Exome sequencing identifies *MPL* as a causative gene in familial aplastic anemia

Amanda J. Walne,[1] Arran Dokal,[1] Vincent Plagnol,[2] Richard Beswick,[1] Michael Kirwan,[1] Josu de la Fuente,[3] Tom Vulliamy,[1] and Inderjeet Dokal[1]

[1]Centre for Paediatrics, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK; [2]University College London Genetics Institute, Gower Place, London, UK; and [3]St Mary's Hospital Paddington, London, UK

ABSTRACT

The primary cause of aplastic anemia remains unknown in many patients. The aim of this study was to clarify the genetic cause of familial aplastic anemia. Genomic DNA of an affected individual from a multiplex consanguineous family was hybridized to a Nimblegen exome library before being sequenced on a GAIIx genome analyzer. Once the disease causing homozygous mutation had been confirmed in the consanguineous family, this gene was then analyzed for mutation in 33 uncharacterized index cases of aplastic anemia (<13 years) using denaturing HPLC. Abnormal traces were confirmed by direct sequencing. Exome sequencing identified a novel homozygous nonsense mutation in the thrombopoietin receptor gene *MPL*. An additional novel homozygous *MPL* mutation was identified in the screen of 33 aplastic anemia patients. This study shows for the first time a link between homozygous *MPL* mutations and familial aplastic anemia. It also highlights the important role of *MPL* in trilineage hematopoiesis.

Key words: aplastic anemia, MPL, exome sequencing.

Citation: Walne AJ, Dokal A, Plagnol V, Beswick R, Kirwan M, de la Fuente J, Vulliamy T, and Dokal I. Exome sequencing identifies MPL *as a causative gene in familial aplastic anemia. Haematologica 2012;97(4):524-528. doi:10.3324/haematol.2011.052787*

Introduction

Aplastic anemia (AA, MIM #609138) is a rare heterogeneous disorder affecting 2-7 people per million per year.[1] In 50-70% of cases, the etiology is unknown. In the remainder, it is either inherited (10-20%) or is secondary to a particular insult (10-20%[2,3]). Some of the inherited cases can be linked to various syndromes, for example, Fanconi anemia (FA) or dyskeratosis congenita (DC), but others remain uncharacterized.[4] Familial AA is an extremely rare inherited subtype affecting multiple individuals in a family. Patients typically only have features of AA; the absence of any somatic features making it distinct from other inherited aplastic anemias. The primary cause of these uncharacterized familial cases is of interest to the scientist and can present a diagnostic conundrum to the clinician.

Until recently the main ways to identify new disease causing genes was to identify candidate gene regions using linkage analysis or by a direct candidate gene approach based on similar biological function. Often success in such traditional approaches required the recruitment of large families with many affected and unaffected individuals.[5,6] These approaches have a low success rate in identifying new disease genes, especially when there are no current genes associated with a particular disease phenotype. Inheritance type and family size also significantly affect the likelihood of success and compound heterozygous mutations cannot be identified using these types of approach. Recent advances, such as next generation sequencing, have made it possible to sequence the entire genome within weeks or, more easily, the entire coding exome within a few days.[7] This new technology now allows the identification of causative mutations in families and disorders that are not amenable to linkage-based positional cloning.[8]

The aim of this study was to use exome sequencing to identify the cause of familial AA in a consanguineous family with 2 affected children. This would then identify a gene to screen in other uncharacterized AA cases.

Design and Methods

Patient selection

DNA samples were received from 10 family members of a large consanguineous family from Tunisia (2 had severe AA: Family 1). Other DNA samples came from individuals aged under 13 years presenting with AA which was not associated with any other bone marrow (BM) failure syndrome. All samples were obtained with informed consent and with the approval of our local ethics committee.

Homozygosity mapping

Genome wide microsatellite analysis was performed on the

The online version of this article has a Supplementary Appendix.

Acknowledgments: we would like to thank all the families and clinicians for biological samples, specifically Dr. Abdelkefi (Tunisia) and Dr. Richards (UK).
Funding: this work was funded by The Wellcome Trust.
Manuscript received on July 29, 2011. Revised version arrived on November 29, 2011. Manuscript accepted on November 30, 2011.
Correspondence: Amanda Walne, Centre for Paediatrics, Blizard Institute of Cell and Molecular Science, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, 4 Newark Street, London, UK. Phone: international +44.207.8822458. E-mail: a.walne@qmul.ac.uk

2 affected children (Family 1: 4-I and 4-III) and where identical homozygous markers were identified, all family members were typed at this and flanking markers. Areas of homozygosity which were consistent with linkage were noted.

Exome sequencing

Approximately 180,000 coding exons from 5µg of genomic DNA from patient 4-III (Family 1) were sequenced on the Illumina Genome Analyser IIx after enrichment using the NimblGen SeqCap EZ exome library. Sequencing data were processed through the Illumina pipeline and unique homozygous changes identified by filtering the resultant data set against variations reported on dbSNP (*www.ncbi.nlm.nih.gov/snp/*), the 1000 genome project (*www.1000genomes.org/*) and an in-house data set.

MPL mutation detection and gene screening

Primers were designed to all coding exons (*Online Supplementary Table S1*). The mutation c.1248G>A p.W416X in exon 8 of *MPL*, was confirmed by direct sequencing using ABI BigDye v3.1. For the other 33 samples, all exons of *MPL* were screened by denaturing high performance liquid chromatography (dHPLC) using the Transgenomic Wave DNA fragment analysis system to detect heteroduplexes following pair-wise pooling. All abnormal elution patterns were sequenced as described above. Any coding variations identified were compared to dbSNP and 1000 genome databases to determine that they had not been described as a rare variation.

Telomere length measurement

Telomere lengths were measured using a multiplex real time PCR method as described previously.[9-10] T/S ratios obtained for the patient samples were compared with healthy control samples and also patients with known *TERC* mutations as examples of short telomeres.

Results and Discussion

Following microsatellite analysis, several regions of homozygosity shared by the 2 affected individuals were identified (Family 1, *Online Supplementary Figure S1*). Given the large number of genes in these intervals, DNA from patient 4-III (Figure 1A) underwent exome sequencing. We filtered the resultant data set against: i) known variations in published databases; ii) an in-house database; iii) excluding any changes not in a homozygous interval; iv) based on gene function; and v) precedence of a disease with similar clinical features being caused by mutations in the same gene. We were able to reduce the number of variations to study to a homozygous mutation c.1248 G>A, p.Trp416Stop in *MPL*. *MPL* (myeloproliferative leukemia virus oncogene) encodes the thrombopoietin (TPO) receptor. The interaction of thrombopoietin with its receptor is largely responsible for megakaryopoiesis, platelet activation, as well as the maintenance of hematopoietic stem cells (HSC). This also initiates intracellular signaling, including activation of the JAK/STAT, RAS/MAPK and P13K/AKT pathways.[11] Clinically, biallelic constitutional mutations in *MPL* have been described in congenital amegakaryocytic

Table 1. Clinical features of the patients with biallelic *MPL* mutations and a comparison with features of CAMT.

Feature Ethnic origin Relationship of parents	CAMT *	Family 1 Tunisia Consanguineous		Family 2 Pakistan Non consanguineous but parents from same village	
Age at presentation	Typically at birth/infancy	4 years (index case)	3 years	18 months (index case)	3 years
Presenting feature	Reduced/absent megakaryocytes and platelets	Severe aplastic anemia	Aplastic anemia with no other clinical abnormalities	Rotavirus infection (abnormal blood count noted)	Family study
Blood counts				(At age 6 yrs)	Thrombocytopenia only
Hb (g/dL)		7.1	NA	11.3	
WBC (×10⁹/L)		2.5		Low	
Platelets (×10⁹/L)		7		14	60
Retriculocytes (×10⁹/L)		15			
Neutrophils (×10⁹/L)		0.4		1.56	
Exclude inherited forms of thrombocytopenia	Yes	FA excluded	FA excluded	FA excluded	FA excluded
Bone marrow cellularity	Normal cellularity (initially)	Hypocellular, no dysmyelopoiesis	Hypocellular	Hypocellular, mild dyserythropoesis	Hypocellular
Other non hematologic abnormalities	Minor heart, eye and brain abnormalities reported in some	None	None	None	None
Treatment	Various including platelet transfusions, matched BMT	Successful HLA identical sib BMT (2 yrs post transplant full donor chimerism)	Successful HLA identical sib BMT (engraftment by day 11)	Close surveillance + supportive care	Close surveillance
MPL mutations	Homozygous or compound heterozygous	c.1248 G>A p.Trp416Stop (homozygous)		c.1180 C>T p. Pro394Ser (homozygous)	

**Definition from Ballmaier and Germeshausen[19]; BMT: bone marrow transplantation; FA: Fanconi anemia; Hb: hemoglobin; NA: not available; WBC: white blood cells.*

thrombocytopenia (CAMT). Patients with CAMT typically present with thrombocytopenia in the first few weeks/months of life.[12] Knockout studies in mice for either *Mpl* or *Tpo* show a significant decrease in the numbers of myeloid, erythroid, megakaryocytic and burst colony forming units, but apart from a reduction in platelet numbers, no reduction in other peripheral blood counts was observed. When progenitor assays are performed on BM from patients with CAMT, there is a significant reduction in the clonogenic cells of all three lineages when compared with BM from healthy individuals, suggesting that the thrombocytopenia and pancytopenia observed in these patients may result from loss of function mutations in *MPL*.[13]

Sequencing the c.1248 G>A, p.Trp416Stop variation in all Family 1 members confirmed this novel *MPL* mutation segregated with disease in an autosomal recessive manner (Figure 1A). We then extended our screen to patients that were reported to have AA in childhood but not necessarily in the neonatal period. From the clinical information available, the diagnosis of CAMT was not considered by the referring physician and it was not always clear if a thrombocytopenic phase preceded the presentation of the AA. A total of 33 patients were selected (mean age at presentation 4.5 years; range 3 months-11 years). All coding exons of *MPL* were screened and a novel missense mutation was identified in one patient. This patient (Family 2, from Pakistan) had a homozygous mutation c.1180 C>T p.Pro394Ser. A study of additional family members showed that a second affected brother was also homozygous for this mutation (Figure 1B). Both mutations identified in this study were clustered in exon 8. This exon is not reported to be widely mutated, and the majority of published muta-

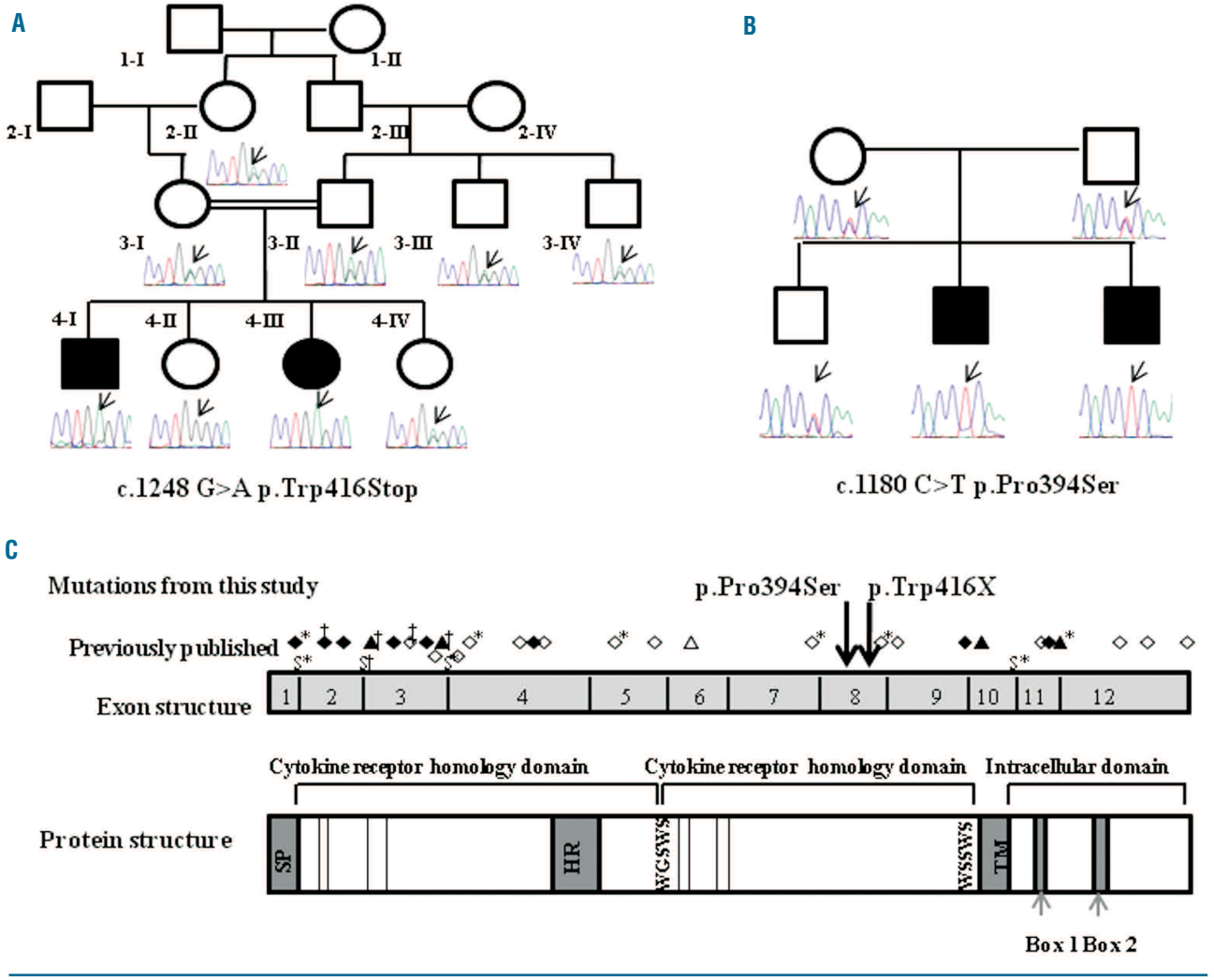

Figure 1. Genetic evaluation of *MPL* mutations in aplastic anemia. (A) Segregation of the c.1248G>A *MPL* mutation in Family 1 as identified by exome sequencing. Filled shapes: affected individuals; arrow: affected base. (B) The homozygous missense mutation c.1180C>T p.Pro394Ser segregates with disease in Family 2. Filled shapes: affected individuals; arrow: affected base. (C) Relative locations of the mutations identified in *MPL* from this study (marked by arrows) and previously published studies. The exon structure and derived protein structure with the functional domains are shown in the lower panel. Filled diamonds: nonsense mutations; open diamond: missense mutations; filled triangles: frame shift mutations; open triangle: 7bp deletion; S: splice site mutations; *: recurrent mutation with <5 mutated alleles reported; †: recurrent mutation with 5-10 mutated alleles reported; ‡: recurrent mutation with >10 mutated alleles reported; SP; signal peptide; TM; transmembrane domain. Adapted from Ballamaier *et al.*[19]

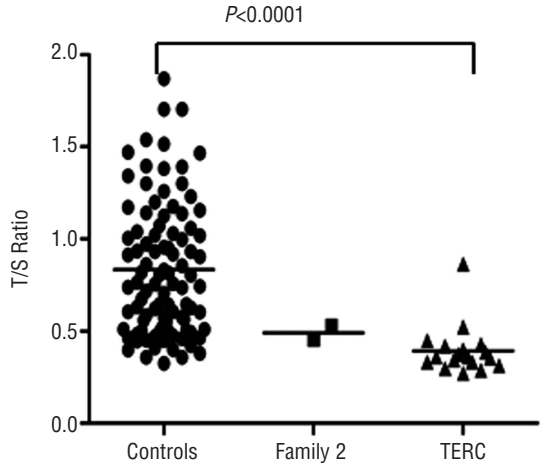

Figure 2. Telomeres are not significantly short in patients with aplastic anemia associated with biallelic *MPL* mutations compared with controls. Telomere lengths are short in patients with *TERC* mutations compared to controls.

tions tend to cluster in exons 2 and 3 (Figure 1C).

Many mutations in *MPL* have been identified in relation to CAMT (Figure 1C[12,14-16]) and this in turn has allowed different subtypes of CAMT to be described, partially based on the mutation type. Type I is the most severe form with early onset pancytopenia, decreased BM activity and very low platelet count. The mutations associated with this subtype involve premature stop codons or frame shift mutations that eliminate receptor signaling by deletion of all or most of the intracellular domain. Type II CAMT is a milder form of the disease with transient increases of platelet count up to nearly normal values in the first year of life and a later onset of BM failure (between 3-6 years). The mutations associated with this subtype include missense mutations or splicing defects resulting in reduced, residual receptor activity. A third type has also been described in which patients have ineffective megakaryopoiesis with no apparent defects in the *MPL* gene.[15,17,18]

In a systematic review of the literature, Ballmaier and Gereshausen report on all the cases of CAMT published since 1990.[19] Of the 71 cases where the age of onset of pancytopenia or BM hypocellularity was reported, the average age was 5.3 months (range 0-96 months). Our patient group differs from the more usual presentation of CAMT, suggesting they could be part of a different subtype of disease (key clinical features are shown in Table 1). As a diagnosis of CAMT was not suspected in the larger group of patients with AA, *MPL* was not an obvious candidate gene to screen without the information obtained from the exome sequencing data from Family 1. In Family 2, the thrombocytopenia was only identified when the children were undergoing clinical investigations for a viral infection. None of these patients had the classical presentation of CAMT.

In Family 1, although having a premature stop mutation that is usually associated with type I CAMT, the age of disease onset is later. Clinically, CAMT patients can be classified as type I if the platelet count in the first year of life is 50×10^9/L or under or, if this information is not present, pancytopenia develops before the age of three years.[15] From this classification it is apparent that the affected individuals in Family 1 have a different clinical course from the more usual presentation of CAMT type I (Table 1). This cannot be easily explained as both the nonsense mutations usually associated with CAMT type I and the one reported here would be predicted not to produce any functional protein; therefore, there would be an absence of receptors, which would impact on the downstream signaling.

In retrospect, Family 2 could be considered to be similar to type II CAMT but this diagnosis only became possible after the homozygous missense mutation had been identified. The age of onset is more variable in type II CAMT and presentation tends to be less severe. Clinically, type II CAMT can be defined by platelet count within the first year of life as transiently increasing to over 50×10^9/L or by developing pancytopenia over the age of six years.

Short telomeres are also a recognized feature of patients with AA.[20] Mutations in genes encoding the telomerase complex result in short telomeres and some of these have been described in patients with apparently acquired severe AA.[21,22] Telomere lengths were assayed to determine if short telomeres were involved in the pathology of AA caused by *MPL* mutations. Due to technical problems, we were only able to assay telomeres from both affected individuals from Family 2. For comparison, telomere lengths reported as T/S ratios are included for healthy controls (n=95, mean 0.83; range 0.32-1.86) and for patients with *TERC* mutations (n=18, mean 0.36; range 0.27-0.86). The T/S ratios for the patients with *MPL* mutations are above the 10th percentile for the normal range (Figure 2) and longer than those seen in patients with *TERC* mutations (which are significantly shorter than controls). This suggests that the AA associated with *MPL* mutations is due more to a reduction of stem cells, resulting in an inability to maintain the pool of blood cells, rather than short telomeres causing increased cellular apoptosis and/or senescence leading to the same reduction in blood cells, as seen in AA patients with *TERC* mutations.

In this study, we have demonstrated exome sequencing of a single individual from a consanguineous family is powerful enough to identify the causative recessive disease gene, particularly when coupled with previous linkage data. This reduces the number of potential variations, as any not in a previously identified homozygous region could be excluded. We identified the causative homozygous *MPL* mutation in a family with familial AA. This gene has been implicated in a similar disease phenotype CAMT, but due to the differences in presentation and course of the disease this diagnosis was not considered. An additional novel *MPL* homozygous mutation was also identified in another family following screening of 33 AA cases. We have shown for the first time that homozygous *MPL* mutations can be found (albeit rarely) in children with familial AA in whom CAMT was not diagnosed or suspected. Additional studies will be needed to further clarify the relationship between CAMT, AA and MPL.

Authorship and Disclosures

The information provided by the authors about contributions from persons listed as authors and in acknowledgments is available with the full text of this paper at www.haematologica.org.

Financial and other disclosures provided by the authors using the ICMJE (www.icmje.org) Uniform Format for Disclosure of Competing Interests are also available at www.haematologica.org.

References

1. Young NS, Scheinberg P, Calado RT. Aplastic anemia. Curr Opin Hematol. 2008;15(3):162-8.
2. Dokal I. Inherited aplastic anaemia/bone marrow failure syndromes. In: Hoffbrand AV, Catovsky, D. Tuddenham E.G.D and Green, A.R, ed. Postgraduate Haematology (ed 6th): Wiley-Blackwell; 2011:186-205.
3. Marsh J and Young N. Acquired aplastic anaemia. In: Hoffbrand AV, Catovsky, D. Tuddenham E.G.D and Green, A.R, ed. Postgraduate Haematology (ed 6th): Wiley-Blackwell; 2011:206-25.
4. Dokal I, Vulliamy T. Inherited bone marrow failure syndromes. Haematologica. 2010;95(8):1236-40.
5. Knight SW, Vulliamy TJ, Heiss NS, Matthijs G, Devriendt K, Connor JM, et al. 1.4 Mb candidate gene region for X linked dyskeratosis congenita defined by combined haplotype and X chromosome inactivation analysis. J Med Genet. 1998;35(12):993-6.
6. Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science. 1987;236(4808):1567-70.
7. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet. 2008;9387-402.
8. Kuhlenbaumer G, Hullmann J, Appenzeller S. Novel genomic techniques open new avenues in the analysis of monogenic disorders. Hum Mutat. 2011;32(2):144-51.
9. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nuc Acids Res. 2009;37(3):e21.
10. Walne AJ, Vulliamy T, Beswick R, Kirwan M and Dokal I. Mutations in C16orf57 and normal-length telomeres unify a subset of patients with dyskeratosis congenita, poikiloderma with neutropenia and Rothmund-Thomson syndrome. Hum Mol Genet. 2010;19(22):4453-61.
11. Kaushansky K. Molecular mechanisms of thrombopoietin signaling. J Thromb Haemost. 2009;7(Suppl 1)235-8.
12. Ihara K, Ishii E, Eguchi M, Takada H, Suminoe A, Good RA, et al. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. Proc Natl Acad Sci USA. 1999;96(6):3132-6.
13. Chou FS, Mulloy JC. The thrombopoietin/MPL pathway in hematopoiesis and leukemogenesis. J Cell Biochem. 2011;112 (6):1491-8.
14. Savoia A, Dufour C, Locatelli F, Noris P, Ambaglio C, Rosti V, et al. Congenital amegakaryocytic thrombocytopenia: clinical and biological consequences of five novel mutations. Haematologica. 2007;92 (9):1186-93.
15. Germeshausen M, Ballmaier M, Welte K. MPL mutations in 23 patients suffering from congenital amegakaryocytic thrombocytopenia: the type of mutation predicts the course of the disease. Hum Mutat. 2006;27(3):296.
16. Ballmaier M, Germeshausen M, Schulze H, Cherkaoui K, Lang S, Gaudig A, et al. c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. Blood. 2001;97(1):139-46.
17. Geddis AE. Congenital amegakaryocytic thrombocytopenia. Pediatr Blood Cancer. 2011;57(2):199-203.
18. Al-Qahtani FS. Congenital amegakaryocytic thrombocytopenia: a brief review of the literature. Clin Med Insights Pathol. 2010;3:25-30.
19. Ballmaier M, Germeshausen M. Advances in the understanding of congenital amegakaryocytic thrombocytopenia. Br J Haematol. 2009;146(1):3-16.
20. Ball SE, Gibson FM, Rizzo S, Tooze JA, Marsh JC, Gordon-Smith EC. Progressive telomere shortening in aplastic anemia. Blood. 1998;91(10):3582-92.
21. Yamaguchi H, Calado RT, Ly H, Kajigaya S, Baerlocher GM, Chanock SJ, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. N Engl J Med. 2005;352(14):1413-24.
22. Marrone A, Sokhal P, Walne A, Beswick R, Kirwan M, Killick S, et al. Functional characterization of novel telomerase RNA (TERC) mutations in patients with diverse clinical and pathological presentations. Haematologica. 2007;92(8):1013-20.