REVIEW

The hemoglobinopathies, molecular disease mechanisms and diagnostics

Cornelis L. Harteveld[1] | Ahlem Achour[1,2] | Sandra J. G. Arkesteijn[1] | Jeanet ter Huurne[1] | Maaike Verschuren[1] | Sharda Bhagwandien-Bisoen[1] | Rianne Schaap[1] | Linda Vijfhuizen[1] | Hakima el Idrissi[1] | Tamara T. Koopmann[1]

[1]Department of Clinical Genetics/LDGA, Leiden University Medical Center, Leiden, The Netherlands

[2]Department of congenital and hereditary diseases, Charles Nicolle Hospital, Tunis, Tunisia

Correspondence
Cornelis L. Harteveld, Department of Clinical Genetics/LDGA, Leiden University Medical Center, Einthovenweg 20, 2333ZC Leiden, The Netherlands.
Email: c.l.harteveld@lumc.nl

Abstract

Hemoglobinopathies are the most common monogenic disorders in the world with an ever increasing global disease burden each year. As most hemoglobinopathies show recessive inheritance carriers are usually clinically silent. Programmes for pre-conception and antenatal carrier screening, with the option of prenatal diagnosis are considered beneficial in many endemic countries. With the development of genetic tools such as Array analysis and Next Generation Sequencing in addition to state of the art screening at the hematologic, biochemic and genetic level, have contributed to the discovery of an increasing number of rare rearrangements and novel factors influencing the disease severity over the recent years. This review summarizes the basic requirements for adequate carrier screening analysis, the importance of genotype–phenotype correlation and how this may lead to the unrevealing exceptional interactions causing a clinically more severe phenotype in otherwise asymptomatic carriers. A special group of patients are β-thalassemia carriers presenting with features of β-thalassemia intermedia of various clinical severity. The disease mechanisms may involve duplicated α-globin genes, mosaic partial Uniparental Isodisomy of chromosome 11p15.4 where the *HBB* gene is located or haplo-insufficiency of a non-linked gene *SUPT5H* on chromosome 19q, first described in two Dutch families with β-thalassemia trait without variants in the *HBB* gene.

KEYWORDS
haematology, haemoglobin, molecular diagnosis, sickle cell disease, thalassemia, *β*-thalassemia intermedia

1 | INTRODUCTION

Haemoglobin is the major protein responsible for oxygen transportation in the human body and the major component of the red blood cell. The adult HbA ($\alpha2\beta2$) is a tetrameric protein for which the coding genes are grouped in two separate globin gene cluster families on different locations in the genome. Hemoglobinopathies, the genetic diseases related to haemoglobin synthesis, constitute the most common monogenic disorders worldwide. The genetic cause of this group of diseases are DNA variants in or near the globin genes, coding for the

globin chains of the tetrameric haemoglobin protein.[1,2] These DNA variants may result in altered synthesis of α- or β-globin (the α- and β-thalassemia syndromes respectively) or structural changes of haemoglobin, causing disease such as sickle cell disease, haemolytic anaemia, erythrocytosis or polycythaemia.

The interaction between thalassemia variants and various structural haemoglobin variants produce a wide range of disorders of varying clinical severity. The most important categories for which genetic counselling is indicated, with eventually the option of prenatal diagnosis, are Thalassemia Major (TM), the Sickle Cell Syndromes, the HbE/β-thalassemia combinations and the α-thalassemia syndromes, such as the lethal Hb Bart's and HbH Hydrops Fetalis Syndromes. The clinical relevance of these forms may differ amongst populations as the incidence is largely population specific. β-Thalassemia Major is a considerable health problem in the Mediterranean and Middle and Far East, resulting in carrier screening programs to prevent the birth of affected children. The HbE/β-thalassemia syndromes are more common in South-East Asia as well as the α0-thalassemia related syndromes, while HbS is more common in sub-Saharan Africa, India and the Middle East.

An estimated 7% of the world population carry a DNA variant which causes a defective haemoglobin synthesis, leading to approximately 300 000 to 400 000 affected newborn babies of which the majority (approx. 300 000) has sickle cell syndromes and a minor part (approx. 40 000) transfusion dependent β-thalassemia major.[3,4] In most populations where hemoglobinopathies are endemic, α- and β-thalassemia co-exists along with various abnormal hemoglobins. Historically hemoglobinopathies are most endemic in subtropical regions of the world due to the presence of malaria, extending from the Mediterranean area, Middle East and India, to South-East Asia. An increasing amount of evidence suggests natural selection favouring the carrier state as carriers tend to survive an infection by *Plasmodium falciparum* induced malaria tropica better than non-carriers. Due to centuries of migration, hemoglobinopathies have become widespread also in formerly non-endemic regions, such as North- and South America and Northern Europe.

Several studies have shown that the impact of global disease burden presented by the hemoglobinopathies is increasing each year.[5] Patients with β-thalassemia intermedia or major as well as sickle cell disease patients require life-long treatment. This involves regular blood transfusions and iron chelation therapy in the case of the β-thalassemia syndromes and management to reduce the painful vaso-occlusive crisis, anaemia, pulmonary hypertension and infections amongst others in Sickle cell patients. Although there is a prospect of curative therapies such as Haematopoietic Stem Cell transplantation and more recently gene-therapy using lentiviral vectors,[6] the lack of suitable donors in the former and high cost or lack of suitable infra-structure for treatment in the latter prevents a worldwide applicability. Due to the impact of these inherited disorders on patients, families and society, many endemic countries have developed pre-conceptional, pre-marital and/or antenatal carrier screening programs to minimize the incidence of new cases.[7] Carrier couples, if identified preferably before an affected child is born, can be offered reproductive options, within the legal framework and accepted local practices and legislation. Carrier screening and prevention of affected newborns has been most effective in several Mediterranean countries, such as Cyprus, Italy and Greece, which have implemented this already from the early 70-ties.[8,9] A renewed increase in number of affected births in these areas is attributable to recent immigration flows.[7]

Although the severe forms of hemoglobinopathies rarely escape from provisional clinical diagnosis by clinicians, the carriers can easily be missed. Carriers are usually asymptomatic and only identified during family analysis because of an affected family-member, when taking part in hemoglobinopathy screening programs or by chance during routine hematologic or biochemical analysis, such as diabetes related HbA1c analysis by HPLC or CE. Molecular analysis of the globin genes will support the definitive diagnosis of patients, carriers and those presenting with atypical hematologic parameters.[10–12] The European Molecular Genetics Quality Network (EMQN) has published a recommendation in 2015 for carrier identification and prenatal diagnosis of hemoglobinopathies, to be used in conjunction with other guidelines and recommendations, such as those of the British Society of Haematology,[13] ENERCA and the UK NHS Sickle Cell and Thalassemia screening program (http://sct.screening.nhs.uk/standardsandguidelines).

Due to the fact that the hemoglobinopathies are highly heterogeneous, the definitive diagnosis of carriers and patients, screening, counselling and prenatal diagnosis are challenging. Not only for clinicians, but also for laboratory staff who have to use wide ranging methodology to come to a definitive diagnosis. With the emerging Next Generation Sequencing tools, previously incomplete diagnosed cases were revisited and lead to the discovery of new mechanisms involved in modulation of disease severity. This review aims to highlight the challenges associated with new technical developments in carrier detection and some uncommon mechanisms leading to intermediate disease phenotypes in carriers of a single β-thalassemia variant.

2 | GENETICS

Haemoglobin is a tetrameric protein composed of two α-like and two β-like globin chains. The α- and β-globin genes are located in two separate clusters on different chromosomes. The duplicated α-globin genes (*HBA1* and *HBA2*) are located on chromosome 16p in the α-globin gene cluster together with the embryonically expressed HBZ1. The β-globin gene cluster, located on chromosome 11p, harbours the prenatally expressed *HBE*, *HBG1* and *HBG2* and the postnatally expressed *HBD* and *HBB* genes.

The embryonic, fetal and adult globin genes are organized along the genome in order of expression and regulated in a developmental stage- and tissue specific manner resulting in different types of haemoglobin, also called the haemoglobin switch (Figure 1). During the embryonic phase, the genes *HBZ1* and *HBE* express the Hb Gower I (ζ2ε2), II (α2ε2) and Hb Portland (ζ2γ2), in the following 6 weeks the *HBZ1* and *HBE* genes are switched off and HbF is expressed (α2γ2).
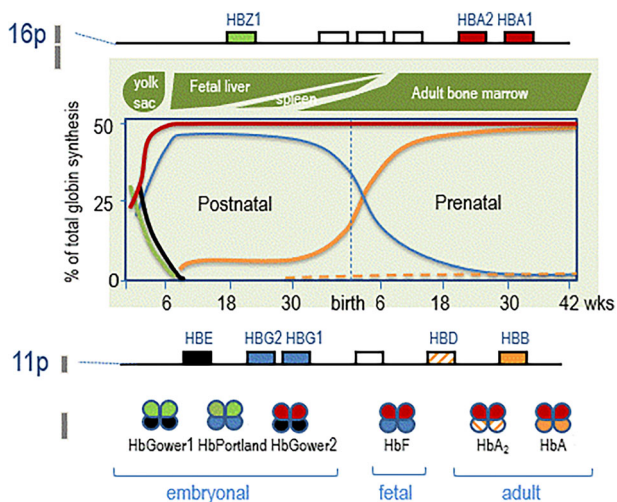
FIGURE 1 Schematic presentation of the different types of haemoglobin molecules synthesized during different stages of development. The graph shows the ontological expression of the α-like and non-α-like-globin genes pre- and postnatally and the steep change in γ- to β-globin synthesis around birth.

Around birth there is a steep repression of γ gene expression in favour of the β-globin gene being expressed, resulting in a decrease in HbF level down to less than 0.5% and an increasing synthesis of HbA ($\alpha 2\beta 2$) up to approx. 97–98%, which is completed approx. 6 months after birth. At the same time *HBD* is expressed and HbA_2 ($\alpha 2\delta 2$) synthesized (2%–3%), which has a diagnostic value in detecting β-thalassemia carriers for which the HbA_2 levels are elevated (4%–8%). Disease phenotypes associated with variants affecting the integrity of the β-globin gene are expressed approximately 6 months after birth, which is also the time that elevated HbA_2 in β-thalassemia carriers can be measured adequately.

Variants in the α- and β-globin genes, that is, the duplicated *HBA1* and *HBA2* and *HBB*, respectively, constitute the majority of clinically relevant hemoglobinopathies, the δ- and γ-globin chains coded by *HBD* and *HBG1* and *HBG2*, respectively, are clinically less important. More than 1600 globin gene variants are described (HbVar, ITHANET) of which around one third are clinically relevant.[11] Genomic variants involve deletions, removing functional genes or regulatory elements, or nucleotide variants in- or nearby globin genes, which interfere with proper translation, transcription or gene expression. The majority of α-thalassemia defects, roughly 80% of all molecular causes, are explained by deletions involving one or both *HBA1* and *HBA2*, deletions of the upstream Major conserved Sequences also called HS-40 are less common. The remaining approx. 20% is due to non-deletion defects in the *HBA1* and *HBA2* genes or rare deletions. The occurrence of specific variants is largely population specific, however certain deletions, such as the $-\alpha^{3.7}$ can be found in almost every population. The β-thalassemia mutations involve mainly nucleotide variants (more than 95%) while (partial) deletions of the β-globin gene with or without the upstream globin genes and β Locus Control region are much less common (approx. 5%).[12,14–16]

Traditionally DNA variants are classified according to the abolished synthesis of globin chains by the affected allele as 'α-zero' (both *HBA1* and *HBA2* deleted in cis) or 'α-plus' (one gene deleted or non-functional) and 'β-zero' or 'β-plus', respectively, the lack of expression or some remaining expression of the mutated gene. Other rearrangements, such as the δ-β hybrid gene or $(\delta\text{-}\beta)^0$ and $(\gamma\text{-}\delta\text{-}\beta)^0$-thalassemia lead to resp. Hb Lepore or Hereditary persistence of Fetal Haemoglobin (HPFH) with different levels of HbF expression. The clinically most frequently occurring Hb variant is HbS, which causes SCD in homozygosity and in combination with a whole range of other common Hb variants, such as HbC, HbD, HbE, HbO-Arab etc. and in combination with β-thalassemia. A variety of less common Hb variants in both the α- and β-globin genes may result in a wide spectrum of clinical conditions, such as erythrocytosis, poly-cythemia and hemo-lytic anaemia.[2]

There are numerous potential interactions between the different variants and resulting in different and sometimes unexpected or unexplained phenotypes. A helpful tool is provided by the Ithanet Portal, Databases and tools, called Ithaphen on the Ithanet web-site in this respect (IthaPhen (ithanet.eu) One of the major challenges when applying screening for the most common variants is to compare the results with the haematology and biochemical results for the full explanation of the phenotype. If there are contradictory results or the genotype does not explain the phenotypic severity, more research is needed. These exceptional cases may unravel novel rearrangement or disease mechanisms explaining the disease phenotype to reach a definitive diagnosis.

3 | LABORATORY HAEMATOLOGY

The routine analysis to diagnose a hemoglobinopathy carrier involves a complete red blood count including Haemoglobin concentration (Hb g/l), Mean Cellular Volume (MCV fl), Mean Corpuscular Haemoglobin (MCH pg), Packed Cell Volume (PCV l/l), and Red Blood Cell count (RBC $\times 10^{12}$ /L). Instruments for analysing most of these parameters can be automated, such as automated electronic cell counters. In addition haptoglobin (Hp) and Zinc Protoporphirin (ZPP) or ferritin are markers used in relationship to the complete red blood cell count (i.e., low Hb, RBC and PCV) to distinguish hemolysis or iron deficiency anaemia (low Hb, MCV, MCH and low/normal RBC and elevated ZPP or low ferritin/SeFe). If the MCV and MCH are low at normal iron levels, indicating the microcytic hypochromia, a presumptive diagnosis of a thalassemia carrier can be made. Hb variants, such as the common HbS, HbC or HbD variants for example may present with normal parameters and remain undetected by hematologic analysis.

To identify Hb variants and to distinguish an α-thalassemia carrier from a β-thalassemia carrier, capillary electrophoresis (CE) or high pressure liquid chromatography (HPLC) is performed. Dedicated devices are available on the market reaching high sensitivity in detecting and quantitating abnormally segregating haemoglobin fractions.[17] Recent years have shown an increasing use of HPLC devices

for the detection of HbA1c in diabetes diagnostics and follow-up. Most of these devices are capable of detecting at least the most common variants like HbS, HbC, HbD and HbE. Equipment designed to detect Hb variants are optimized for a superior separation and quantification of abnormal Hb fractions[18] and suitable for hemoglobinopathy diagnostics. Caution should be taken to use a presumptive identification of an Hb variant as other (rare) variants may migrate similarly on CE or HPLC. Other methods, such as sickle cell test for HbS or molecular confirmation is recommended for a definitive diagnosis. The quantification of the HbA_2 is important to distinguish a β- from an α-thalassemia carrier.[12] In the case of low or normal HbA_2 (2.0%–3.1%) it is likely an α-thalassemia, an elevated HbA_2 (>3.5%) is likely a β-thalassemia carrier and an HbA_2 between 3.1% and 3.5% is indecisive and could be either an α- or normal HbA_2 β-thalassemia carrier. Molecular analysis is needed to confirm a presumptive diagnosis (see Figure S1, Flow chart 'Patient and carrier diagnostics hemoglobinopathies').

4 | MOLECULAR DIAGNOSTICS

Each population endemic for α- and β-thalassemia has a usually limited spectrum of 10–20 frequently occurring DNA variants, along with perhaps one or two clinically significant haemoglobin variants. It is therefore common practice to choose methods for a population specific targeted mutation detection strategy, such as Allele Refractory Multiplex-PCR (ARMS) or Reverse dot-blot based assays such as commercially available population specific strip-assays.[19,20]

However, due to worldwide migration and the evolvement of large multi-ethnic societies, the mutation spectrum increased and more generic methods to detect disease-causing DNA variants are necessary. Automated Direct Sanger sequencing has long been the standard for the detection of disease specific DNA variants. A comprehensive PCR protocol as well as primers to amplify the HBA1, HBA2 and HBB used in our laboratory can be found in Table S1.[10]

Deletions are a common molecular cause of α-thalassemia and the seven most common deletions ($-\alpha^{3.7}$, $-\alpha^{4.2}$, $-^{SEA}$, $-^{MED1}$, $-^{THAI}$, $-^{FIL}$ and $-(\alpha)^{20.5}$) can easily be detected using a multiplex PCR strategy[21–25]. Other rearrangements which may play a deteriorating role in disease severity, such as the α-gene triplication ($\alpha\alpha\alpha^{anti\ 3.7}$) can also be detected by adding specific primers to the multiplex PCR. Protocols and primers used for screening the most common α-thalassemia deletions in our laboratory have been reviewed previously and updated with novel primers in Table S2.[10]

Commercially available Multiplex Ligation-dependent Probe Amplification (MLPA) assays are frequently used for the detection of (unknown) rearrangements in the α- and β-globin gene clusters[26] (www.mrcholland.com/technology/mlpa). This may eventually be replaced by Next Generation Sequencing (NGS) based Copy Number Variation (CNV) detection, which becomes increasingly successful using better software tools to interpret NGS obtained data.[27]

Now with the evolving technology of Next Generation Sequencing many molecular diagnostic labs have implemented Inherited Disease Panels, Whole Exome- and Whole Genome Sequencing[28] for a variety of genetic diseases. For the hemoglobinopathies the high homology between the HBA1 and HBA2 and between HBD and HBB, puts a challenge to NGS methods as short sequence reads interfere with the specificity of the highly homologous genes. As the globin genes are relatively small and NGS technology expensive, there are still many labs applying Sanger sequencing and MLPA as routine methods in diagnostics for hemoglobinopathies. This may change over the coming years when NGS is embraced as a universal applicable and affordable DNA sequence strategy in the genetic laboratory performing diagnostics for a variety of genetic diseases not only for the hemoglobinopathies.

5 | MOLECULAR BASIS AND HETEROGENEITY OF THE HEMOGLOBINOPATHIES

According to the EMQN recommendations for hemoglobinopathies diagnostics all hemoglobinopathy cases should be considered in relationship to the hematologic screening results and confirmed that the genotype and phenotype are consistent.[12]

With the methods and lab flow presented previously (see Figure S1, Flow chart 'Patient and carrier diagnostics hemoglobinopathies') almost all cases can be adequately diagnosed. However, attention needs to be paid to those cases for which the genotype–phenotype correlation are inconsistent and remain unexplained by the finds at the molecular level. These cases should be studied further and may lead to the discovery of additional rearrangements, involvement of other genes or novel disease mechanisms. Evaluation of family history and haematology may be relevant in these cases.

The atypical phenotype of β-thalassemia intermedia varies between the clinically silent β-thalassemia trait and severe hemolytic anaemia requiring regular blood-transfusion-chelation therapy. Usually this is genotype dependent, for instance compound heterozygosity for combinations of different β^+ and β^0-thalassemia determinants, combinations with α^0-thalassemia or high HbF expression due to co-inheritance of determinants causing Hereditary Persistence of Fetal Haemoglobin. A specifically interesting group are the β-thalassemia carriers expressing a more severe phenotype than the expected β-thalassemia trait. Instead of a mild microcytic hypochromic anaemia fitting the carrier state of the specific HBB variant, patients may present with features of β-thalassemia intermedia, characterized by moderate to severe microcytosis, with or without hemolysis and anaemia, and sometimes even transfusion dependency. A few cases we came across during our routine diagnostic practice and international collaboration are reviewed here. These involve β-thalassemia carriers presenting with features of β-thalassemia intermedia, due to duplicated α-globin genes, rare cases of patients born as carriers but developing adult onset β-thalassemia intermedia due to a mosaic partial Uniparental Isodisomy (UPiD) and carriers of β-thalassemia with haploinsufficiency of SUPT5H showing clinical features of a moderate β-thalassemia intermedia.[29–33]
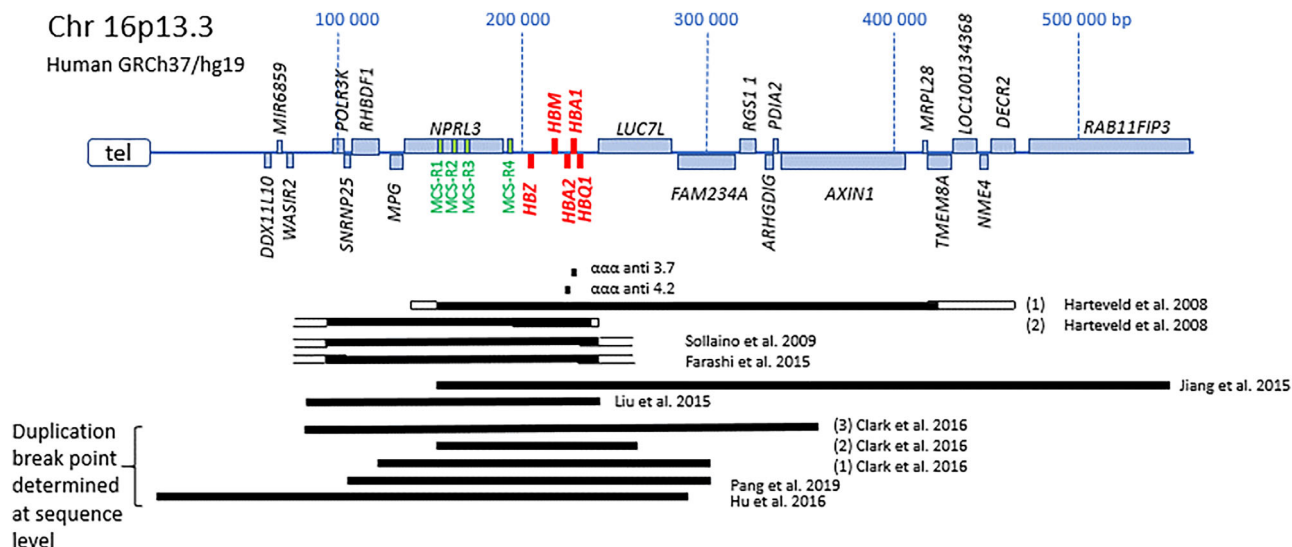
FIGURE 2 Summary of duplications of the α-globin gene locus on chromosome 16p13.3. The duplicated area is indicated as a black bar. The region of uncertainty when the duplication is not entirely mapped is indicated in white. The open lines indicate that the extend of the duplication is unknown. From the sequence data available a head-to-tail orientation was evident. The last duplication by Hu et al. (2016) was found to be integrated at position 94 054 bp, upstream of POLR3K.

6 | β-THALASSEMIA INTERMEDIA DUE TO EXTRA α-GLOBIN GENES

Co-inheritance of α-thalassemia reduces the chain imbalance and severity of disease in β-thalassemia homo- and compound heterozygotes. Increased α-globin synthesis on the other hand increases chain imbalance converting a typically asymptomatic carrier state into a thalassemia intermedia. This is most evident from the α- β-chain ratio measured by an isotopic procedure using tritium-labelled Leucine in a fraction enriched for reticulocytes,[34] however only very few labs have this technique up and running. The majority of diagnostic labs rely on the interpretation of the haematology and clinical phenotype of family-members in informative pedigrees. A profound difference in Hb, RBC, MCV, MCH, HbF level and clinical severity between sibs with seemingly identical genotypes raise suspicion and necessitates further DNA analysis.

The existence of α-globin gene triplications (ααα $^{anti\ 3.7}$ and ααα $^{anti\ 4.2}$) and the aggravating effect in carriers of β-thalassemia variants causing a pronounced anaemia is known for quite a while.[35-38] After the introduction of MLPA as a standard tool in molecular diagnostics for thalassemia, a variety of cases emerged to have duplications of the complete α-gene cluster, including the HS-40. The alleles thus carrying four α-genes in cis result from a segmental duplication of chromosomal areas ranging from 120 to 400 kb including a variety of non-haemoglobin related neighbouring genes as well[30]. A summary of duplications is shown in Figure 2.

It is remarkable that carriers of these extra copies of α-genes without the β-gene defect show no hematologic abnormalities in the families studied, and thus will not be easily identified by routine analysis as described previously. On the other hand, when the α-gene duplication is inherited along with a β-thalassemia defect, the effect is undeniable. The clinical severity is largely dependent on the type of β-thalassemia variant (β$^+$ or β^0-thalassemia trait) and on the number of extra α-genes, which is sometimes increased further by inheritance of a common α-triplication in addition to the segmental duplication. Two cases have been studied by FISH analysis using cosmid probes in the α-gene cluster[29]. This revealed no chromosomal translocations and it was therefore expected that the duplications were sequential. Five cases were studied at sequence level using either breakpoint PCR and direct sequencing or NGS techniques and the orientation of the duplications studied appeared head-to-tail.[30,39-44] More cases need to be studied at the molecular level to determine duplication breakpoints and orientation to speculate on the mechanism.

As carriers of α-globin gene duplications are clinically asymptomatic, it may occur more frequently but remain unnoticed. The presence of these segmental duplications only become evident when in combination with β-thalassemia. It may explain at least part of many unclear cases of severe thalassemia intermedia in β-thalassemia carriers. That other mechanisms are responsible for increased clinical severity in carriers became evident from a study by Origa et al.[45] Out of 33 patients showing unexplained β-thalassemia intermedia, 19 were found to have duplications of the α-globin genes, but for 14 patients the reason for the more severe phenotype remained obscure.[45]

7 | ADULT-ONSET β-THALASSEMIA INTERMEDIA IN PREVIOUSLY ASYMPTOMATIC CARRIERS

An intriguing category of patients is those born as asymptomatic carriers but developing adult onset β-thalassemia intermedia with blood transfusion dependency later in life. Thalassemia and sickle cell
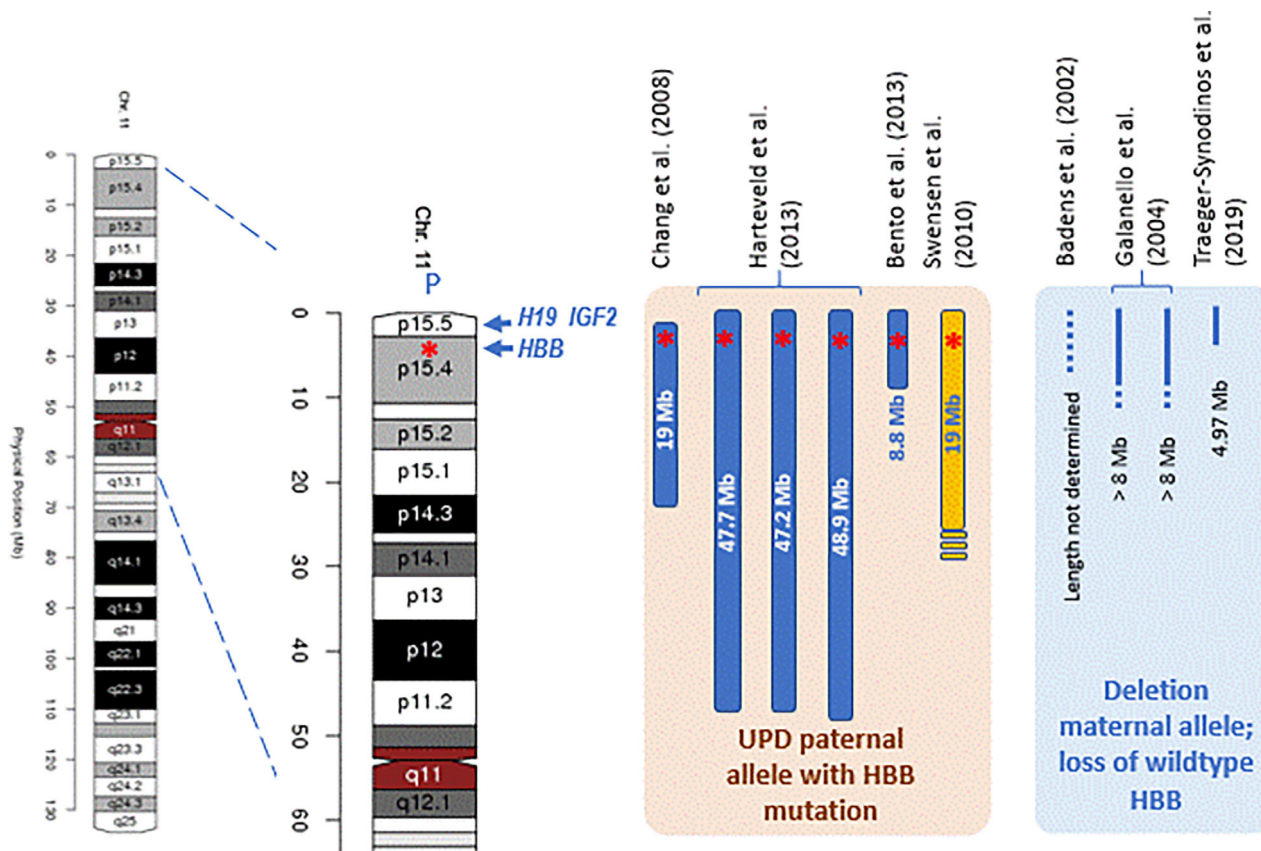
FIGURE 3 Schematic presentation of the short arm of chromosome 11 with summary of partial Uniparental Iso-Disomy (UPD) and segmental deletions of chromosome 11p (the scale is in Mb). The position of the *HBB*, *H19* and *IGF2* are indicated. The paternally inherited variant in the *HBB* gene is indicated with an asterisk. The blue bars indicate the segmental Uniparental isodisomic region of the paternal allele carrying the β^0-thalassemia variant, in orange the HbS variant detected by Swensen et al. in a boy with late-onset sickle cell disease. The deletions of the maternally inherited allele are indicated as lines. Stippled lines indicate uncertainty in length. All patients show mosaicism for either UPD or deletions in the DNA derived from blood isolated when developing adult-onset of β-thalassemia intermedia or major.

disease are typical examples of autosomal recessive diseases with Mendelian inheritance wherein each parent contributes one DNA variant to an affected offspring. However, in rare instances the child may develop such a disorder when only one parent transmits the mutation. Uniparental disomy (UPD) may unmask a recessive mutation or, alternatively, a deletion may occur de novo in cells of a person born as an asymptomatic heterozygote. The first case of UPD related β-thalassemia major was reported by Beldjord et al. in 1992 in a child born with Beckwith-Wiedemann Syndrome and β-thalassemia major.[46] However in 2008 Chang first reported an adult-onset β-thalassemia major patient, born as a clinically silent carrier of HBB: c.52A > T inherited from her father. The lady was diagnosed with β-thalassemia major and became transfusion dependent at the age of 28 years.[33] DNA analysis revealed almost complete homozygosity for the paternally inherited mutation, while her mother was completely normal. A somatic de novo deletion was excluded by fluorescence in situ hybridization (FISH), showing no loss of signal of the *HBB* region. Subsequent SNP genotyping assays revealed homozygosity for the region on 11p14.3 to 15.5 where the β-globin gene locus is located, as well as *H19* and *IGF2*, which play a role in Beckwith Wiedemann Syndrome (BWS). Interestingly, mosaic segmental paternal isodisomy

of 11p15.5 has been demonstrated in 20% of BWS cases. BWS is a disorder in which growth regulation is disturbed leading to macrosomia, hemihyperplasia and an increased risk for developing embryonic tumours. This syndrome is characterized by imprinting and differential expression of maternally and paternally inherited chromosomes. None of the patients developing adult-onset β-thalassemia major have clinical features of BWS.

Additional cases of UPD involving the short arm of chromosome 11 were found in patients showing paternal inheritance of a single β-thalassemia mutation and developing a transfusion dependent β-thalassemia major later in life (Figure 3). Three individuals from Italian descendance, one carrier of HBB:c.230delCfs and two carriers of HBB:c.118C > T, became transfusion dependent at the age of 30, 43 and 51 years old, respectively, showing almost complete homozygosity for the mutant allele.[31] ArrayCGH covering genome wide 730 K SNPs demonstrated that almost the entire short arm of chromosome 11 was derived from one parent. The degree of mosaicism and parental origin were determined by SNP ratio and comparison to the parental SNPs revealed resp. 65%, 70% and 80% of DNA exhibiting UPD of the paternally inherited allele at least 47 Mb in length. An additional case from Portugal carrying the paternally

inherited HBB:c.48G > A was diagnosed as β-thalassemia trait at the age of 7 years old and became regularly transfused from the age of 21 years every 3–4 weeks.[47] She appeared 80% mosaic for a 8.8 Mb segmental uniparental isodisomy including the entire HBB gene cluster and amongst others the H19 and IFG2 genes.

A second category consists of four independent cases of adult-onset β-thalassemia due to a mosaic somatic deletion of the maternal allele unmasking the paternally inherited allele with the HBB mutation, three cases of HBB:c.118C > T and one HBB:c.315 + 1G > A (Figure 3). Due to the deletion of the compensating maternal allele, the cells behave like β-thalassemia major cells and in concordance to the UPD cases there is loss of the imprinted allele. Remarkably in two cases, the level of mosaicism was estimated 10% and 20% (resp. Galanello et al and Badens et al.)[48,49]; both cases were anaemic with hepatosplenomegaly and growth failure, but received no blood transfusions. A third case (male, 42 years old) was splenectomized at age 40 and received no transfusions.[48] The fourth case was a female from Greek origin, showing the smallest deletion of approx. 4.97 Mb (11p15.4-11p15.5) and 90% mosaicism for the paternally inherited HBB:c.315 + 1G > A. She became blood transfusion dependent at the age of 34 years old.[32]

All cases without exception were born as β⁰- or severe β⁺-thalassemia carriers and developed an adult onset of β-thalassemia major. The β-thalassemia mutant was paternally inherited and loss of the maternally inherited allele involves the HBB, H19 and IGF2 genes. The H19 and IGF2 are prone to parental imprinting. It was concluded that the β-thalassemia major phenotype in these patients was associated with mosaic paternal UPID or a partial deletion of the maternal chromosome 11p15 only found in the patient's haematopoietic tissue and not in other tissues investigated (bone marrow, hair follicles, oral mucosal epithelium). Chang et al. (2008) postulate that loss of imprinting of H19 DMR the maternal allele as a consequence of the UPD of the paternal allele may result in IGF2 activation and over-expression, which could play a role in a cellular proliferation and growth of the clone homozygous for the paternally inherited mutation. Through proliferation this clone became the dominant haematopoietic cell in the bone marrow and peripheral blood, resulting in adult-onset β-thalassemia Major and transfusion dependency later in life.[33] This mechanism could also explain the maternal deletion cases as hemizygosity for the paternal allele involves loss of the imprinted maternal allele. A lesson to be learned from these cases is that analysis of the HBB gene variants should not be restricted to the blood, but other tissues should be investigated as well to unmask these somatic events.

8 | SUPT5H HAPLO-INSUFFICIENCY DOWN-REGULATING β-GLOBIN GENE EXPRESSION

Family tree analysis is a useful tool in determining genetic traits in thalassemia inheritance. Sometimes family members express a more severe phenotype than expected from the genotype. Two Greek families, one from Crete and the other from Macedonia showed an incompletely resolved moderate β-thalassemia intermedia in three family members.[50] A mother and daughter expressed lower Hb, MCV and MCH than usual with HBB:c.118C > T, while a 2 years old boy presented with moderate β-thalassemia intermedia being just a carrier of the common HBB:92 + 1G > A, just like his father who was only mildly affected. Strikingly, the HbA₂ was elevated between 8% and 12% in these individuals while normally the HbA₂ is between 4% and 6% in carriers of the afore mentioned HBB mutations. Full hematologic and HPLC/CE analysis revealed other family members with typical β-thalassemia trait but without deletions or point mutations in the HBB genes.

A similar phenomenon was found in two large Dutch families, showing a heritable β-thalassemia trait independently segregating from the HBB locus. Whole Exome Sequencing revealed a strong association with heterozygosity for "null" -mutations in a gene located on chromosome 19q previously not associated with β-globin gene expression or thalassemia called SUPT5H. Additional families and individuals showing a β-thalassemia trait without mutations in the HBB gene were examined for mutations in SUPT5H and eight different mutations have been found disrupting splice site consensus sequences or deletions/duplications causing frameshifts and premature termination of translation.[50]

The SUPT5H gene codes for the human Spt5 protein which forms a dimer with hSpt4 as part of the DSIF complex, regulating mRNA processing and transcription elongation by RNA polymerase II.[51–53] Although SUPT5H is ubiquitously expressed, it is highly expressed in testis and bone marrow. Animal and cellular studies suggest that 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole Sensitive Inducing Factor (DSIF) functions in a tissue specific manner.[54–57] The embryonic erythropoiesis in zebrafish is regulated by the transcription elongation factor Foggy/Spt5 through gata1 gene regulation. Erythrocytes of zSpt5 Knock-Down embryo's showing decreased expression suggest a disruption of erythroid differentiation probably through repression of GATA1 expression. The human Spt5 deficiency caused by haplo-insufficiency may act similarly through repression of GATA1 expression, causing down-regulation of HBB genes in trans. Further studies are needed to determine the exact role of SUPT5H in explaining the reduced β-gene expression in carriers and compound heterozygotes with true β-thalassemia determinants in individuals presenting with moderate β-thalassemia intermedia.

9 | CONCLUSION

The identification of carriers and offering prenatal diagnosis to couples at risk constitute valuable health services to offer informed reproductive choices. In contrast to most genetic diseases carriers of hemoglobinopathies can easily be detected using routine hematologic screening and protein separation and quantification by automated HPLC or CE. In our experience the strategy described here, including confirmation by DNA sequencing of the α- and β-globin genes and gap-PCR and MLPA analysis to detect deletions and duplications, is

effective for screening and definitive diagnosis of most cases of sickle cell disease, Hb variants, α- and β-thalassemia. The introduction of NGS will probably not change much to this concept as the globin genes are relatively small and covered completely by Direct Sanger Sequencing. However, care should be taken to correlate well the clinical phenotype with the hematologic, biochemical and genetic findings. If discrepancies occur NGS based techniques have proven to be an efficient tool to uncover novel molecular interactions in complex thalassemia cases. We have highlighted a few examples from our own experience of complex β-thalassemia carriers with discordant genotype–phenotype correlation, which were brought to our attention through international collaboration. Some of these cases remained unsolved for years, until the technology improved by the introduction of MLPA, arrayCGH or the rapidly evolving NGS techniques.

Due to the extreme heterogeneity characteristic the hemoglobinopathies, screening and definitive diagnosis can be challenging for laboratories. They need to apply a wide range of well validated tests by an experienced team of geneticists, haematologists and laboratory-staff. In addition, the clinical information derived from family history and pedigrees are mandatory in searching for genomic associations related to the disease phenotype. International collaboration proved essential in the examples mentioned, international initiatives to establish a network of professionals working on a particular rare disease, such as IthanNet and ENERCA united under the umbrella of Euro-BloodNet are helpful in bringing professionals into contact, promoting collaborations.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Even though data sharing is not directly applicable to this article as no new data were generated or analyzed in this study, the data presented are available.

ORCID

Cornelis L. Harteveld https://orcid.org/0000-0002-5135-6491

REFERENCES

1. Williams TN, Weatherall DJ. World distribution, population genetics, and health burden of the hemoglobinopathies. *Cold Spring Harb Perspect med*. 2012;2(9):a011692.
2. Steinberg MH, Forget BG, Higgs DR, Nagel RL. *Disorders of Hemoglobin*. 2nd ed. Genetics, Pathophysiology and Clinical Management; 2009.
3. Haldane JB. The rate of mutation of human genes. *Hereditas*. 1949;35:267-273.
4. Allison AC. Protection afforded by sickle-cell trait against subtertian malareal infection. *Br med J*. 1954;1(4857):290-294.
5. Piel FB, Patil AP, Howes RE, et al. Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nat Commun*. 2010;1:104.
6. Zittersteijn HA, Harteveld CL, Klaver-Flores S, et al. A small key for a heavy door: genetic therapies for the treatment of hemoglobinopathies. *Front Genome Ed*. 2020;2:617780.
7. Angastiniotis M, Petrou M, Loukopoulos D, et al. The prevention of thalassemia revisited: a historical and ethical perspective by the thalassemia international federation. *Hemoglobin*. 2021;45(1):5-12.
8. Cao A, Kan YW. The prevention of thalassemia. *Cold Spring Harb Perspect Med*. 2013;3(2):a011775.
9. Angastiniotis MA, Hadjiminas MG. Prevention of thalassaemia in Cyprus. *Lancet*. 1981;1(8216):369-371.
10. Harteveld CL. State of the art and new developments in molecular diagnostics for hemoglobinopathies in multiethnic societies. *Int J Lab Hematol*. 2014;36(1):1-12.
11. Traeger-Synodinos J, Harteveld CL. Preconception carrier screening and prenatal diagnosis in thalassemia and hemoglobinopathies: challenges and future perspectives. *Expert Rev Mol Diagn*. 2017;17(3):281-291.
12. Traeger-Synodinos J, Harteveld CL, Old JM, et al. EMQN best practice guidelines for molecular and haematology methods for carrier identification and prenatal diagnosis of the haemoglobinopathies. *Eur J Hum Genet*. 2015;23(4):426-437.
13. Ryan K, Bain BJ, Worthington D, et al. Significant haemoglobinopathies: guidelines for screening and diagnosis. *Br J Haematol*. 2010;149(1):35-49.
14. Giardine B, Borg J, Higgs DR, et al. Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nat Genet*. 2011;43(4):295-301.
15. Thein SL, Wood WG. *The Molecular Basis of Beta Thalassemia, Delta-Beta Thalassemia, and Hereditary Persistence of Fetal Hemoglobin*. Cambridge University Press; 2009.
16. Thein SL. Dominant beta thalassaemia: molecular basis and pathophysiology. *Br J Haematol*. 1992;80(3):273-277.
17. van DP LE, Bakker-Verweij M, de KM BU, Harteveld CL, et al. Evaluating five dedicated automatic devices for haemoglobinopathy diagnostics in multi-ethnic populations. *Int J Lab Hematol*. 2009;31(5):484-495.
18. Traeger-Synodinos J, Harteveld CL. Advances in technologies for screening and diagnosis of hemoglobinopathies. *Biomark med*. 2014;8(1):119-131.
19. Angastiniotis M, Eleftheriou A, Galanello R, et al. *Prevention of thalassaemias and Other Haemoglobin Disorders*. Thalassaemia International Federation; 2016.
20. Puehringer H, Najmabadi H, Law HY, et al. Validation of a reverse-hybridization StripAssay for the simultaneous analysis of common alpha-thalassemia point mutations and deletions. *Clin Chem Lab med*. 2007;45(5):605-610.
21. Liu YT, Old JM, Miles K, Fisher CA, Weatherall DJ, Clegg JB. Rapid detection of alpha-thalassaemia deletions and alpha-globin gene triplication by multiplex polymerase chain reactions. *Br J Haematol*. 2000;108(2):295-299.
22. Chong SS, Boehm CD, Higgs DR, Cutting GR. Single-tube multiplex-PCR screen for common deletional determinants of alpha-thalassemia. *Blood*. 2000;95(1):360-362.
23. Tritipsombut J, Phylipsen M, Viprakasit V, et al. A single-tube multiplex gap-polymerase chain reaction for the detection of eight beta-globin gene cluster deletions common in Southeast Asia. *Hemoglobin*. 2012;36(6):571-580.

24. Craig JE, Barnetson RA, Prior J, Raven JL, Thein SL. Rapid detection of deletions causing delta beta thalassemia and hereditary persistence of fetal hemoglobin by enzymatic amplification. *Blood*. 1994;83(6):1673-1682.

25. Craig JE, Barnetson R, Weatherall DJ, Thein SL. Rapid detection of a 13.4-kb deletion causing delta beta thalassemia in an Egyptian family by polymerase chain reaction. *Blood*. 1993;81(3):861-863.

26. Harteveld CL, Voskamp A, Phylipsen M, et al. Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification. *J med Genet*. 2005;42(12):922-931.

27. Roca I, Gonzalez-Castro L, Fernandez H, Couce ML, Fernandez-Marmiesse A. Free-access copy-number variant detection tools for targeted next-generation sequencing data. *Mutat Res Rev Mutat Res*. 2019;779:114-125.

28. Achour A, Koopmann TT, Baas F, Harteveld CL. The evolving role of next-generation sequencing in screening and diagnosis of hemoglobinopathies. *Front Physiol*. 2021;12:686689.

29. Harteveld CL, Refaldi C, Cassinerio E, Cappellini MD, Giordano PC. Segmental duplications involving the alpha-globin gene cluster are causing beta-thalassemia intermedia phenotypes in beta-thalassemia heterozygous patients. *Blood Cells Mol Dis*. 2008;40(3):312-316.

30. Clark B, Shooter C, Smith F, et al. Beta thalassaemia intermedia due to co-inheritance of three unique alpha globin cluster duplications characterised by next generation sequencing analysis. *Br J Haematol*. 2018;180(1):160-164.

31. Harteveld CL, Refaldi C, Giambona A, et al. Mosaic segmental uniparental isodisomy and progressive clonal selection: a common mechanism of late onset beta-thalassemia major. *Haematologica*. 2013;98(5):691-695.

32. Traeger-Synodinos J, Bosch CAJ, Vrettou C, et al. Adult-onset beta-thalassaemia intermedia caused by a 5-Mb somatic clonal segmental deletion in haemopoietic stem cells involving the beta-globin locus. *Br J Haematol*. 2019;186(6):e165-e170.

33. Chang JG, Tsai WC, Chong IW, Chang CS, Lin CC, Liu TC. {beta}-thalassemia major evolution from {beta}-thalassemia minor is associated with paternal uniparental isodisomy of chromosome 11p15. *Haematologica*. 2008;93(6):913-916.

34. Giordano PC, van Delft P, Batelaan D, Harteveld CL, Bernini LF. Haemoglobinopathy analyses in The Netherlands: a report of an in vitro globin chain biosynthesis survey using a rapid, modified method. *Clin Lab Haematol*. 1999;21(4):247-256.

35. Henni T, Belhani M, Morle F, et al. Alpha globin gene triplication in severe heterozygous beta thalassemia. *Acta Haematol*. 1985;74(4):236-239.

36. Oron V, Filon D, Oppenheim A, Rund D. Severe thalassaemia intermedia caused by interaction of homozygosity for alpha-globin gene triplication with heterozygosity for beta zero-thalassemia. *Br J Haematol*. 1994;86(2):377-379.

37. Goossens M, Dozy AM, Embury SH, et al. Triplicated alpha-globin loci in humans. *Proc Natl Acad Sci U S A*. 1980;77(1):518-521.

38. Traeger-Synodinos J, Kanavakis E, Vrettou C, et al. The triplicated alpha-globin gene locus in beta-thalassaemia heterozygotes: clinical, haematological, biosynthetic and molecular studies. *Br J Haematol*. 1996;95(3):467-471.

39. Jiang H, Liu S, Zhang YL, Wan JH, Li R, Li DZ. Association of an alpha-globin gene cluster duplication and heterozygous beta-thalassemia in a patient with a severe thalassemia syndrome. *Hemoglobin*. 2015;39(2):102-106.

40. Liu S, Jiang H, Wu MY, Zhang YL, Li DZ. Thalassemia intermedia caused by 16p13.3 sectional duplication in a beta-thalassemia heterozygous child. *Pediatr Hematol Oncol*. 2015;32(5):349-353.

41. Sollaino MC, Paglietti ME, Perseu L, Giagu N, Loi D, Galanello R. Association of alpha globin gene quadruplication and heterozygous beta thalassemia in patients with thalassemia intermedia. *Haematologica*. 2009;94(10):1445-1448.

42. Farashi S, Vakili S, Faramarzi Garous N, et al. Copy number variations of six and seven alpha-globin genes in a family with intermedia and major thalassemia phenotypes. *Expert Rev Hematol*. 2015;8(5):693-698.

43. Pang D, Shang X, Cai D, et al. Thalassaemia intermedia caused by coinheritance of a beta-thalassaemia mutation and a de novo duplication of alpha-globin genes in the paternal allele. *Br J Haematol*. 2019;186(4):620-624.

44. Hu L, Shang X, Yi S, et al. Two novel copy number variations involving the alpha-globin gene cluster on chromosome 16 cause thalassemia in two Chinese families. *Mol Genet Genomics*. 2016;291(3):1443-1450.

45. Origa R, Sollaino MC, Borgna-Pignatti C, et al. Alpha-globin gene quadruplication and heterozygous beta-thalassemia: a not so rare cause of thalassemia intermedia. *Acta Haematol*. 2014;131(3):162-164.

46. Beldjord C, Henry I, Bennani C, Vanhaeke D, Labie D. Uniparental disomy: a novel mechanism for thalassemia major. *Blood*. 1992;80(1):287-289.

47. Bento C, Maia TM, Milosevic JD, Carreira IM, Kralovics R, Ribeiro ML. Beta thalassemia major due to acquired uniparental disomy in a previously healthy adolescent. *Haematologica*. 2013;98(1):e4-e6.

48. Galanello R, Perseu L, Perra C, et al. Somatic deletion of the normal beta-globin gene leading to thalassaemia intermedia in heterozygous beta-thalassaemic patients. *Br J Haematol*. 2004;127(5):604-606.

49. Badens C, Mattei MG, Imbert AM, et al. A novel mechanism for thalassaemia intermedia. *Lancet*. 2002;359(9301):132-133.

50. Achour A, Koopmann T, Castel R, et al. A new gene associated with a beta-thalassemia phenotype: the observation of variants in SUPT5H. *Blood*. 2020;136(15):1789-1793.

51. Stachora AA, Schafer RE, Pohlmeier M, Maier G, Ponstingl H. Human Supt5h protein, a putative modulator of chromatin structure, is reversibly phosphorylated in mitosis. *FEBS Lett*. 1997;409(1):74-78.

52. Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell*. 2006;21(2):227-237.

53. Yamaguchi Y, Wada T, Watanabe D, Takagi T, Hasegawa J, Handa H. Structure and function of the human transcription elongation factor DSIF. *J Biol Chem*. 1999;274(12):8085-8092.

54. Komori T, Inukai N, Yamada T, Yamaguchi Y, Handa H. Role of human transcription elongation factor DSIF in the suppression of senescence and apoptosis. *Genes Cells*. 2009;14(3):343-354.

55. Krishnan K, Salomonis N, Guo S. Identification of Spt5 target genes in zebrafish development reveals its dual activity in vivo. *PLoS One*. 2008;3(11):e3621.

56. Taneda T, Zhu W, Cao Q, et al. Erythropoiesis is regulated by the transcription elongation factor Foggy/Spt5 through gata1 gene regulation. *Genes Cells*. 2011;16(2):231-242.

57. Yang Q, Liu X, Zhou T, Cook J, Nguyen K, Bai X. RNA polymerase II pausing modulates hematopoietic stem cell emergence in zebrafish. *Blood*. 2016;128(13):1701-1710.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.