

Article

Clinical application of a protocol based on universal next-generation sequencing for the diagnosis of beta-thalassaemia and sickle cell anaemia in preimplantation embryos



Nada Kubikova ^{a,*}, Dhruvi Babariya ^b, Jonas Sarasa ^c, Katharina Spath ^b,
Samer Alfarawati ^{b,1}, Dagan Wells ^a

^a University of Oxford, Nuffield Department of Women's and Reproductive Health, John Radcliffe Hospital, Level 3, Women's Centre, Oxford, OX3 9DU, UK

^b CooperGenomics, Institute of Reproductive Sciences, Oxford Business Park North, Alec Issigonis Way, Oxford OX4 2HW, UK

^c iGLS, C/Britania 7, Alicante, 03540, Spain



Nada obtained an MSc degree in Clinical Embryology from the University of Oxford. She is currently a DPhil student at the Nuffield Department of Women's and Reproductive Health, University of Oxford. Her research focuses on application of new technologies for detection, diagnosis and correction of disease in preimplantation embryos.

KEY MESSAGE

For monogenic diseases displaying high degree of mutation diversity, like beta-thalassaemia, traditional preimplantation genetic testing protocols require numerous distinct tests specific to individual couples and time-consuming and expensive customization. These limitations can be overcome using next-generation sequencing technology, sequencing the entire disease-causing gene along with closely linked polymorphic markers.

ABSTRACT

Research question: Mutations of the beta-globin gene (*HBB*) cause beta-thalassaemia and sickle cell anaemia. These are the most common cause of severe inherited disease in humans. Traditional preimplantation genetic testing protocols for detecting *HBB* mutations frequently involve labour intensive, patient-specific test designs owing to the wide diversity of disease-associated *HBB* mutations. We, therefore, asked the question whether a universally applicable preimplantation genetic testing method can be developed to test for *HBB* gene mutations.

Design: A multiplex polymerase chain reaction protocol was designed, allowing simultaneous amplification of multiple overlapping DNA fragments encompassing the entire *HBB* gene sequence in addition to 17 characterized, closely linked single nucleotide polymorphisms (SNP). Amplicons were then analysed using a next-generation sequencing method, revealing mutations and SNP genotypes. The protocol was extensively validated, optimized and eventually clinically applied on whole-genome amplified DNA derived from embryos of three couples carrying different combinations of beta-thalassaemia mutations.

Results: The *HBB* mutation status and associated SNP haplotypes were successfully determined in all 21 embryos. Analysis of 141 heterozygous sites showed no instances of allele dropout and the test displayed 100% concordance compared with the results obtained from karyomapping. This

* Corresponding author.

E-mail address: nada.kubikova@bnc.ox.ac.uk (N Kubikova).

¹ Present address: Oxford Science Park, Magdalen Centre, Robert Robinson Ave, Oxford OX4 4GA, UK.

<https://doi.org/10.1016/j.rbmo.2018.05.005>

suggests that the combination of trophectoderm biopsy and highly sensitive next-generation sequencing may provide superior accuracy than typically achieved using traditional preimplantation genetic testing methods. Importantly, no patient-specific test design or optimization was needed.

Conclusions: It is hoped that protocols that deliver almost universally applicable low-cost tests, without compromising diagnostic accuracy, will improve patient access to preimplantation genetic testing, especially in less affluent parts of the world.

© 2018 Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.

Introduction

Preimplantation genetic testing (PGT) was developed for couples at high risk of transmitting an inherited genetic disorder to their offspring, providing them with an alternative to prenatal diagnosis and termination of affected pregnancy (Handyside et al., 1990). The most common reason for referral for PGT of a single gene disorder is the desire to avoid conditions affecting haemoglobin synthesis, caused by mutations in the *HBB* gene (Moutou et al., 2014). These disorders represent the most common class of gene defect worldwide. Sickle cell anaemia, a consequence of the Glu6Val point mutation in the *HBB* gene, which results in production of abnormal haemoglobin S, affects about 1 in 500 African-Americans and, in some parts of West Africa, carrier frequencies within the population are estimated to be as high as one in five (Agasa et al., 2010; Tshilolo et al., 2008). Another disorder caused by *HBB* mutation, beta-thalassaemia, is an autosomal recessive disorder common across a broad swathe of the world, from West Africa and the Mediterranean, through the Middle East and across Central and South East Asia (Angastiniotis and Modell, 1998). A wide spectrum of *HBB* mutations are associated with beta-thalassaemia, giving rise to significant phenotypic heterogeneity (Cao and Galanello, 2010).

Demand for PGT of inherited *HBB* defects is increasing, but the vast diversity of potential affected genotypes presents a problem for traditional PGT methods. These typically involve creation of individual protocols, each customized for the detection of a specific combination of mutations. Such strategies have mostly relied on polymerase chain reaction (PCR) methods capable of simultaneously amplifying DNA fragments encompassing the mutation site(s) as well as DNA polymorphisms located in close proximity to the mutant gene. Identification of informative polymorphic markers, PCR primer design, optimizations of single-cell multiplex PCR to ensure all DNA fragments are sufficiently amplified for subsequent analysis and protocol validation, require significant investments of time and resources. Frequently, combinations of mutations and informative polymorphisms are unique to an individual family, meaning that the protocol eventually developed, following extensive work in the laboratory, may only be of use for a single couple.

Genetic methodologies are evolving rapidly and offer the possibility of delivering PGT with higher accuracy and lower cost than was possible using earlier techniques. Chief among the new wave of methods is next-generation sequencing (NGS). This technology has already been adapted for use in assisted reproductive technology, providing a robust platform for the detection of aneuploidy in embryo biopsy specimens (Fiorentino et al., 2014a, 2014b; Wells et al., 2014; Zheng et al., 2015). Recently, NGS has also been used for the diagnosis of single gene mutations in embryos, although the protocols used clinically still require significant customization in designing and optimizing multiplex-PCR for the amplification of mutation sites and polymorphisms (Chen et al., 2016; Ren et al., 2016; Treff et al., 2013;

Yan et al., 2015). Here, we describe a novel PGT protocol, based on NGS technology, for the detection of virtually all mutations responsible for beta-thalassaemia and sickle cell anaemia. This provides a single cost-effective method applicable to most couples seeking PGT for these conditions.

Materials and methods

Patients, IVF and embryo biopsy

Before clinical implementation of the protocol on embryo biopsies, the technique was validated on genomic DNA obtained from five family trios (each composed of the mother, father and a child or prenatal sample) and two couples, together carrying 14 different *HBB* mutations. After protocol optimization, three of the families, all healthy carriers of beta-thalassaemia, used the test clinically for the purpose of PGT. In all three cases, the patients underwent ovarian stimulation, and oocytes were collected and fertilized using intracytoplasmic sperm injection. The resulting embryos were biopsied and vitrified at cleavage or blastocyst stages. In two cases, no other family members were available for testing, and one couple had a previous affected pregnancy; a sample of amniotic fluid was included in the study along with the DNA samples extracted from the parents. Analysis of this additional sample allowed the phase of linked polymorphisms to be determined, i.e. revealed which alleles were associated with parental mutations. All patients underwent PGT for beta-thalassaemia in different IVF centres and gave consent for NGS-based PGT to be carried out in parallel with karyomapping (a validated single nucleotide polymorphism SNP-array based technique) (Ben-Nagi et al., 2017; Giménez et al., 2015; Konstantinidis et al., 2015; Natesan et al., 2014). DNA obtained from parents and embryo biopsies was tested at Reprogenetics UK. The diagnosis of beta-thalassaemia in embryos has been previously licensed by the Human Fertilisation and Embryology Authority. The study was approved by Aspire IRB on 26 August 2015 (reference number PGSP-2015).

Genomic DNA extraction and whole-genome amplification

Genomic DNA was extracted from 4 ml parental blood and from an amniotic fluid sample using the QIAamp DNA Blood Mini Kit (Qiagen, Germany), using the standard protocols recommended by the manufacturer. Extracted DNA and cell(s) obtained from embryo biopsies were lysed as previous described (Thornhill et al., 2015) and subjected to whole-genome amplification with using a REPLI-g Single Cell kit (Qiagen) according to the manufacturer's instructions.

Multiplex polymerase chain reaction, library construction and next-generation sequencing

An aliquot of each multiple displacement amplification (MDA) product was further amplified to enrich the regions of interest in a multiplex

PCR reaction comprising eight pairs of primers designed to encompass the entire *HBB* gene locus (Table S1). As well as these amplicons, which covered all the amino acid coding regions of the gene and all splice donor and acceptor sites, an additional 12 PCR fragments were also amplified, encompassing 17 single nucleotide polymorphisms (SNP) flanking the *HBB* gene (Table S2). All the targeted SNP were known to have a high degree of heterozygosity (minor allele frequency values higher than 0.24), meaning a high likelihood of informativity for genetic linkage analysis in individual families. For this aspect of the design, we used the NCBI dbVAR database (<https://www.ncbi.nlm.nih.gov/dbvar>) in combination with the ENSEMBL genome browser (<https://www.ensembl.org/index.html>). The multiplex PCR amplification included 0.5 µl of MDA product in a total reaction volume of 50 µl, containing 22 µl of nuclease-free water, 25 µl of 2 x Qiagen master mix (Qiagen, Germany), and 2.5 µl of primer mix adjusted to the concentration of 2.5 µM (Eurogentec, UK). Amplifications were carried as follows: initial denaturation step at 95°C for 15 min, followed by 55 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 90 s, and extension at 72°C for 1 min, with a final extension step at 60° for 10 min. Alongside the clinical samples, we assessed three samples each consisting of a single buccal cell (of normal genotype) and three further samples each composed of five isolated cells from the same source. These were subjected to targeted PCR using the same amplification conditions, but without carrying out whole-genome amplification (WGA) before the multiplex PCR. This confirmed that the protocol could be used on embryo biopsy specimens directly, avoiding the additional cost of WGA.

The DNA sequencing libraries were prepared using Truseq PCR Free Kit (Illumina) according to the manufacturer's standard work flow. The steps included size selection and clean-up, end repair, adenylation of 3' ends, ligation of indexed adapters and normalization and pooling of the libraries. Cluster generation was carried out on a MiSeq (Illumina) instrument using paired end synthesis 500 cycles MiSeq Reagent kit v2 (Illumina). In parallel, MDA products were processed using karyomapping according to the standard manufacturer's protocol (Infinium Karyomapping Assay Protocol Guide, Illumina).

Data analysis

Sequence alignment files for all indexed libraries were obtained from the Illumina MiSeq Reporter software in BAM format. The BAM files were loaded into the Integrative Genomic Viewer (IGV, Broad Institute) to visualize the sequence alignment using the h19 human genome assembly as a reference. Regions of interest were inspected after creating region tracks and importing all the SNP and mutation positions into the software. On the basis of the total number of reads for that position, the percentage of bases that did not correspond to the reference nucleotide was determined. The sequences obtained from embryonic DNA were screened for heterozygous loci and compared with the parental DNA to assess allele dropout (ADO) and informativity of selected SNP. For mutations, the aligned sequences were screened for genotype call at every mutation position in each family. For both categories, the analysis included the determination of sequence coverage at positions of interest and the identification of unexpected variants and polymorphisms within the sequenced DNA, which could serve as additional linkage markers.

Results

Validation study using the new next-generation sequencing protocol: genotyping of familial beta-thalassaemia mutations and linkage analysis

All 14 different mutations and all 19 patient genotypes were successfully detected in the *HBB* gene in the five family trios (mother, father, child/prenatal sample) and the two couples. The protocol successfully delivered the complete sequence of the *HBB* gene for all samples tested. The parental genotypes obtained, as well as the mutation status of existing children/prenatal samples (when available) were concordant with the previously determined genotypes in all seven families (Figure 1).

In addition to the direct mutation detection, all 17 targeted SNP outside the *HBB* locus (no more than 14 kb away) were examined in addition to patient-specific SNP found within the *HBB* locus. The polymorphisms informative in each couple were determined, looking for SNP in which one parent was heterozygous and the other was homozygous. For the five families with existing children or prenatal samples, informative SNP were examined in the offspring's DNA and used to deduce which parental alleles were associated with the mutant and normal copies of the gene. An example of a family specific haplotype analysis based on the SNP genotyping data produced by the NGS protocol is provided in Figure 2.

HBB gene mutation detection in clinical samples

After the validation study, three of the families proceeded to clinical implementation of the newly designed protocol. In the first couple, the mother carried a substitution c.92 + 6T>C whereas the father carried a c.118C>T substitution. In the second couple, the mother carried a c.93 – 21G>A substitution and the father carried a c.316 – 106C>G substitution. The second couple had a previous affected pregnancy and a sample of amniotic fluid was analysed and confirmed as compound heterozygote for the two parental mutations tested. In the third couple, the mother carried an insertion in codons c.27_28insG, whereas the father carried a c.92 + 5G>C substitution. In case 1, a single blastomere was biopsied from cleavage stage embryos whereas in cases 2 and 3, about five cells from the trophectoderm were biopsied from blastocyst stage embryos. All samples were whole-genome amplified and the tests were run in parallel with a standard karyomapping protocol (Konstantinidis et al., 2015). All 21 embryo samples amplified after MDA and NGS analysis was successful for each of these. The read depth for the sequenced amplicons was >1000x in all cases except for embryo 9 where no reads were obtained and karyomapping analysis confirmed nullisomy of chromosome 11. The obtained genetic status at the mutation sites were concordant with the results acquired from karyomapping. Of note, application of the targeted amplification protocol to single buccal cells and clumps, followed by NGS, generated a mean coverage of beta-globin sequences of 888x (1061x for single cells samples and 716x for clumps of cells), which was equivalent to the coverage obtained from the clinical samples subjected to WGA before targeted PCR. These results suggest that it is possible to use this protocol on embryo biopsies directly, without compromising the read depth, eliminating WGA and thus further reducing costs.

Mutation analysis														
Mutation	HBB:c.17_18delC T	HBB:c.135delC	HBB:c.92+5G>C	HBB:c.126_129de ICTTT	HBB:c.93-21G>A	HBB:c.92G>A	HBB:c.316- 106C>G	HBB:c.118C>T	HBB:c.92+5G>C	HBB:c.27_28insG	HBB:c.118C>T	HBB:c.92+6T>C	HBB:c.316-106C>G	HBB:c.93-21G>A
Position	5248235	5247987	5248155	5247996	5248050	5248159	5247062	5248004	5248155	5248224	5248004	5248154	5247062	5248050
Sample														
F1 Male	G/-													
F1 Female		G/-												
F1 Son	G/G	G/G												
F2 Male			C/G											
F2 Female				A/-										
F2 Son			C/G	A/-										
F3 Male					C/T									
F3 Female						C/T								
F3 Daughter					C/C	C/T								
F4 Male							C/G							
F4 Female								A/G						
F4 Daughter							C/G	A/G						
F5 Male									C/G					
F5 Female										A+C				
F5 Prenatal Sample									C/G	A+C				
F6 Male											G/A			
F6 Female												G/A		
F7 Male													G/C	
F7 Female														T/C

Figure 1 – Genotypes for beta-thalassaemia mutations in families referred for preimplantation genetic testing. Each pair of mutations for which the parents are carriers is indicated in the top row with their respective annotations. The seven families are separated by different shades of grey fill. Carrier genotypes are indicated in white and the affected genotypes in black. Five of the seven families already had a child/prenatal sample available, which could be analysed in parallel with parental DNA samples, assisting in the validation of mutation detection and determination of the phase of linked single nucleotide polymorphisms.

Linkage analysis in clinical samples

In case 1, 10 informative SNP were identified and used to assist determination of embryo mutation status: five associated with the maternal mutation and five associated with the paternal mutation (Figure 3a). All SNP assessed by the protocol described here were located outside the *HBB* gene but within 14 kb of upstream and downstream distance. It is worth noting that, for case 1, although the

diagnoses of individual embryos obtained from targeted NGS were concordant with those obtained from the standard PGT procedure, the results from karyomapping were suboptimal, associated with low call rates for individual SNP (in 50% of the embryos) and high ADO rates (>40%) in all embryo samples tested, likely a consequence of suboptimal WGA after blastomere biopsy, degradation of the biopsy specimens, or both. In embryo 9, no diagnosis was possible owing to complete absence of chromosome 11 and in embryo 13 only paternal

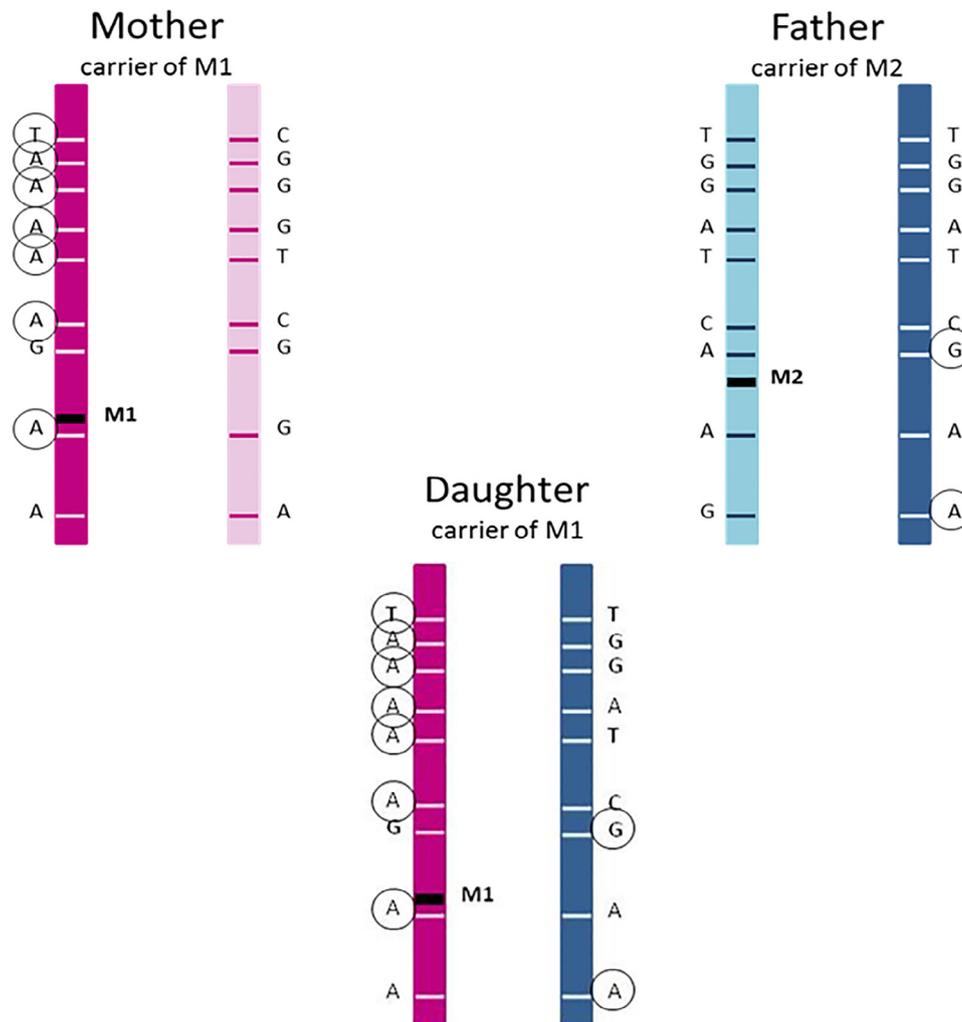


Figure 2 – Demonstration of the principle of linkage analysis and targeted mutation detection. For both parents, carrier states were confirmed by direct mutation detection. Linkage analysis identified nine informative single nucleotide polymorphism (SNP) near the *HBB* gene in this family. The father is heterozygous for two SNP for which the mother is homozygous (blue in circle) whereas the mother is heterozygous for seven SNP for which the father is homozygous (pink in circle). The inheritance of parental chromosomes in the daughter is thus inferred from nine SNPs as well as direct mutation detection. The daughter has inherited the affected maternal haplotype, associated with the M1 mutation (c.92G>A), and unaffected paternal haplotype, and is therefore a carrier of M1. This was further confirmed by the detection of the maternal c.92G>A mutation in the daughter.

alleles linked to *HBB* were detected. The karyomapping confirmed the presence of chromosome 11 monosomy, explaining this observation. Despite using the same low-quality WGA template, the targeted NGS showed excellent performance in all 13 embryo samples that had provided a karyomapping and NGS result, with no incidence of ADO observed in any of the informative SNP and mutation loci and >1000x sequencing coverage for the entire *HBB* gene and flanking polymorphisms. In case 2, 13 informative SNP were identified, out of which two were located within the *HBB* locus. Two of these were inherited with the paternal mutation, whereas the rest were inherited with the maternal mutation. For one embryo (trophectoderm sample 6) only paternal alleles for SNP linked to *HBB* were detected. Karyomapping analysis confirmed the presence of chromosome 11 monosomy (Figure 3b). In case 3, 10 informative SNP were analysed, two located within the *HBB* gene sequence itself (Figure 3c). Six of these were associated with the paternal mutation and four with the maternal mutation. Interestingly, three of the maternal SNPs were

not among those deliberately targeted by the protocol and appear to be rare variants, which are not present within the dbVAR database.

Surprisingly, of 141 heterozygous sites sequenced no instances of ADO occurred in any of the clinical samples tested using the NGS protocol, including the poor-quality WGA products from case 1. This contrasts with an ADO rate of around 12.5% for SNP genotyped using karyomapping in case 3 and exceeding 40% in case 1, and suggests that the targeted NGS approach is highly sensitive and of excellent diagnostic accuracy.

Discussion

Next-generation sequencing technology has developed rapidly in recent years and has found numerous research and clinical applications. In the context of PGT, NGS has principally been used for detecting

(A) Case 1	SNP ID	Position	Mother	Father	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	
Linked SNP	rs3813727	5255912	G/G	A/G	G/G	A/G	A/G	A/G	A/G	A/G	A/G	G/G	no reads / copy of chromosome 11 detected	G/G	G/G	G/G	G	A/G	
	rs4910736	5258989	A/A	A/C	A/A	A/C	A/C	A/C	A/C	A/C	A/C	A/A		A/A	A/A	A/A	A/A	C	A/C
	rs2105819	5259727	C/C	C/G	C/C	C/G	C/G	C/G	C/G	C/G	C/G	C/C		C/C	C/C	C/C	C/C	G	C/G
	rs7936823	5250168	A/A	A/G	A/A	A/G	A/G	A/G	A/G	A/G	A/G	A/A		A/A	A/A	A/A	A/A	G	A/G
	rs6578588	5252251	C/C	C/T	C/C	C/T	C/T	C/T	C/T	C/T	C/T	C/C		C/C	C/C	C/C	C/C	T	C/T
	rs7945118	5236417	C/G	G/G	C/G	C/G	C/G	G/G	C/G	C/G	C/G	C/G		C/G	C/G	G/G	G/G	G	C/G
	rs34220818	5236851	C/T	T/T	C/T	C/T	C/T	T/T	C/T	C/T	C/T	C/T		C/T	C/T	T/T	T/T	T	C/T
	rs12364872	5244144	A/G	G/G	A/G	A/G	A/G	G/G	A/G	A/G	A/G	A/G		A/G	A/G	G/G	G/G	G	A/G
	rs7110263	5246512	T/G	G/G	T/G	T/G	T/G	G/G	T/G	T/G	T/G	T/G		T/G	T/G	G/G	G/G	G	T/G
	rs10837626	5244299	T/A	A/A	T/A	T/A	T/A	A/A	T/A	T/A	T/A	T/A		T/A	T/A	A/A	A/A	A	T/A
Female mutation	c.92+6T>C	5248154	G/A	A/A	G/A	G/A	G/A	A/A	G/A	G/A	G/A	G/A	G/A	A/A	A/A	A	G/A		
Male mutation	c.118C>T	5248004	G/G	G/A	G/G	G/A	G/A	G/A	G/A	G/A	G/A	G/G	G/G	G/G	G/G	A	G/A		
Diagnosis			C	C	MC	A	A	PC	A	A	A	MC	CA	MC	U	U	CA	A	
(B) Case 2	SNP ID	Position	Mother	Father	E4	E6	E7												
Linked SNP	rs4910543	5258827	C/G	C/C	C/G	C	C/C												
	rs4910735	5258852	A/G	A/A	A/G	A	A/A												
	rs4910544	5258856	A/T	A/A	A/T	A	A/A												
	rs4910736	5258989	A/C	A/A	A/C	A	A/A												
	rs2105819	5259727	C/G	C/C	C/G	C	C/C												
	rs11036364	5249004	A/G	A/A	A/G	A	A/A												
	rs6578588	5252251	C/T	C/C	C/T	C	C/C												
	rs10837620	5243559	G/A	G/G	A/G	G	G/G												
	rs12364872	5244144	G/A	A/A	A/A	A	G/A												
	rs10837628	5244404	A/G	A/A	A/G	A	A/A												
	rs10837631	5246356	T/A	T/T	A/T	T	T/T												
	rs7480526	5247733	C/A	A/A	A/A	A	C/A												
	rs63750628	5248281	G/G	G/A	G/A	A	G/G												
Female mutation	c.93-21G>A	5248050	T/C	C/C	C/C	C	T/C												
Male mutation	c.316-106C>G	5247062	G/G	G/C	G/C	C	G/G												
Diagnosis			C	C	PC	CA	MC												
(c) Case 3	SNP ID	Position	Mother	Father	TE1	TE2	TE4	TE5											
Linked SNP	rs713040	5248243	G/G	G/A	G/G	G/A	G/A	G/A											
	rs1609812	5247141	A/A	A/G	A/A	A/G	A/G	A/G											
	rs10837631	5246356	A/T	A/A	A/A	A/A	A/A	T/A											
	rs7110263	5246512	G/G	G/T	G/G	G/T	G/T	G/T											
	rs12574989	5246514	C/C	C/T	C/C	C/T	C/T	C/T											
	rs10837626	5244299	A/A	A/T	A/A	A/T	A/T	A/T											
	N/A	5243559	G/A	G/G	G/G	G/G	G/G	G/A											
	N/A	5243613	C/T	C/C	C/C	C/C	C/C	C/T											
	N/A	5236740	G/A	A/A	G/A	G/A	G/A	A/A											
	rs7945118	5236417	G/G	G/C	G/G	G/C	G/C	G/C											
Female mutation	c.27_28insG	5248224	A+C	A	A+C	A+C	A+C	A											
Male mutation	c.92+5G>C	5248155	C/C	C/G	C/C	C/G	C/G	C/G											
Diagnosis			C	C	MC	A	A	PC											

Figure 3 – Results of mutation and linkage analysis: (a) case 1; (b) case 2; (c) case 3. Single nucleotide polymorphisms where heterozygosity was detected in one parent but not the other were selected for identification of disease-associated alleles in the embryo whole-genome amplification products subjected to targeted multiplex amplification in three clinical cases. In green are the alleles associated with the paternal *HBB* mutation and in red are the alleles inherited together with the maternal *HBB* mutation. Black alleles represent the disease-free genotype. A, affected embryo; C, carrier; CA, chromosomally abnormal; E, embryo, MC, carrier of maternal mutation; N/A, not applicable, the single nucleotide polymorphism identifiers do not exist for these variants as they were not present within the dbVAR database; PC, carrier of paternal mutation; SNP, single nucleotide polymorphisms; U, unaffected embryo.

chromosome abnormality (Fiorentino et al., 2014a, 2014b; Wells et al., 2014; Zheng et al., 2015), although interest in its application for the diagnosis of inherited single gene disorders has been growing (Chen et al., 2016; Ren et al., 2016; Treff et al., 2013; Yan et al., 2015). In this study, we designed, optimized and implemented an NGS protocol for PGT of beta-thalassaemia and sickle-cell anaemia that is applicable to most couples at risk of having children affected by these disorders.

Preimplantation genetic testing strategies involving WGA of embryo biopsy specimens, followed by targeted re-amplification and NGS, have recently been reported, allowing successful diagnosis of monogenic disorders. Yan et al. (2015) and Ren et al. (2016) developed strategies for the preimplantation diagnosis of selected inherited disorders, combining direct mutation detection and linkage analysis, and reported live births after the clinical use of their methods. Although NGS-based approaches are expected to eventually replace conventional PGT methods, one of the major challenges with the current technology lies in the extensive test customization required before clinical application. In some cases, the effort required to create a custom-NGS protocol, involving targeted mutation detection and simultaneous analysis of linked polymorphisms, are substantial. Indeed, they may necessitate as much, or more, work than required for a conventional PGT protocol. This makes such approaches impractical for the diagnosis of rare disorders or conditions caused by a wide spectrum of mutations, as in such cases the protocols created after extensive laboratory work may only be applicable to a single family.

In a previous study, Chen et al. (2016) described an NGS-based PGT approach requiring less customization. This was achieved by using a specially designed sequence capture array followed by NGS to provide data on the genotype of over 24,000 SNP. The information gathered was subsequently used to assemble haplotypes, allowing diagnosis of embryos that inherited mutant copies of the *PKD2* gene from their parents based upon linkage analysis. This sort of NGS strategy is particularly useful because it delivers high fidelity, sensitivity, and throughput and, as it focuses on the inheritance of common polymorphisms rather than family specific mutations, requires little work-up for each case. One of the limitations of this approach, however is the cost associated with obtaining an adequate depth of sequencing when assessing a large number of genomic regions. In the present study, although the entire sequence of the *HBB* gene was sequenced at high depth, as well as additional multiple closely linked polymorphic sites, the total proportion of the genome investigated remained small, equivalent to less than 10 kb of DNA sequence. This allows simultaneous analysis of large numbers of samples in a single sequencing run, reducing costs. The potential throughput is also much greater compared with PGT methods currently in routine use, such as karyomapping.

Although PGT methods based upon linkage analysis have the advantage of providing a more generic approach with less patient-specific work-up, a challenge for such strategies is that they cannot be used in cases in which the phase of the SNP alleles cannot be determined, i.e. where it is unclear which alleles are located on the same chromosome as the mutant gene. Deduction of phase requires DNA samples from the patients requesting PGT, and also from additional family members who have been previously tested and are of known mutation status (close relatives such as the parents or children of the couple are ideal). Testing of polymorphisms in these extra samples allows the inheritance of specific alleles to be traced through the family, revealing those consistently associated with normal and mutant gene copies. A lack of DNA samples from close relatives is a common occurrence in PGT; in our experience, this is a consideration in about a

one-quarter of all referred cases. Sometimes, the couple are reluctant to discuss the fact that they are undergoing PGT with other members of the family, other times key relatives may be deceased or unavailable for other reasons, and on some occasions a patient may carry a de-novo mutation, not present in any relative. Another issue for PGT strategies based entirely on linkage analysis is the possibility that informative alleles, permitting the two chromosomal copies to be distinguished, may be difficult to find. This is a particular problem when offering PGT to consanguineous families as many parts on the genome may be identical owing to shared ancestry. In instances such as those outlined, PGT cannot be carried out without direct detection of the causative mutation(s) in the embryos produced.

In the present study, rather than targeting specific mutations, of which there are a large number in *HBB*, we created a method that provides information on the entirety of the coding region and splice junctions of the gene, as well as selected flanking sequences containing sites of common polymorphism. This approach permits direct detection of virtually all *HBB* gene mutations, effectively eliminating the requirement for patient-tailored test design. The ability to trace the inheritance of defective *HBB* genes by using multiple linked polymorphisms, as well as direct mutation detection, results in a highly redundant test, greatly increasing diagnostic accuracy. In the context of a PGT case, the mutation sites and polymorphisms can each provide their own independent diagnosis. Linkage analysis alone can potentially provide a reliable diagnosis in circumstances where parental mutations have not been identified before PGT being undertaken, in cases in which mutations are refractory to detection, or in instances where one of the mutation sites fails to amplify appropriately from the embryo biopsy sample as a result of ADO or other technical problems. The extreme proximity to the *HBB* gene of the SNP tested (≤ 14 kb in all cases) makes it highly unlikely that meiotic recombination would ever occur between the polymorphic sites and the sites of mutations (indeed this would be impossible for the intragenic SNP analysed). The diagnoses should be highly resistant to errors owing to problems caused by failure of individual loci to amplify, preferential amplification and ADO. During this investigation, however, considering the embryo biopsy samples with a normal number of chromosome 11 copies, no ADO was detected at any of the 141 heterozygous sites sequenced (0%). Rates of ADO after MDA are influenced by the number of cells within the biopsy specimen and other technical factors. Hou et al. (2015) used the same type of MDA protocol as used in the present study and detected a 12.5% ADO rate on single cells re-sequenced at 30x sequence depth. In the present study, the incidence of ADO observed after karyomapping was generally of a similar level, although more than 50% of SNP loci were affected in some samples. The fact that ADO was so low in the present study could be attributable to the high sequencing depth used ($>1000\times$). In theory, this should increase the sensitivity for the detection of alleles, which are substantially under-represented owing to extreme preferential amplification of the alternate allele in a heterozygous sample. Indeed, this was shown to be the case in some of the cleavage stage biopsied embryos analysed in case 1 (embryos 1, 2, 3 and 5), where preferential amplification of several alleles generated over 97% of total reads (Table S3). In addition to the 17 well-characterized polymorphisms specifically targeted by the PGT protocol described here, an extra three previously uncharacterized intragenic sequence variants/polymorphisms were detected (at positions chr11:5243559, chr11:5243613 and chr11:105236740). These provided a useful additional source of linkage data, further supplementing the diagnosis. The ability of NGS to detect novel polymorphisms and variants unique

to individual couples reduces the risk of encountering low informativity when using the method.

The use of WGA before targeted PCR and NGS, also meant that a resource of material was available for further testing if desired, e.g. repeat of the original PGT analysis. Furthermore, if desired, low-pass next generation sequencing of the WGA templates generated using multiple displacement amplification can be used to establish the cytogenetic status of the embryo (Wells et al., 2014). The combination of minimal work-up and high throughput provided by this protocol resulted in an extremely economical test. The issue of cost is of great relevance in this particular case, given the fact that many regions of the world where *HBB* mutations are of high prevalence are relatively resource-poor. The experience from the present study confirms that NGS can provide a rapid, streamlined and potentially cost-effective solution for couples seeking to use PGT to avoid genetic disease transmission. It is expected that, in the future, additional protocols similar to the one described here will be developed for the testing of other single gene disorders where mutation heterogeneity leads to problems for conventional PGT methods. Strategies for PGT based upon NGS technology have the potential to provide the lowest costs without compromising accuracy and are therefore likely to become increasingly popular.

Acknowledgement

DW was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre Programme. NK was supported by the University of Oxford Clarendon Fund.

Appendix: Supplementary material

Supplementary data to this article can be found online at [doi:10.1016/j.rbmo.2018.05.005](https://doi.org/10.1016/j.rbmo.2018.05.005).

ARTICLE INFO

Article history:

Received 11 October 2017

Received in revised form 30 April 2018

Accepted 2 May 2018

Declaration: The authors report no financial or commercial conflicts of interest.

Keywords:

Beta-thalassaemia

Next generation sequencing

Preimplantation genetic diagnosis

Sickle-cell anaemia

REFERENCES

Agasa, B., Bosunga, K., Opara, A., Tshilumba, K., Dupont, E., Vertongen, F., Cotton, F., Gulbis, B., 2010. Prevalence of sickle cell disease in a northeastern region of the Democratic Republic of Congo: what

- impact on transfusion policy? *Transfus. Med.* 20, 62–65. doi:10.1111/j.1365-3148.2009.00943.x.
- Angastiniotis, M., Modell, B., 1998. Global epidemiology of hemoglobin disorders. *Ann. N. Y. Acad. Sci.* 850, 251–269. doi:10.1111/j.1749-6632.1998.tb10482.x.
- Ben-Nagi, J., Wells, D., Doye, K., Loutradi, K., Exeter, H., Drew, E., Alfarawati, S., Naja, R., Serhal, P., 2017. Karyomapping: a single centre's experience from application of methodology to ongoing pregnancy and live-birth rates. *Reprod. Biomed. Online* 35, 264–271. doi:10.1016/j.rbmo.2017.06.004.
- Cao, A., Galanello, R., 2010. Beta-thalassaemia. *Genet. Med.* 12, 61–76. doi:10.1097/GIM.0b013e3181cd68ed.
- Chen, S.-C., Xu, X.-L., Zhang, J.-Y., Ding, G.-L., Jin, L., Liu, B., Sun, D.-M., Mei, C.-L., Yang, X.-N., Huang, H.-F., Xu, C.-M., 2016. Identification of PKD2 mutations in human preimplantation embryos in vitro using a combination of targeted next-generation sequencing and targeted haplotyping. *Sci. Rep.* 6, 25488. doi:10.1038/srep25488.
- Fiorentino, F., Biricik, A., Bono, S., Spizzichino, L., Cotroneo, E., Cottone, G., Kokocinski, F., Michel, C.E., 2014a. Development and validation of a next-generation sequencing-based protocol for 24-chromosome aneuploidy screening of embryos. *Fertil. Steril.* 101, 1375–1382, e2. doi:10.1016/j.fertnstert.2014.01.051.
- Fiorentino, F., Bono, S., Biricik, A., Nuccitelli, A., Cotroneo, E., Cottone, G., Kokocinski, F., Michel, C.-E., Minasi, M.G., Greco, E., 2014b. Application of next-generation sequencing technology for comprehensive aneuploidy screening of blastocysts in clinical preimplantation genetic screening cycles. *Hum. Reprod.* 29, 2802–2813. doi:10.1093/humrep/deu277.
- Giménez, C., Sarasa, J., Arjona, C., Vilamajó, E., Martínez-Pasarell, O., Wheeler, K., Valls, G., Garcia-Guixé, E., Wells, D., 2015. Karyomapping allows preimplantation genetic diagnosis of a de novo deletion undetectable using conventional PGD technology. *Reprod. Biomed. Online* 31, 770–775. doi:10.1016/j.rbmo.2015.08.017.
- Handyside, A.H., Kontogianni, E.H., Hardy, K., Winston, R.M., 1990. Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature* 344, 768–770. doi:10.1097/00006254-199107000-00024.
- Hou, Y., Wu, K., Shi, X., Li, F., Song, L., Wu, H., Dean, M., Li, G., Tsang, S., Jiang, R., Zhang, X., Li, B., Liu, G., Bedekar, N., Lu, N., Xie, G., Liang, H., Chang, L., Wang, T., Chen, J., Li, Y., Zhang, X., Yang, H., Xu, X., Wang, L., Wang, J., 2015. Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. *Gigascience* 4, 37. doi:10.1186/s13742-015-0068-3.
- Konstantinidis, M., Prates, R., Goodall, N.-N., Fischer, J., Tecson, V., Lemma, T., Chu, B., Jordan, A., Armenti, E., Wells, D., Munné, S., 2015. Live births following Karyomapping of human blastocysts: experience from clinical application of the method. *Reprod. Biomed. Online* 31, 394–403. doi:10.1016/j.rbmo.2015.05.018.
- Moutou, C., Goossens, V., Coonen, E., De Rycke, M., Kokkali, G., Renwick, P., Sengupta, S.B., Vesela, K., Traeger-Synodinos, J., 2014. ESHRE PGD Consortium data collection XII: Cycles from January to December 2009 with pregnancy follow-up to October 2010. *Hum. Reprod.* 29, 880–903. doi:10.1093/humrep/deu012.
- Natesan, S.A., Handside, A.H., Thornhill, A.R., Ottolini, C.S., Sage, K., Summers, M.C., Konstantinidis, M., Wells, D., Griffin, D.K., 2014. Live birth after PGD with confirmation by a comprehensive approach (karyomapping) for simultaneous detection of monogenic and chromosomal disorders. *Reprod. Biomed. Online* 29, 600–605. doi:10.1016/j.rbmo.2014.07.007.
- Ren, Y., Zhi, X., Zhu, X., Huang, J., Lian, Y., Li, R., Jin, H., Zhang, Y., Zhang, W., Nie, Y., Wei, Y., Liu, Z., Song, D., Liu, P., Qiao, J., Yan, L., 2016. Clinical applications of MARSALA for preimplantation genetic diagnosis of spinal muscular atrophy. *J. Genet. Genomics* 43, 541–547. doi:10.1016/j.jgg.2016.03.011.
- Thornhill, A.R., Handside, A.H., Ottolini, C., Natesan, S.A., Taylor, J., Sage, K., Harton, G., Cliffe, K., Affara, N., Konstantinidis, M., Wells,

- D., Griffin, D.K., 2015. Karyomapping—a comprehensive means of simultaneous monogenic and cytogenetic PGD: comparison with standard approaches in real time for Marfan syndrome. *J. Assist. Reprod. Genet.* 32, 347–356. doi:10.1007/s10815-014-0405-y.
- Treff, N.R., Fedick, A., Tao, X., Devkota, B., Taylor, D., Scott, R.T., 2013. Evaluation of targeted next-generation sequencing-based preimplantation genetic diagnosis of monogenic disease. *Fertil. Steril.* 99, 1377–1384, e6. doi:10.1016/j.fertnstert.2012.12.018.
- Tshilolo, L., Kafando, E., Sawadogo, M., Cotton, F., Vertongen, F., Ferster, A., Gulbis, B., 2008. Neonatal screening and clinical care programmes for sickle cell disorders in sub-Saharan Africa: Lessons from pilot studies. *Public Health* 122, 933–941. doi:10.1016/j.puhe.2007.12.005.
- Wells, D., Kaur, K., Grifo, J., Glassner, M., Taylor, J.C., Fragouli, E., Munne, S., 2014. Clinical utilisation of a rapid low-pass whole genome sequencing technique for the diagnosis of aneuploidy in human embryos prior to implantation. *J. Med. Genet.* 51, 553–562. doi:10.1136/jmedgenet-2014-102497.
- Yan, L., Huang, L., Xu, L., Huang, J., Ma, F., Zhu, X., Tang, Y., Liu, M., Lian, Y., Liu, P., Li, R., Lu, S., Tang, F., Qiao, J., Xie, X.S., 2015. Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. *Proc. Natl. Acad. Sci.* 112, 15964–15969. doi:10.1073/pnas.1523297113.
- Zheng, H., Jin, H., Liu, L., Liu, J., Wang, W.-H., 2015. Application of next-generation sequencing for 24-chromosome aneuploidy screening of human preimplantation embryos. *Mol. Cytogenet.* 8, 38. doi:10.1186/s13039-015-0143-6.