bjh short report

Thalassaemia intermedia caused by coinheritance of a β-thalassaemia mutation and a *de novo* duplication of α-globin genes in the paternal allele

Dejian Pang,[1,*] Xuan Shang,[1,2,3,*] Decheng Cai,[1] Fei Zhu,[1] Yi Cheng,[1] Jianmei Zhong,[1] Sheng Yi,[4] Qianqian Zhang[1] and Xiangmin Xu[1,2,3]

[1]Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, [2]Guangdong Genetics Testing Engineering Research Centre, [3]Guangdong Provincial Key Laboratory of Single Cell Technology and Application, Guangzhou, Guangdong and [4]Prenatal Diagnostic Centre, Guangxi Zhuang Autonomous Region Women and Children Care Hospital, Nanning, Guangxi, China

Received 1 January 2019; accepted for publication 5 March 2019

Correspondence: Xiangmin Xu, Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, Guangdong, China.
E-mail: xixm@smu.edu.cn

*These authors contributed equally to this study.

Summary

Next generation sequencing identified a *de novo*, 204 kb, tandem duplication ($\alpha\alpha\alpha\alpha^{204}$) in the α-globin gene cluster of a Chinese thalassaemia intermedia patient. Haplotype analysis showed that the duplicated chromosome was of paternal origin. Molecular analysis of genomic DNA from the patient's lymphocytes, hair follicles, buccal mucosa cells, his father's lymphocytes and sperm cells excluded the possibility of somatic or germinal mosaicism. The analysis also indicated that this duplication arose during spermatogenesis. The microhomology in the breakpoint was found and suggested that this duplication could be formed by a coupled homologous and non-homologous recombination mechanism.

Keywords: *de novo* mutation, α-globin gene, thalassaemia intermedia, spermatogenesis, recombination.

In the clinic, researchers and clinicians sometimes detect thalassaemia patients whose inheritance mode contradicts Mendelian laws. Many such patients can be explained by two mechanisms: (i) β-thalassaemia carriers can co-inherit a duplication of α-globin (*HBA*) genes that results in thalassaemia intermedia (TI), or (ii) A *de novo* point mutation or large deletions combined with another mutation, inherited from the parents, which lead to thalassaemia (Hu *et al*, 2016; Shang & Xu, 2017). However, the origin of the *de novo* mutation is rarely analysed and there are still a small number of cases that are difficult to explain. Previous studies have indicated that *de novo* duplications are not rare at the cellular level but those found in the α-globin gene cluster are never reported as the real cause of thalassaemia (Lam & Jeffreys, 2007). Here, we report the first discovery of a thalassaemia case related to a *de novo* large duplication. The patient reported here had TI and was heterozygous for the *HBB*:c.316-197C>T β-thalassaemia variant plus a *de novo* duplication of approximately 204 kb in the α-globin gene cluster ($\alpha\alpha\alpha\alpha^{204}$). A possible mechanism for the generation of this duplication was also explored.

Patient and methods

The patient was a 6-year-old boy from Lianzhou City in Guangdong Province, southern China. He was diagnosed as having TI due to moderate anaemia (Hb: 67 g/l, Hb A_2: 3·4%, Hb F: 12·5%; Fig 1A), but no hepatosplenomegaly (Karimi *et al*, 2014) and had received a total of two blood transfusions before reaching 6 years of age. Iron deficiency was excluded. Detection of common α-globin gene (*HBA*) mutations and 17 known β-globin gene (*HBB*) mutations, found in the Chinese population, identified the patient as heterozygous for the *HBB*: c.316-197C>T β-thalassaemia variant, which was inherited from his father. All members of this family were referred to our laboratory for further study with approval from the Ethics

doi: 10.1111/bjh.15958

Committee of Nanfang Hospital of Southern Medical University and with informed consent. Peripheral blood samples were collected from all members of the family and haematological parameters were analysed as previously described (Hu *et al*, 2016). The α/preβ+β globin chain ratio of all subjects' peripheral blood samples were analysed by reversed-phase high-performance liquid chromatography as described in our previous study (Wan *et al*, 2012). The patient's hair follicle, buccal mucosa and semen samples from the father were also obtained with informed consent.

Genomic DNA was isolated using the TIANamp Micro DNA Kit (DP316; TIANGEN Biotech, Beijing, China). Analysis of *HBB* mutations was performed by direct DNA sequencing. Copy number variations (CNVs) of the α- and β-globin gene clusters were analysed by multiplex ligation dependent probe amplification (MLPA) (Hu *et al*, 2016). Next generation sequencing (NGS) was also used to confirm the *HBA* rearrangement region and other mutations of modifier genes (Shang *et al*, 2017). To verify the duplication breakpoints, three pairs of primers were designed upstream and downstream of this rearrangement region, based on the NGS results, as follows (Fig 1C):

F1 (5′-GCAAAGACCCTCAGGACACA-3′),
R1 (5′-TACAGCGTGGTGGGAATGAC-3′),
F2 (5′-CTGGTAGTTAGAGTGGGATGA-3′),
R2 (5′-ACAGCAGCACAAAATGGA-3′),
F3 (5′-GGCTGTTGTTAATAATGAACCAAT-3′)
R3 (5′-CTGTAATCGCAGCACTCTGG-3′)

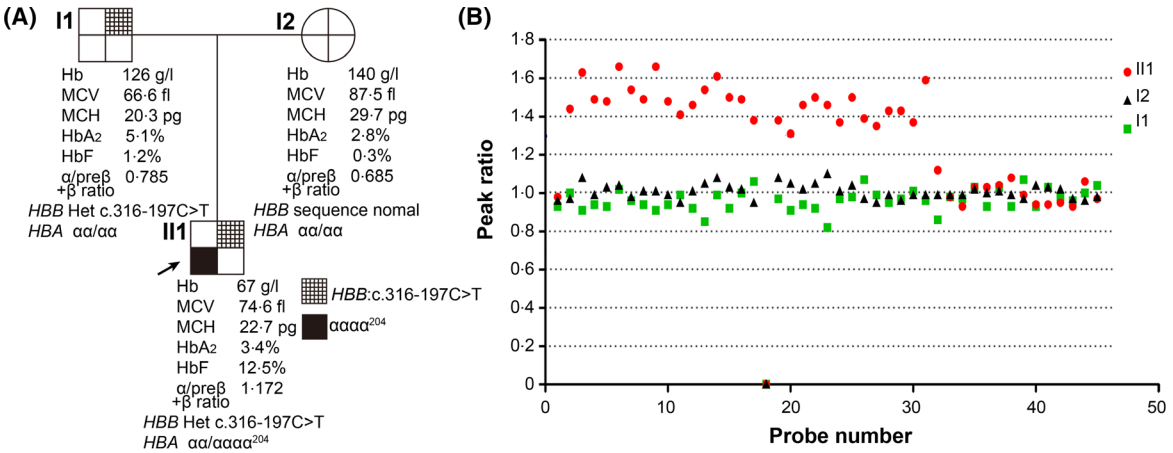
Six single nucleotide polymorphisms (SNPs) were detected near the breakpoint and a haplotype analysis was performed in order to confirm the source of chromosome duplication (details in Appendix S1). To exclude germline mosaicism, the genomic DNA from the fathers' semen sample (I1) was amplified with primers F2 and R2. To confirm paternity and maternity, 19 short tandem repeat loci and a segment of the X-Y homologous amelogenin gene (*AMEL*) were co-amplified using the Microreader™ 20A ID System (Microread Genetics Incorporation, Suzhou, China).

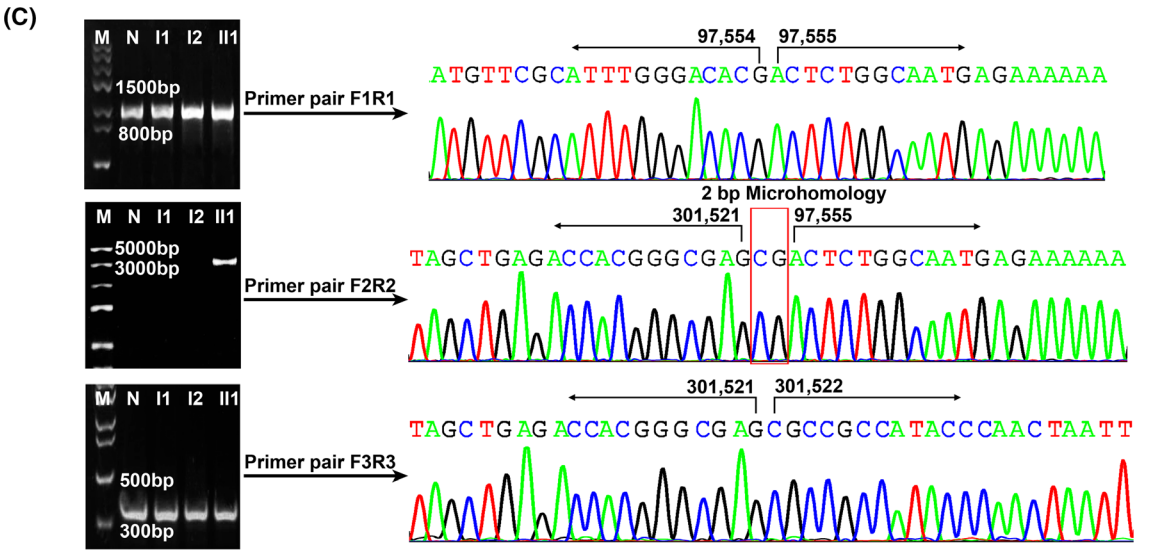
Results and discussion

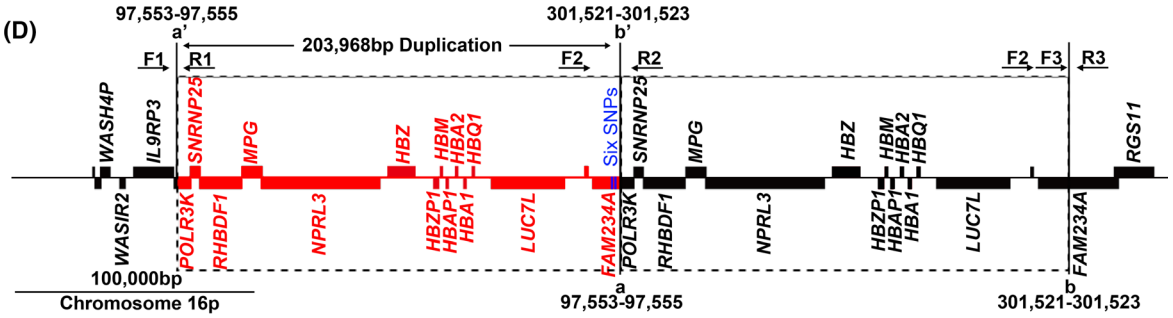
The phenotype and genotype data from this family is summarized in Fig 1A. The haemoglobin chain analysis showed that the α/preβ+β globin chain ratio of the patient (II1) was greater than that of his parents, which is consistent with their phenotype. However, in normal individuals, the phenotypic effect of this duplication on the mean corpuscular haemoglobin and mean corpuscular volume is unclear due to the absence of a heterozygous duplication. Previous reports in heterozygotes with comparable duplications suggest a normal haematological phenotype (Liu *et al*, 2015; Hu *et al*, 2016; Clark *et al*, 2018).

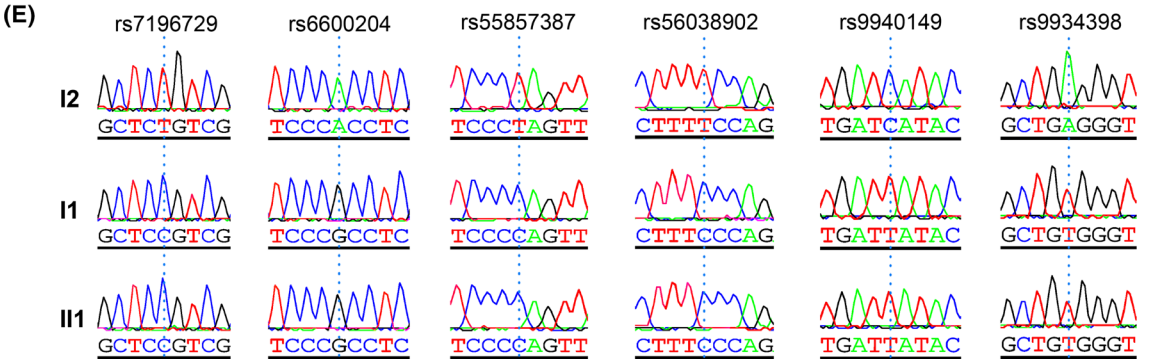
Direct sequencing of *HBB* gene in the patient identified a single common β-thalassaemia variant; *HBB*:c.316-197C>T, which was inherited from his father. A large novel duplication, which includes the *HBA* genes, was detected by MLPA analysis but the range could not be determined (Fig 1B). NGS analysis estimated the duplication size to be approximately 200 kb and the position to be chr16: 97 615–299 525 (hg19). The breakpoint positions were further confirmed by a strategy that has been described previously (Hu *et al*, 2016). The F1 and R1 primers generated a 950 bp product. The F2 and R2 primers generated a 3500 bp product and the F3 and R3 primers generated a 350 bp product. Sequencing of these products revealed a 203 968 bp tandem duplication segment from positions 97 553–97 555 to 301 521–301 523 (Fig 1C, D). This duplication increases the copy number of the α-globin genes from four to six in this patient, which explains the intermediate β-thalassaemia status. It arose as a *de novo* duplication in this patient as neither parent carries this variation and non-paternity was excluded by a paternity test (Table SI). Duplication in the α-globin gene cluster can easily be ignored as the cause of β-thalassaemia, unless it is co-inherited with a defect of the *HBB* gene. A few groups of duplications in the α-globin gene cluster have previously been reported (Figure S1). Most of the studies have been unable to verify the precise breakpoints due to previous technical limits. Our findings, and those from other studies (Hu *et al*, 2016; Clark *et al*, 2017), have proved that combinations of MLPA, array comparative genomic hybridization (array CGH) and NGS assays are very useful for identifying the known or unknown rearrangements in α- or β-globin gene clusters. This duplication not only widens the spectrum of CNVs in globin genes but it is also the first report of a *de novo* large duplication, although some *de novo* point mutations have previously been found at this locus (Shang & Xu, 2017).

Fig 1. The family pedigree and identification of the duplication. (**A**) Pedigree of the Lianzhou family. The patient is indicated by an arrow. The phenotypic and genotypic data is listed. (**B**) Multiplex ligation dependent probe amplification results for the family. Scattering of II1 (red circles) shows that the duplication region is between the site of probe 2 and probe 31 (chr16:163 528–289 854, hg19). This is not observed in I1 and I2. (**C**) Identification of the breakpoints of the duplication. The breakpoint was amplified by primers F2 and R2. A unique 1300 bp product was amplified in II1 but not in I1, I2 or the normal control. The sequencing result shows that the duplication starts between position 97 553 and 97 555 and ends between 301 521 and 301 523 (*middle*). Polymerase chain reaction was performed using the primers F1R1 and F3R3. (**D**) Schematic representation of the duplication that involves the α-globin gene cluster in II1. The 203 968 bp tandem duplication, from **a′** to **b′,** was found between positions 97 553–97 555 and 301 521–301 523. (**E**) Haplotype analysis of six single nucleotide polymorphisms near the breakpoint shows that the duplicated chromosome was derived from the father (I1). Hb, haemoglobin; Het, heterozygous; MCH, mean corpuscular haemoglobin; MCV, mean corpuscular volume.









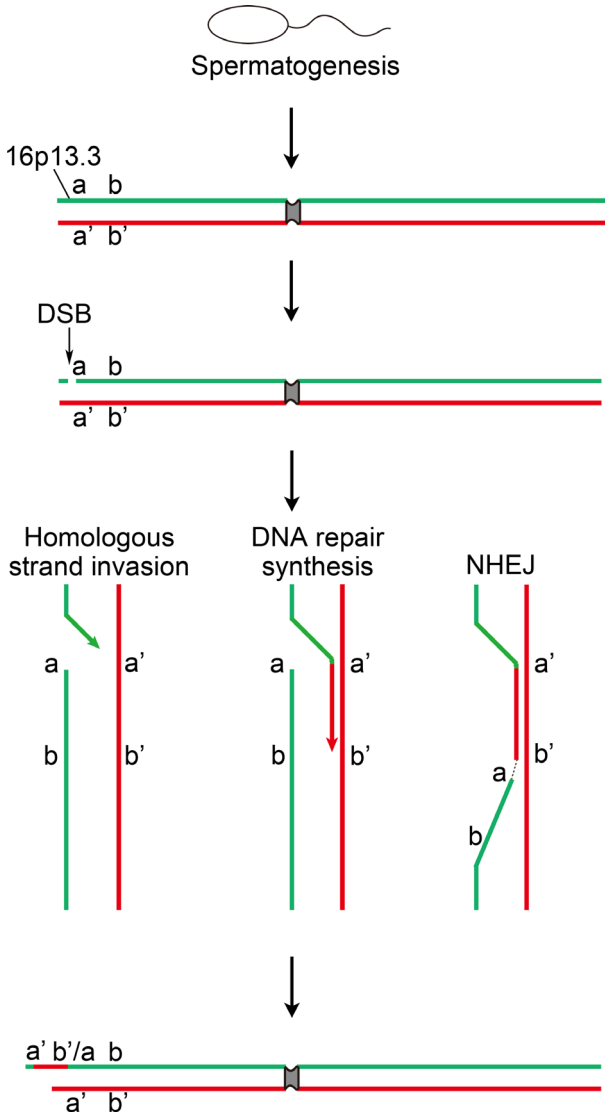


Fig 2. Model for partial tandem duplication of the α-globin locus. The rearrangement originally occurred on chromosome 16p during spermatogenesis in I1. The event was initiated by a double-strand break (DSB) within or near the duplication fragment during the period of meiosis. To one side of the break, the strand of one end invades into the homologous element on the intrachromosomal sister chromatid and DNA synthesis extends beyond the site of the DSB for the length of the duplication region (**a′b′**), using the invasive homologous element as the template. Lastly, the extended end, **b′,** is re-joined by nonhomologous end joining (NHEJ) to the free end, **a′,** of the DSB.

Analysis of DNA obtained from lymphocytes from the patient was consistent with the testing of his hair root and oral mucosal exfoliated cells (Figure S2A). This excludes the possibility of a somatic mosaicism and indicates that the mutation could originate from the germline. Genotyping data from six SNPs near the breakpoint showed that the duplicated chromosome was derived from the father (Fig 1E). Further study excluded germline mosaicism as we could not detect the duplication in germ cells from the father (Figure S2B). It is, therefore, reasonable to suggest that the cause of the duplication was an event during spermatogenesis that produced a mutant sperm. At the cellular level, Lam and Jeffreys (2007) found that *de novo* duplication is not rare in sperm and the $\alpha\alpha\alpha^{anti3.7}$ duplication often arises due to ectopic recombination between homologous chromosomes. But this finding was not proved at the individual level. Our research highlighted a clinical case with TI caused by coinheritance of a β-thalassaemia mutation and a *de novo* duplication of *HBA* genes in the paternal allele. However, the generation mechanism for $\alpha\alpha\alpha\alpha^{204}$ is inconsistent with $\alpha\alpha\alpha^{anti3.7}$, because only 2 bp (CG) microhomology (Fig 1D) was found at the breakpoints of $\alpha\alpha\alpha\alpha^{204}$. In contrast, non-allelic homologous recombination usually requires up to 200 bp of sequence homology between proximal and distal breakpoints (Liskay *et al*, 1987; Hastings *et al*, 2009). Another mechanism of coupled homologous and non-homologous recombination, by repair of a double-strand break (DSB), may provide a more suitable explanation for this tandem duplication (Richardson & Jasin, 2000). In general, sperm have a higher chance of acquiring DNA damage than other cells, as sperm divide continuously throughout their reproductive lifetime and the testicular environment is more prone to the toxic effects of oxidative stress (Crow, 2000; Sharma *et al*, 2015). Hence, a model based on this mechanism is proposed to describe the origin of this tandem duplication (Fig 2), in accordance with previous research on Pelizaeus-Merzbacher Disease (Woodward *et al*, 2005). Due to the 2 bp microhomology between the proximal and distal breakpoints of $\alpha\alpha\alpha\alpha^{204}$, non-homologous end joining (NHEJ) is considered as the most appropriate repair paths to create tandem duplication in chromosomes because this model requires one to four base pairs or even no homology between proximal and distal breakpoints.

In summary, we have identified a *de novo* duplication by NGS and have analysed the mechanism by which it was generated. This is the first clinical case to indicate that *de novo* duplications of the α-globin gene cluster in sperm could be generated by coupled homologous and non-homologous recombination. This data may be valuable in providing accurate diagnosis and genetic counselling in similar cases.

Acknowledgements

This study was supported by research funding from the National Key R&D Programme of China (2018YFA0507800; 2018YFA0507803) and National Natural Science Foundation of China (81870148).

Author contribution

D.P, D.C, F.Z, Y.C, J.Z, S.Y and Q.Z collected samples and performed the research. D.P, X.S and X.X designed the study, analysed the data and wrote the paper. X.X supervised the research. All authors reviewed, edited and approved the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Summary of duplications in the α-globin locus, taken from the literature.

Figure S2. Mosaicism analysis of the $\alpha\alpha\alpha\alpha^{204}$ duplication. MLPA analysis of the hair follicle and oral mucosa from II1 (A), detection of duplication in the sperm from I1 (B)

Table SI. Analysis of paternity test.

Appendix S1. Patient and methods.

References

Clark, B.E., Shooter, C., Smith, F., Brawand, D. & Thein, S.L. (2017) Next-generation sequencing as a tool for breakpoint analysis in rearrangements of the globin gene clusters. *International Journal of Laboratory Hematology*, **39**, 111–120.

Clark, B., Shooter, C., Smith, F., Brawand, D., Steedman, L., Oakley, M., Rushton, P., Rooks, H., Wang, X., Drousiotou, A., Kyrri, A., Hadjigavriel, M., Will, A., Fisher, C., Higgs, D.R., Phylipsen, M., Harteveld, C., Kleanthous, M. & Thein, S.L. (2018) Beta thalassaemia intermedia due to co-inheritance of three unique alpha globin cluster duplications characterised by next generation sequencing analysis. *British Journal of Haematology*, **180**, 160–164.

Crow, J.F. (2000) The origins, patterns and implications of human spontaneous mutation. *Nature Reviews. Genetics*, **1**, 40–47.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M. & Ira, G. (2009) Mechanisms of change in gene copy number. *Nature Reviews. Genetics*, **10**, 551–564.

Hu, L., Shang, X., Yi, S., Cai, R., Li, Z., Liu, C., Liang, Y., Cai, D., Zhang, F. & Xu, X. (2016) Two novel copy number variations involving the α-globin gene cluster on chromosome 16 cause thalassemia in two Chinese families. *Molecular Genetics and Genomics*, **291**, 1443–1450.

Karimi, M., Cohan, N., De Sanctis, V., Mallat, N.S. & Taher, A. (2014) Guidelines for diagnosis and management of Beta-thalassemia intermedia. *Pediatric Hematology Oncology*, **31**, 583–596.

Lam, K.W. & Jeffreys, A.J. (2007) Processes of *de novo* duplication of human alpha-globin genes. *Proceedings of the National Academy Sciences of the United States of America*, **104**, 10950–10955.

Liskay, R.M., Letsou, A. & Stachelek, J.L. (1987) Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics*, **115**, 161–167.

Liu, S., Jiang, H., Wu, M.-Y., Zhang, Y.-L. & Li, D.-Z. (2015) Thalassemia intermedia caused by 16p13.3 sectional duplication in a β-thalassemia heterozygous child. *Pediatric Hematology and Oncology*, **32**, 349–353.

Richardson, C. & Jasin, M. (2000) Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. *Molecular and Cell Biology*, **20**, 9068–9075.

Shang, X. & Xu, X. (2017) Update in the genetics of thalassemia: What clinicians need to know. *Best Practice & Research. Clinical Obstetrics Gynaecology*, **39**, 3–15.

Shang, X., Peng, Z., Ye, Y., Asan, Zhang, X., Chen, Y., Zhu, B., Cai, W., Chen, S., Cai, R., Guo, X., Zhang, C., Zhou, Y., Huang, S., Liu, Y., Chen, B., Yan, S., Chen, Y., Ding, H., Yin, X., Wu, L., He, J., Huang, D., He, S., Yan, T., Fan, X., Zhou, Y., Wei, X., Zhao, S., Cai, D., Guo, F., Zhang, Q., Li, Y., Zhang, X., Lu, H., Huang, H., Guo, J., Zhu, F., Yuan, Y., Zhang, L., Liu, N., Li, Z., Jiang, H., Zhang, Q., Zhang, Y., Juhari, W., Hanafi, S., Zhou, W., Xiong, F., Yang, H., Wang, J., Zilfalil, B.A., Qi, M., Yang, Y., Yin, Y., Mao, M. & Xu, X. (2017) Rapid targeted next-generation sequencing platform for molecular screening and clinical genotyping in subjects with hemoglobinopathies. *EBioMedicine*, **23**, 150–159.

Sharma, R., Agarwal, A., Rohra, V.K., Assidi, M., Abu-Elmagd, M. & Turki, R.F. (2015) Effects of increased paternal age on sperm quality, reproductive outcome and associated epigenetic risks to offspring. *Reproductive Biology and Endocrinology*, **13**, 35.

Wan, J.H., Tian, P.L., Luo, W.H., Wu, B.Y., Xiong, F., Zhou, W.J., Wei, X.C. & Xu, X.M. (2012) Rapid determination of human globin chains using reversed-phase high-performance liquid chromatography. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences*, **901**, 53–58.

Woodward, K.J., Cundall, M., Sperle, K., Sistermans, E.A., Ross, M., Howell, G., Gribble, S.M., Burford, D.C., Carter, N.P., Hobson, D.L., Garbern, J.Y., Kamholz, J., Heng, H., Hodes, M.E., Malcolm, S. & Hobson, G.M. (2005) Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. *American Journal of Human Genetics*, **77**, 966–987.

