Human Reproduction, Vol.37, No.11, pp. 2700–2708, 2022
Advance Access Publication on September 23, 2022 https://doi.org/10.1093/humrep/deac208

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

Embryo tracking system for high-throughput sequencing-based preimplantation genetic testing

Wanwisa van Dijk[1,†], Kasper Derks[1,†], Marion Drüsedau[1,†], Jeroen Meekels[1], Rebekka Koeck[1,2], Rick Essers[1,2], Joseph Dreesen[1,2], Edith Coonen[1,3], Christine de Die-Smulders[1,2], Servi J.C. Stevens[1,2], Han G. Brunner[1,2,4], Arthur van den Wijngaard[1,2], Aimée D.C. Paulussen[1,2], and Masoud Zamani Esteki[1,2,*]

[1]Department of Clinical Genetics, Maastricht University Medical Center+ (MUMC+), Maastricht, The Netherlands [2]Department of Genetics and Cell Biology, GROW School for Oncology and Reproduction, Maastricht University, Maastricht, The Netherlands [3]Center for Reproductive Medicine, Maastricht University Medical Centre+, GROW School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands [4]Department of Human Genetics, Radboud University Medical Centre, Nijmegen, The Netherlands

*Correspondence address. Laboratory of Cellular Genomic Medicine, Department of Clinical Genetics, Maastricht UMC+, P. Debyelaan 25, 6229 HX Maastricht, The Netherlands. Tel: +31-43-38-75306; E-mail: masoud.zamaniesteki@mumc.nl https://orcid.org/0000-0003-3909-0050

Submitted on March 23, 2022; resubmitted on September 1, 2022; editorial decision on September 9, 2022

STUDY QUESTION: Can the embryo tracking system (ETS) increase safety, efficacy and scalability of massively parallel sequencing-based preimplantation genetic testing (PGT)?

SUMMARY ANSWER: Applying ETS-PGT, the chance of sample switching is decreased, while scalability and efficacy could easily be increased substantially.

WHAT IS KNOWN ALREADY: Although state-of-the-art sequencing-based PGT methods made a paradigm shift in PGT, they still require labor intensive library preparation steps that makes PGT cost prohibitive and poses risks of human errors. To increase the quality assurance, efficiency, robustness and throughput of the sequencing-based assays, barcoded DNA fragments have been used in several aspects of next-generation sequencing (NGS) approach.

STUDY DESIGN, SIZE, DURATION: We developed an ETS that substantially alleviates the complexity of the current sequencing-based PGT. With (n = 693) and without (n = 192) ETS, the downstream PGT procedure was performed on both bulk DNA samples (n = 563) and whole-genome amplified (WGAed) few-cell DNA samples (n = 322). Subsequently, we compared full genome haplotype landscapes of both WGAed and bulk DNA samples containing ETS or no ETS.

PARTICIPANTS/MATERIALS, SETTING, METHODS: We have devised an ETS to track embryos right after whole-genome amplification (WGA) to full genome haplotype profiles. In this study, we recruited 322 WGAed DNA samples derived from IVF embryos as well as 563 bulk DNA isolated from peripheral blood of prospective parents. To determine possible interference of the ETS in the NGS-based PGT workflow, barcoded DNA fragments were added to DNA samples prior to library preparation and compared to samples without ETS. Coverages and variants were determined.

MAIN RESULTS AND THE ROLE OF CHANCE: Current PGT protocols are quality sensitive and prone to sample switching. To avoid sample switching and increase throughput of PGT by sequencing-based haplotyping, six control steps should be carried out manually and checked by a second person in a clinical setting. Here, we developed an ETS approach in which one step only in the entire PGT procedure needs the four-eyes principal. We demonstrate that ETS not only precludes error-prone manual checks but also has no effect on the genomic landscape of preimplantation embryos. Importantly, our approach increases efficacy and throughput of the state-of-the-art PGT methods.

[†]These authors are joint first authors and contributed equally to this work.

LIMITATIONS, REASONS FOR CAUTION: Even though the ETS simplified sequencing-based PGT by avoiding potential errors in six steps in the protocol, if the initial assignment is not performed correctly, it could lead to cross-contamination. However, this can be detected *in silico* following downstream ETS analysis. Although we demonstrated an approach to evaluate purity of the ETS fragment, it is recommended to perform a pre-PGT quality control assay of the ETS amplicons with non-human DNA, such that the purity of each ETS molecule can be determined prior to ETS-PGT.

WIDER IMPLICATIONS OF THE FINDINGS: The ETS-PGT approach notably increases efficacy and scalability of PGT. ETS-PGT has broad applicative value, as it can be tailored to any single- and few-cell sequencing approach where the starting specimen is scarce, as opposed to other methods that require a large number of cells as the input. Moreover, ETS-PGT could easily be adapted to any sequencing-based diagnostic method, including PGT for structural rearrangements and aneuploidies by low-pass sequencing as well as non-invasive prenatal testing.

STUDY FUNDING/COMPETING INTEREST(S): M.Z.E. is supported by the EVA (Erfelijkheid Voortplanting & Aanleg) specialty program (grant no. KP111513) of Maastricht University Medical Centre (MUMC+), and the Horizon 2020 innovation (ERIN) (grant no. EU952516) of the European Commission.

TRIAL REGISTRATION NUMBER: N/A.

Key words: preimplantation genetic testing / IVF embryos / massively parallel sequencing / haplotyping / copy-number profiling / haplarithmisis / genotyping-by-sequencing / clinical sequencing

Introduction

Since the birth of the first *in vitro* fertilized (IVF) baby in 1978 (Steptoe and Edwards, 1978), more than 8 million individuals have been conceived via IVF. This continues to increase due to various demographic factors, including advanced parental age. Preimplantation genetic testing (PGT) has evolved from locus- and family-specific genetic testing, e.g. PCR- and FISH-based PGT methods, to more sophisticated generic approaches, e.g. genome-wide haplotyping methods (Handyside *et al.*, 1990, 2010; Zamani Esteki *et al.*, 2015; Backenroth *et al.*, 2019). Currently, PGT is performed for monogenic disorders (PGT-M), structural rearrangements (PGT-SR) and aneuploidies (PGT-A). Over the last few years, the demand for PGT has increased rapidly due to the continuous discovery of new disease genes and pathogenic mutations (Gilissen *et al.*, 2014), the development of massively parallel sequencing PGT (sequencing-based PGT) methods that have broadened the scope of PGT practice (Schobers *et al.*, 2021), increased public awareness of reproductive options, and the broader availability and accessibility of preconception carrier testing (Sallevelt *et al.*, 2021). For instance, in our center, the number of PGT-M requests increased from 58 in 2009 to 432 in 2019, signifying that generic and scalable PGT is indispensable.

Whole-genome amplification (WGA) methods in combination with high-throughput single nucleotide polymorphism (SNP) profiling platforms, including SNP-array and massively parallel sequencing, have enabled PGT at the single-cell resolution (Zamani Esteki *et al.*, 2015; Vermeesch *et al.*, 2016; Masset *et al.*, 2019; De Witte *et al.*, 2022; Masset *et al.*, 2022). Recently, we demonstrated sequencing-based haplarithmisis (Masset *et al.*, 2019), allowing simultaneous haplotyping and copy-number typing, such that all forms of PGT (PGT-M, PGT-A and PGT-SR) can be performed in a single assay. As a result, the required time for a PGT work-up was reduced drastically. However, these sequencing-based PGT methods are still laborious and prone to specimen provenance errors, i.e. sample switching, cross-contamination or product carryover. This is due to the increased number of wet-lab steps of these methods as compared to the traditional locus-specific PCR-based approaches, which adds to the possibility of human error and misdiagnosis (Wilton *et al.*, 2009). Although good laboratory practice in sample handling and laboratory automation is employed to minimize specimen provenance errors, multiple control steps are still essential. For instance, sample switching can affect 3% of samples in clinical laboratory testing (Pfeifer and Liu, 2013; Sehn *et al.*, 2015). Previously, several methods for sample tracking and the detection of cross-contamination and for single-cell DNA and RNA sequencing have been developed (Xu *et al.*, 2012; Quail *et al.*, 2014; Cusanovich *et al.*, 2015; Mulqueen *et al.*, 2021). However, none have been proven to be suitable for tracing rare cells in a clinical setting, e.g. single- or few-cell DNA samples derived from human preimplantation embryos.

To minimize specimen provenance errors and increase the scalability of our sequencing-based PGT procedure, we developed an embryo tracking system (ETS)-PGT (Fig. 1a). We adapted sample tracking, using spiked-in short DNA probes (Quail *et al.*, 2014), and developed an innovative, easy-to-use approach that makes sequencing-based PGT more robust with higher throughput. ETS-PGT is unique due to its incorporation of the ETS fragments with (i) an extra 20-nucleotide sequence that allows the restriction enzyme to bind, (ii) an adjacent restriction site that is specific for the sequencing-based PGT procedure, (iii) an extra primer binding site that makes sample tracking universal for any sequencing-based wet-lab protocol and (iv) a complementary, integrative computational pipeline that automatically traces the embryos. Here, we show that the ETS eliminates the necessity of the four-eyes principal for six crucial control steps in sequencing-based PGT, allowing not only higher quality assurance but also increasing the scalability of the process by enabling a fully robotized comprehensive PGT.

Materials and methods

Patients with informed consent and embryo biopsies

All couples were counseled by clinical geneticists at Maastricht University Medical Centre (MUMC+) and enrolled in the diagnostic PGT procedure (licensed by the Dutch Ministry of Health, Welfare and Sport CZ-TSZ-291208) after signing an informed consent form.

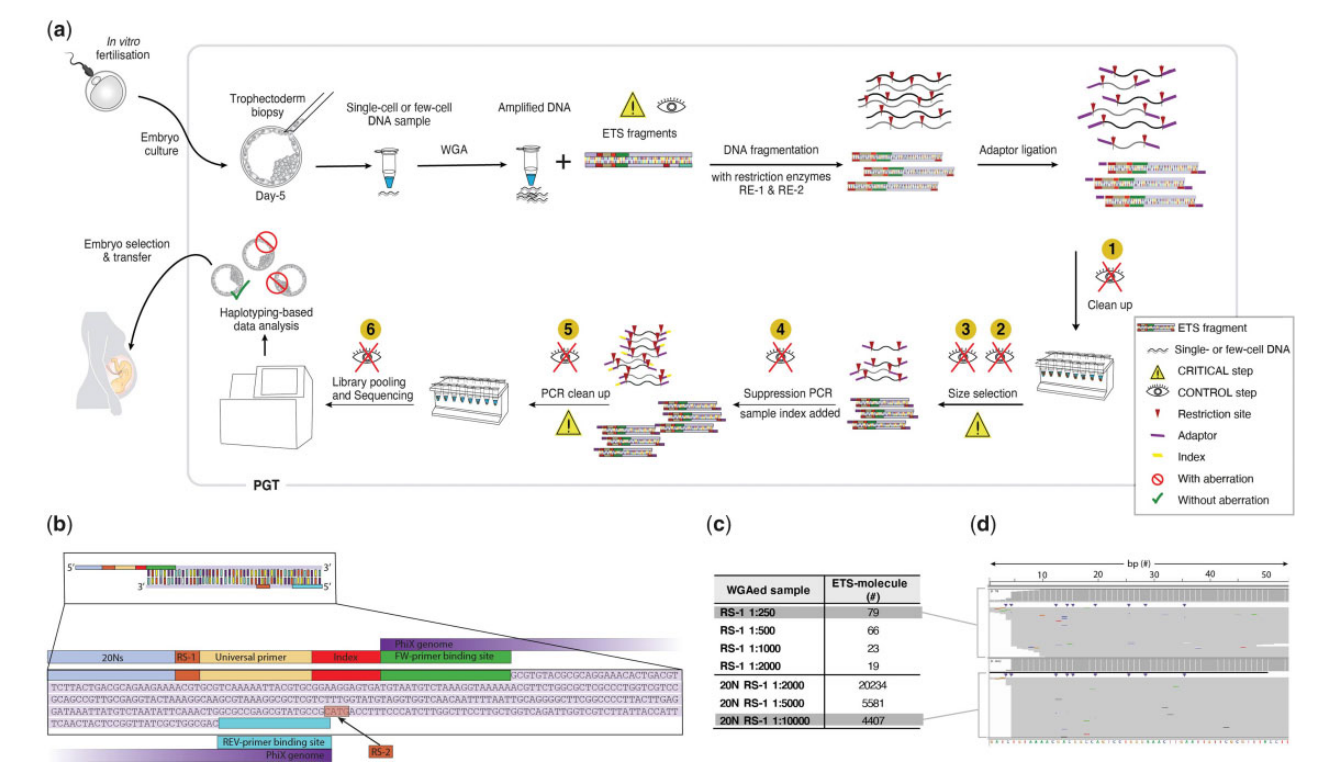

Figure 1. ETS-PGT design, mass production and implementation in PGT practice. (**a**) Implementation of the ETS in sequencing-based PGT procedure increases quality assurance and efficiency of sequencing-based PGT by avoiding controls in six crucial Control steps. (**b**) ETS fragments are based on the 429 nt amplicons of the PhiX 174 genome (purple) which contain the forward (FW, green) and reverse (REV, turquoise) primer binding site and the second restriction site (RS-2, brown). Added at the 3′ end is a unique 11 nt Illumina index (red), an 18 nt universal primer binding site (yellow), the first restriction site (RS-1, brown) and a 20 nt random nucleotide sequence (blue). (**c**) Dilution and number of retrieved ETS fragment with and without 20 random nucleotides. (**d**) Low-yield (upper panel) and high-yield (lower panel) ETS without and with 20-nt, respectively. ETS, embryo tracking system; PGT, preimplantation genetic testing; RE-1, restriction enzyme one; RE-2, restriction enzyme two.

Couples suitable for the sequencing-based PGT procedure were included from December 2019 to December 2020 (Table I). Oocytes were fertilized by means of ICSI and embryos that had developed to the blastocyst stage, showing a distinct inner cell mass and trophectoderm, underwent laser-assisted trophectoderm biopsy in G-MOPS PLUS (Vitrolife) at Day 5/6 post-fertilization (Masset *et al.*, 2019). Biopsy samples containing five to eight trophectoderm cells were further subjected to genetic analysis.

PGT procedure and library preparation

The library preparation method for massively parallel sequencing, with an adapted form of the haplarithmisis algorithm, called OnePGT solution, has been previously described (Masset *et al.*, 2019). Surplus whole embryos were collected in a total of 2 µl washing buffer (Ca2+ and Mg2+ free phosphate-buffered saline with 0.2% polyvinylpyrrolidone (Sigma-Aldrich Chemie BV)). Bulk DNA samples of parents and phasing references were isolated from peripheral blood. All WGA products and genomic DNA from parents and references were then processed using OnePGT solution (Agilent Technologies) according to the manufacturer's instructions. Briefly, 500 ng of (whole-genome

Table I Samples used for clinical validation of the ETS-PGT approach.

Inheritance mode[a]	Indication[b] (#)	PGT samples (#)			
		Without ETS		With ETS	
		Bulk	WGAed	Bulk	WGAed
AD	61	83	39	220	154
AR	47	47	1	149	94
XL	10	11	1	37	30
XL/AR	1	0	0	3	0
AD/AR	4	3	0	10	3
Total	123	144	41	419	281

ETS, embryo tracking system; PGT, preimplantation genetic testing; WGAed, whole-genome amplified few-cell DNA samples; Bulk, bulk DNA samples.
[a]AD, autosomal dominant disorders; AR, autosomal recessive disorders; XL, X-linked disorders; XL/AR, both autosomal recessive and X-Linked disorders; AD/AR, both autosomal recessive and dominant disorders.
[b]Genetic indications (genes) per inheritance mode.

amplified (WGAed) or bulk) DNA was fragmented through restriction enzyme digestion, adapter-ligated, size-selected with PippinHT (Sage Science, USA) and PCR-amplified to yield a reduced representation library per sample. Per sequencing run, libraries of 24 samples were pooled equimolarly and sequenced on an Illumina NextSeq 500 using the High output (2 × 150 bp) kit (Illumina). The ETS is scalable as it has the capacity to pool 96 samples at a time using the 96 different ETS devised fragments (see Supplementary Table SI). A total of 885 samples (322 WGAed DNA from embryo biopsies and 563 peripheral blood DNA samples) were included in this study. Of the embryo trophectoderm biopsies, 308 were derived from '4–8 cells', 13 samples were from '9–15 cells' and 2 samples were from '16–25 cells'.

ETS amplicons preparation

The unique index sequence of the ETS amplicons was generated as described previously (Quail *et al.*, 2014). ETS amplicons were prepared by PCR using PhiX174 RF II DNA (New England Biolabs) as template DNA. In brief, each amplification reaction consisted of 200 ng of PhiX174 DNA (New England Biolabs), 0.5 µM of ETS-indexes forward primer, 0.5 µM ETS universal reverse primer and Q5 hot start high-fidelity 2× master mix (New England Biolabs). PCRs were performed on a Labcycler thermal cycler (Sensoquest GmbH) using the following conditions: 98°C for 30 s, 35 cycles of 98°C for 10 s, 60°C for 20 s, 72°C for 30 s and a final elongation step at 72°C for 30 s. A total of three PCR reactions were performed for each ETS amplicon. PCR products of each ETS amplicon were pooled and purified using the QIAquick PCR Purification Kit (Qiagen) and eluted in 30 µl of Qiagen elution buffer. Fragment concentration was measured using the Qubit™ dsDNA HS Assay Kit (Invitrogen). Each ETS amplicon was adjusted to a final concentration of 3 ng/µl, aliquoted and stored at −20°C as stock ETS plates. ETS amplicons were diluted further to 0.03 ng/µl. ETS fragments (Fig. 1a) and the entire ETS design (Fig. 1, Supplementary Table SI) were optimized and validated for sequencing-based PGT.

Haplarithmisis-based PGT

Demultiplexed sequencing data of both ETS indexes and sequencing-based PGT were mapped to the human reference genome, GRCh37/hg19, complemented with the sequences of all ETS amplicons. Subsequently, the number and purity of the expected ETS fragments for each sample were computed. Purity (PUR) represents the percentage of an ETS fragment for a sample 's':

$$PUR_s = \frac{\sum_e eETS_{s,e}}{\sum_d dETS_{s,d}} \times 100$$

where eETS is the number of expected ETS fragments and dETS is the number of detected ETS for sample 's', i.e. dETS is total number of both expected and unexpected ETS fragments. The index sequence of ETS amplicons was extracted using samtools (version 1.2) (Li *et al.*, 2009). Data from sequencing-based PGT samples were analyzed using our analytical pipeline which includes a pre-PGT test and several quality control steps to ascertain genome-wide copy-number and haplotype profiles for each embryo. Briefly, we applied haplotypecaller from the GATK tool (McKenna *et al.*, 2010) to extract the genomic locations annotated in the dbSNP database (version 150). Using the R-function extract.gt (vcfR package bioconductor), the coverage of the genomic locations was calculated per sample. We then applied haplarithmisis as described previously (Zamani Esteki *et al.*, 2015).

In silico tracking of embryos

The index sequence of ETS amplicons were extracted from the alignment (bam) files using samtools (version 1.2) (Li *et al.*, 2009). Then, filtered for ETS amplicons with a minimal of 50 reads and the ETS amplicon with the highest number of reads was reported as a percentage of total ETS amplicons detected. The reported ETS amplicon is matched with the added ETS amplicon.

Other statistical analysis and visualization

The breadth and depth of coverage were compared per sample type (bulk or WGAed) using Welch's *t*-test which is robust to sample groups with unequal variances and sizes. Non-parametric allele drop out (ADO) and allele drop in (ADI) rates in WGAed samples were compared with ETS (n = 241) and without ETS (n = 51) using Wilcoxon signed-rank tests. Substandard samples (n = 5) were excluded based on QC-by-parents criteria (Zamani Esteki *et al.*, 2015) and the interquartile range above Q3 + 1.5 interquartile range (IQR). For genome haplarithm visualization, we applied adapted visualization modules of siCHILD.

Results

Embryo tracking system design

For the ETS fragments, we made use of the PhiX 174 DNA sequence as a template and amplified them with uniquely designed PCR primers that are specific to the PhiX genome, such that each restriction enzyme could cleave the DNA in only one location. PCR primer pairs with an optimal melting temperature and single restriction site, resulted in a PhiX amplicon length of 429 bp, including a forward primer at PhiX genome position 742 and a reverse primer at PhiX genome position 1138. The identification of a PhiX endogenous 3′ restriction site (RS-2) and the addition of a 5′ restriction site (RS-1) allowed us to make these fragments compatible with sequencing-based PGT. We further optimized the fragments by adding an 11 nt unique index (Kozarewa and Turner, 2011; Quail *et al.*, 2011; Bronner *et al.*, 2014) adjacent to the PhiX forward primer binding site as well as a universal primer (5′-GGCGTCCATCTCGAAG-3′) between RS-1 and the index. For optimal binding of the restriction enzyme, 20 random nucleotides were added at the 5′-end prior to RS-1 (Fig. 1b).

ETS fragment quantity and purity for next-generation sequencing-based PGT

The amount of ETS fragments added to the samples was optimized by making a dilution range in the WGAed DNA samples (Fig. 1c). The low yield of the ETS molecules indicated a suboptimal digestion when the molecule started directly with the restriction site RS-1 (Fig. 1d, upper panel). However, the incorporated 20-nucleotide fragment enabled optimal restriction enzyme binding (Fig. 1d, lower panel). Reducing to a 1:10 000 dilution resulted in median of 2475 (IQR = 1378–4120) ETS molecules detected after sequencing (Fig. 2a). In addition, we calculated the purity of the ETS fragments for

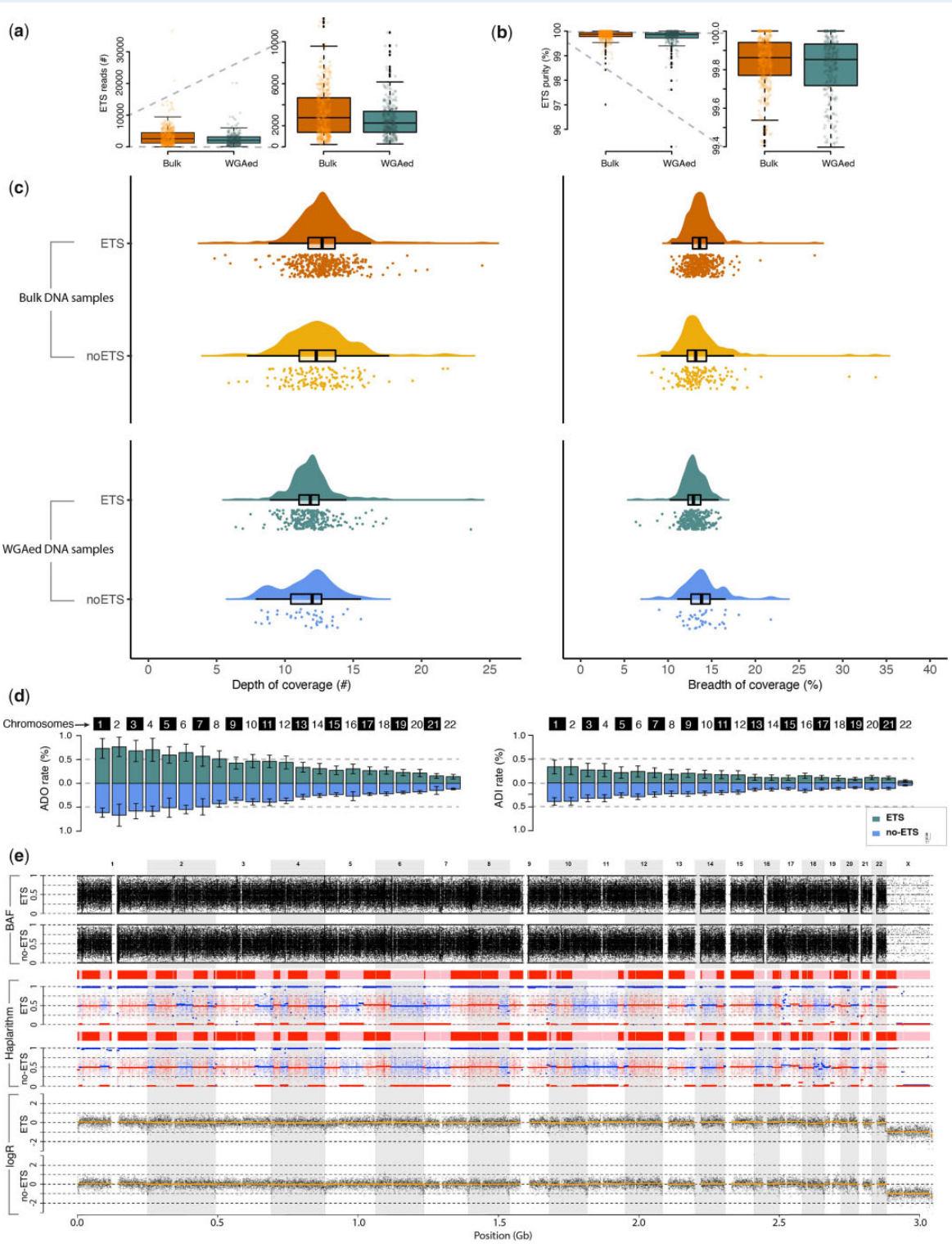

Figure 2. Application of the embryo tracking system (ETS) for preimplantation genetic testing (PGT) from sequenced reads to reconstructed haplotypes. (**a**) Number of ETS fragments and (**b**) their purity in all the samples (n = 885) that were processed by the ETS-PGT protocol. (**c**) Average depth and breadth of coverage with and without ETS for whole-genome amplified (WGAed) (n = 322) and bulk (n = 563) DNA samples. (**d**) Allelic dropout (ADO) and allelic drop in (ADI) rates following QC-by-parents analysis that computes mendelian inconsistencies across all WGAed samples (n = 241 with ETS and n = 51 without ETS). (**e**) Genome-wide profiles with and without ETS-PGT. From top to bottom, we show B allele frequency (BAF) profiles, maternal haplotypes and haplarithms and logR (relative copy number) values of embryo B2, each with and without ETS, respectively (see also Supplementary Fig. S1).

Table II Ruling out possible risks via ETS in a sequencing-based PGT haplotyping procedure.

NGS-based PGT step (see Fig. 1a)	Without ETS				With ETS			
	Sample swap risk		Sample cross-contamination		Sample swap risk		Sample cross-contamination	
	Presence	Detected	Presence	Detected	Presence	Detected	Presence	Detected
DNA fragmentation	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Adaptor ligation	No	N/A	No	N/A	No	N/A	No	N/A
Clean up #1	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Size selection	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Suppression PCR	Yes	No	Yes	No	Yes	Yes	Yes	Yes
PCR clean up #2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

During the NGS-based PGT, different types of risks can be introduced, but ETS can detect all those risks (see also Fig. 1a). Orange represents the possible risks or failure to detect those risks. Green represents no risk or ability to detect the risk. Grey represents steps with no risk introduced.
ETS, embryo tracking system; PGT, preimplantation genetic testing; NGS, next-generation sequencing.

each sample and 689 out of 693 samples (99.42%) had a purity >98% (Fig. 2b, Supplementary Fig. S1). The small deviation from 100% purity could be due to amplification and/or sequencing errors. Overall, the ETS fragment purity in all the 693 samples had a median of 99.86% (IQR = 99.75–99.94%).

ETS implementation and clinical validation for PGT

Adding ETS fragments to WGAed or bulk DNA samples before sequencing-based PGT eliminates the necessity of the four-eyes principle at six Control steps (Fig. 1a), thus facilitating accurate and scalable PGT. The Critical steps indicate quality sensitive steps. At these steps, the DNA concentration is measured to detect if it falls within the QC-range of each specific step. The Control steps are the ones that four-eyes principle should be applied to avoid sample swap. In the process, both the DNA samples and ETS fragments are registered *in silico*. During the cleanup with magnetic beads, the purified product is removed from the wells with beads. At Control steps 2 and 3, the products are size selected and transferred back to new wells and after dilution suppression PCR is performed and indexes are added (Control step 4). Subsequently, the product is cleaned (Control step 5), similar to Control step 2. At the last Control step 6, 24 samples are pooled and prepared for a sequencing run (see also Table II).

To clinically evaluate ETS-PGT, we analyzed WGAed DNA samples from IVF preimplantation embryos (n = 322) of couples (n = 162) who opted for PGT with 123 different genetic indications (Table I and Supplementary Table SII). The PGT procedure was performed on DNA samples with and without the ETS. By adding 0.06 ng of ETS fragments to 500 ng DNA samples, we observed comparable depth of coverage (bulk ETS: 12.76 ± 1.86 SD versus bulk without ETS: 12.47 ± 2.33 SD, $P = 0.176$ Welch's *t*-test, and WGAed ETS: 11.88 ± 1.63 SD versus WGAed without ETS: 11.55 ± 1.88 SD, $P = 0.283$ Welch's *t*-test) but slightly lower breadth of coverage in the WGAed ETS samples as compared to WGAed samples processed without ETS (bulk ETS: 13.70% ± 1.50% SD versus 13.65% ± 2.94% SD, $P = 0.846$ Welch's *t*-test, and WGAed ETS: 12.96 ± 1.24 SD versus WGAed without ETS: 13.96 ± 2.16 SD, $P = 0.005$ Welch's *t*-test) (Fig. 2c). The accuracy of the assayed SNP calls was measured by computation of WGA artifact using parental SNP calls to determine mendelian inconsistencies (Zamani Esteki *et al.*, 2015). We found comparable ADO and ADI WGA artifacts ($P = 0.153$, Wilcoxon signed-rank test). Genome-wide ADO rates with and without ETS were 9.48% (±2.53% SD) and 8.05% (±1.36% SD), respectively. Genome-wide ADI rates with and without ETS were 3.86% (±1.73% SD) and 4.45% (±0.86% SD), respectively (Fig. 2d). Furthermore, the resulting genome-wide haplotype calls were 99.04% (±0.12% SD) concordant in embryos (n = 3) of one family for which we performed PGT with and without ETS on the same WGAed DNA samples (Fig. 2e).

ETS rules out the chance of sample switching during PGT procedure

Without the four-eyes principle at the six error-prone Control steps, it is not possible to detect sample switching. To test our system, we intentionally mixed samples with no, wrong or mixed ETS fragments in both WGAed and bulk DNA samples. Our computational pipeline could easily trace switched and mixed samples (Supplementary Fig. S2) such that by adding ETS fragments to DNA samples prior to the sequencing-based PGT process, all the possible sample switching or cross-contamination events could be detected (Table II), thereby, eliminating the necessity of controls at six critical steps.

To evaluate sensitivity of detecting contamination of WGAed DNA samples with other DNA samples without ETS, we made different admixtures of two sibling embryos (Fig. 3): one normal diploid embryo (Embryo1) and the other with trisomy Chr 22 (Embryo2). Since our approach is genome wide, this allowed us to examine admixtures of both (i) normal diploid and trisomy (Fig. 3a), as well as (ii) normal disomic chromosomes of Embryo1 and Embryo2 (Fig. 3b). This experiment revealed that contaminations <10% cannot be detected without ETS and that only contaminations >30% can be detected in discrete haplotypes with false haplotype blocks and false positive crossover sites, albeit distortion of B-allele frequency values in haplarithms is indicative of contaminations. The 50:50% admixture showed a combination of both

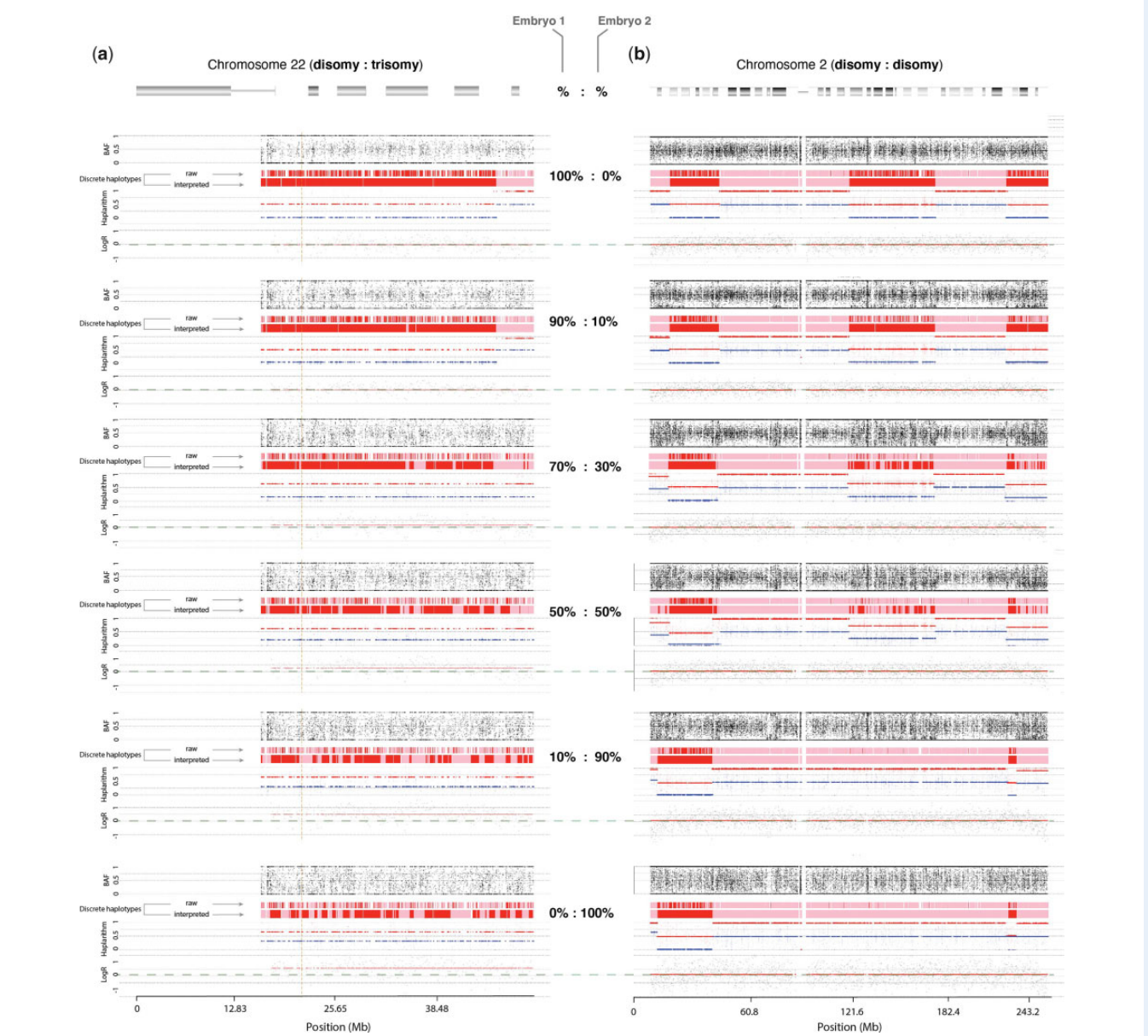

Figure 3. Detection of sample contamination through admixing whole-genome amplified (WGAed) samples of two sibling embryos in different proportions (diploid Embryo1 and Embryo2 with Chromosome (Chr) 22 trisomy) without the embryo tracking system (ETS). (**a**) Normal disomic Chr22 of Embryo1 and Chr22 trisomy of Embryo2. (**b**) Normal Chr2 disomy of Embryo1 and Embryo2. Per mixture experiment (from top to bottom), we show B allele frequency (BAF) profile, raw and interpreted maternal haplotypes, maternal haplarithm and relative copy-number profile. Green dashed lines depict the baseline copy number 2 (logR = 0, i.e. disomic chromosome).

embryos, indicating that PGT-M cannot be performed (Fig. 3b). Moreover, contaminations <10% with WGAed DNA sample of a trisomy cannot be accurately detected in the logR-values (Fig. 3a).

Discussion

In the presence of PGT guidelines that provide best laboratory and clinical practice for traditional PGT approaches (Thornhill *et al.*, 2005), misdiagnoses still occur with estimated rates of <1% and <5% for FISH- and PCR-based PGT, respectively (Wilton *et al.*, 2009). The rate of misdiagnosis and adverse outcomes of sequencing-based PGT have not yet been reported. However, a number of laboratory errors, such as tube switching, could occur leading to a PGT-misdiagnosis with potentially devastating effects, such as the transfer of an affected embryo (Wilton *et al.*, 2009). In the improved PGT international guidelines, to avoid misdiagnosis, it is currently recommended to have an extra observer during labeling and sample identification steps for

PGT quality control and assurance (ESHRE PGT Consortium Steering Committee *et al.*, 2020). Here, we demonstrate that ETS-PGT enables detection of very low contamination of DNA samples from two different embryos (<1%, Fig. 2b), even if the embryos are from the same couple (sibling embryos). Contaminations with unrelated (embryo) DNA samples can be detected with our rigorous QC-by-parents criteria (Zamani Esteki *et al.*, 2015), such that samples with >15% mendelian inconsistencies and <80% concordant SNP calls are labeled as substandard and diagnosis is not performed. When DNA sample mixing occurs, samples with more than 2% of unexpected ETS fragments will be excluded from the downstream PGT analysis. Without the ETS, however, <10% of contamination can go unnoticed when a euploid WGAed DNA sample is contaminated by an aneuploid one (Fig. 3a) and only >30% contamination can be detected with certainty when two euploid WGAed DNA samples are mixed (Fig. 3b). In the case of contamination during ETS index preparation, i.e. contamination of ETS molecules, the similar low purity percentage in all samples receiving the same contaminated ETS amplicons will be observed. While our purity determination approach would prevent misdiagnosis, it could lead to repetition of a PGT run, as library preparation and sequencing would have to be repeated with the new WGA aliquot of the same biopsy together with highly pure ETS fragments (>98%). Therefore, determining the purity of ETS amplicons prior to PGT implementation using a pre-PGT quality control assay with non-human DNA is recommended.

Clinical massively parallel sequencing increasingly leads to discovery of novel pathogenic variants (Gilissen *et al*, 2014) and therefore inherently increases PGT requests by couples with such mutations. PGT by sequencing-based haplotyping could alleviate this high demand, as genome-wide PGT methods are generic, i.e. they do not require family- and locus-specific designs. Although stringent laboratory procedures are effective to reduce the risk of sample switching and cross-contamination, they are cost-prohibitive and still subject to errors. Here, we developed and clinically implemented ETS-PGT that effectively increases sequencing-based PGT quality assurance and throughput. ETS can easily be utilized in any restriction enzyme-based protocol, including automated sequencing-based PGT to support increased PGT requests in the future with a shorter turnaround time. ETS is not only suitable for PGT procedures but could also be implemented in other next-generation sequencing-based parallel procedures, such as non-invasive prenatal testing, single-molecule molecular inversion probes and whole-genome sequencing. We envision that ETS-PGT will rapidly be adopted in IVF clinics and will be incorporated into PGT best practice guidelines.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

The data underlying in this article cannot be shared publicly in order to protect the privacy of the individuals who participated in the study. The anonymized data can be requested through the corresponding author.

Acknowledgements

We gratefully thank all the families who participated in this study. We thank Ping Cao and Darina Obukhova for critical reading of the manuscript.

Authors' roles

K.D., J.D., A.D.C.P., A.v.d.W. and M.Z.E. designed the study. W.v.D. and K.D. contributed to ETS fragment design. W.v.D., M.D. and J.M. performed the experiment. K.D., J.M., R.K., R.E., E.C. and M.Z.E. analyzed and interpreted the data. W.v.D. and M.Z.E. drafted the manuscript. W.v.D., M.D., C.d.D.-S., S.S., H.G.B., A.D.C.P. and M.Z.E. contributed to writing and critical revision of the manuscript.

Funding

M.Z.E. is supported by the EVA (Erfelijkheid Voortplanting & Aanleg) specialty program (grant no. KP111513) of Maastricht University Medical Centre (MUMC+) and the Horizon 2020 innovation (ERIN) (grant no. EU952516) of the European Commission.

Conflict of interest

M.Z.E is co-inventor on patent applications: ZL910050-PCT/EP2011/060211-WO/2011/157846 'Methods for haplotyping single cells' and ZL913096-PCT/EP2014/068315-WO/2015/028576 'Haplotyping and copy-number typing using polymorphic variant allelic frequencies'.

References

Backenroth D, Zahdeh F, Kling Y, Peretz A, Rosen T, Kort D, Zeligson S, Dror T, Kirshberg S, Burak E *et al.* Haploseek: a 24-hour all-in-one method for preimplantation genetic diagnosis (PGD) of monogenic disease and aneuploidy. *Genet Med* 2019;**21**:1390–1399.

Bronner IF, Quail MA, Turner DJ, Swerdlow H. Improved protocols for illumina sequencing. *Curr Protoc Hum Genet* 2014;**80**:18.2.1-42.

Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 2015;**348**:910–914.

De Witte L, Raman L, Baetens M, De Koker A, Callewaert N, Symoens S, Tilleman K, Vanden Meerschaut F, Dheedene A, Menten B. GENType: all-in-one preimplantation genetic testing by pedigree haplotyping and copy number profiling suitable for third-party reproduction. *Hum Reprod* 2022;**37**:1678–1691.

ESHRE PGT Consortium Steering Committee, F Carvalho, E Coonen, V Goossens, G Kokkali, C Rubio, M Meijer-Hoogeveen, C Moutou, N Vermeulen, M. De Rycke ESHRE PGT Consortium good practice recommendations for the organisation of PGT. *Hum Reprod Open* 2020;**2020**:hoaa021.

Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A *et al.* Genome sequencing identifies major causes of severe intellectual disability. *Nature* 2014;**511**:344–347.

Handyside AH, Harton GL, Mariani B, Thornhill AR, Affara N, Shaw MA, Griffin DK. Karyomapping: a universal method for genome wide analysis of genetic disease based on mapping crossovers between parental haplotypes. *J Med Genet* 2010;**47**:651–658.

Handyside AH, Kontogianni EH, Hardy K, Winston RM. Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature* 1990;**344**:768–770.

Kozarewa I, Turner DJ. 96-plex molecular barcoding for the Illumina Genome Analyzer. *Methods Mol Biol* 2011;**733**:279–298.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;**25**:2078–2079.

Masset H, Ding J, Dimitriadou E, Debrock S, Tsuiko O, Smits K, Peeraer K, Voet T, Zamani Esteki M, Vermeesch JR. Single-cell genome-wide concurrent haplotyping and copy-number profiling through genotyping-by-sequencing. *Nucleic Acids Res* 2022;**50**:e63.

Masset H, Zamani Esteki M, Dimitriadou E, Dreesen J, Debrock S, Derhaag J, Derks K, Destouni A, Drusedau M, Meekels J *et al.* Multi-centre evaluation of a comprehensive preimplantation genetic test through haplotyping-by-sequencing. *Hum Reprod* 2019;**34**:1608–1619.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M *et al.* The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;**20**:1297–1303.

Mulqueen RM, Pokholok D, O'Connell BL, Thornton CA, Zhang F, O'Roak BJ, Link J, Yardımcı GG, Sears RC, Steemers FJ *et al.* High-content single-cell combinatorial indexing. *Nat Biotechnol* 2021;**39**:1574–1580.

Pfeifer JD, Liu J. Rate of occult specimen provenance complications in routine clinical practice. *Am J Clin Pathol* 2013;**139**:93–100.

Quail MA, Otto TD, Gu Y, Harris SR, Skelly TF, McQuillan JA, Swerdlow HP, Oyola SO. Optimal enzymes for amplifying sequencing libraries. *Nat Methods* 2011;**9**:10–11.

Quail MA, Smith M, Jackson D, Leonard S, Skelly T, Swerdlow HP, Gu Y, Ellis P. SASI-Seq: sample assurance Spike-Ins, and highly differentiating 384 barcoding for Illumina sequencing. *BMC Genomics* 2014;**15**:110.

Sallevelt S, Stegmann APA, de Koning B, Velter C, Steyls A, van Esch M, Lakeman P, Yntema H, Esteki MZ, de Die-Smulders CEM *et al.* Diagnostic exome-based preconception carrier testing in consanguineous couples: results from the first 100 couples in clinical practice. *Genet Med* 2021;**23**:1125–1136.

Schobers G, Koeck R, Pellaers D, Stevens SJC, Macville MVE, Paulussen ADC, Coonen E, van den Wijngaard A, de Die-Smulders C, de Wert G *et al.* Liquid biopsy: state of reproductive medicine and beyond. *Hum Reprod* 2021;**36**:2824–2839.

Sehn JK, Spencer DH, Pfeifer JD, Bredemeyer AJ, Cottrell CE, Abel HJ, Duncavage EJ. Occult specimen contamination in routine clinical next-generation sequencing testing. *Am J Clin Pathol* 2015;**144**:667–674.

Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. *Lancet* 1978;**2**:366.

Thornhill AR, deDie-Smulders CE, Geraedts JP, Harper JC, Harton GL, Lavery SA, Moutou C, Robinson MD, Schmutzler AG, Scriven PN *et al.*; ESHRE PGD Consortium. ESHRE PGD Consortium 'Best practice guidelines for clinical preimplantation genetic diagnosis (PGD) and preimplantation genetic screening (PGS)'. *Hum Reprod* 2005;**20**:35–48.

Vermeesch JR, Voet T, Devriendt K. Prenatal and pre-implantation genetic diagnosis. *Nat Rev Genet* 2016;**17**:643–656.

Wilton L, Thornhill A, Traeger-Synodinos J, Sermon KD, Harper JC. The causes of misdiagnosis and adverse outcomes in PGD. *Hum Reprod* 2009;**24**:1221–1228.

Xu W, Gao H, Seok J, Wilhelmy J, Mindrinos MN, Davis RW, Xiao W. Coding SNPs as intrinsic markers for sample tracking in large-scale transcriptome studies. *Biotechniques* 2012;**52**:386–388.

Zamani Esteki M, Dimitriadou E, Mateiu L, Melotte C, Van der Aa N, Kumar P, Das R, Theunis K, Cheng J, Legius E *et al.* Concurrent whole-genome haplotyping and copy-number profiling of single cells. *Am J Hum Genet* 2015;**96**:894–912.