascon[1,2]*

Familial Pseudohyperkalemia (FP) is a dominant red cell trait characterized by increased serum [K^+] in whole blood stored at or below room temperature, without additional hematological abnormalities. Functional gene mapping and sequencing analysis of the candidate genes within the 2q35–q36 critical interval identified—in 20 affected individuals among three multigenerational FP families—two novel heterozygous missense mutations in the *ABCB6* gene that cosegregated with disease phenotype. The two genomic substitutions altered two adjacent nucleotides within codon 375 of ABCB6, a porphyrin transporter that, in erythrocyte membranes, bears the Langereis blood group antigen system. The ABCB6 R375Q mutation did not alter the levels of mRNA or protein, or protein localization in mature erythrocytes or erythroid precursor cells, but it is predicted to modestly alter protein structure. ABCB6 mRNA and protein levels increase during *in vitro* erythroid differentiation of $CD34^+$ erythroid precursors and the erythroleukemia cell lines HEL and K562. These data suggest that the two missense mutations in residue 375 of the ABCB6 polypeptide found in affected individuals of families with chromosome 2-linked FP could contribute to the red cell K^+ leak characteristic of this condition. Am. J. Hematol. 00:000–000, 2012. © 2012 Wiley Periodicals, Inc.

Introduction

Familial pseudohyperkalemia (FP) is a dominant red cell trait characterized by increased serum [K^+] measured in whole-blood specimens stored at or below room temperature. This dominantly inherited trait is not accompanied by clinical symptoms or biological signs except for borderline abnormalities of red cell shape [1]. FP Lille was described in a large family of Flemish origin with morphologically normal red cells [2,3]. In this family, cation leak measured *in vitro* in the presence of ouabain and bumetanide showed normal K^+ efflux at 37°C, which increased greatly at 22 and 9°C [3]. The subsequently reported asymptomatic cases of FP Chiswick and FP Falkirk [4] were remarkable for increased MCV. FP Lille was mapped to 2q35–q36 by genome-wide search [5].

FP is considered as a subtype of the larger group of leaky red blood cell (RBC) disorders that include Southeast Asian ovalocytosis [6], dehydrated hereditary stomatocytosis (DHSt) [7], overhydrated hereditary stomatocytosis (OHSt) without or with neurological symptoms [8], and cryohydrocytosis (CHC). There appears to be a continuum between FP and DHSt that may be associated with pseudohyperkalemia [7]. Several temperature-dependent patterns of cation leak have been characterized in these conditions by measurements of ^{86}Rb influx insensitive to ouabain and bumetanide [9]. FP stands out among the leaky RBC disorders for its mild clinical and hematological phenotype and its minimal changes in cell shape.

Mutations in several genes have been shown to cause red cell cation leak disorders. These include *SLC4A1* (an ion exchange) in CHC and in atypical forms of hereditary spherostomatocytosis [10,11], *RHAG* (Rh-associated glycoprotein) in isolated stomatin-deficient OHSt [12,13] *GLUT1* (glucose transporter 1) in echinocytosis with paroxysmal dyskinesia [14], stomatin-deficient cryohydrocytosis [15] or in CHC [16] or pseudohyperkalemia and hemolysis [17] with neurological symptoms, and *PIEZO1* (mechanosensitive cation channel protein FAM38A) in DHSt [18]. These findings suggest that distinct missense mutations in various red cell membrane solute transporters or channels generate cation leak pathways either through the mutant proteins themselves or by deregulating one or more independent cation permeabilities of the red cell membrane.

Here, we report that in FP Lille, linked to chromosome 2q, and in two other FP families, the causal mutations reside in the same codon (375) of the ABCB6 gene, encoding the ABCB6 polypeptide reported to be a porphyrin transporter [19]. This protein was recently identified in the RBC membrane [20], where it displays the Langereis blood group [21].

Materials and Methods

Cases reports. FP Lille was first diagnosed in a mother and daughter [2,3] from a large family of Flemish descent. The carriers were hematologically normal. Temperature-dependent ^{86}Rb influx showed a shallow slope pattern [5]. The responsible gene was mapped to 2q35–q36 (very close to marker D2S1338), based on the analysis of 23 carriers (including the above-mentioned daughter as individual II.8)

[1]CEINGE, Biotecnologie Avanzate, Naples, Italy; [2]Department of Biochemistry and Medical Biotechnologies, "Federico II" University of Naples, Naples, Italy; [3]Division of Nephrology, Beth Israel Deaconess Medical Center, Boston, Massachusetts; [4]Division of Molecular and Vascular Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts; [5]Department of Medicine, Harvard Medical School, Boston, Massachusetts; [6]UMR_S 779, INSERM, Faculté de Médecine Paris-Sud, Université Paris-Sud, 94275 Le Kremlin-Bicêtre, Paris, France; [7]Department of Laboratory Medicine, Children's Hospital Boston, and Harvard Medical School, Boston, Massachusetts; [8]Department of Medicine, University of Verona, Piazzale Lo Scuro 10, Verona, Italy

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: Nothing to report

*Correspondence to: Achille Iolascon, CEINGE Biotecnologie Avanzate, Via Gaetano Salvatore, 486, Naples 80145, Italy. E-mail: achille.iolascon@unina.it

Contract grant sponsor: Italian Ministero dell'Università e della Ricerca (MIUR); Contract grant sponsor: Telethon (Italy); Contract grant number: GGP09044, MUR-PS 35-126/Ind; Contract grant sponsor: Regione Campania, DGRC 1901/200; Contract grant sponsor: The Doris Duke Charitable Foundation.

Received for publication 17 October 2012; Accepted 17 October 2012

Am. J. Hematol. 00:000–000, 2012.

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/ajh.23357