**Human Reproduction, Vol.38, No.3, pp. 511–519, 2023**
Advance Access Publication on January 10, 2023 https://doi.org/10.1093/humrep/deac273

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

# Preclinical workup using long-read amplicon sequencing provides families with *de novo* pathogenic variants access to universal preimplantation genetic testing

**Olga Tsuiko [1,2,†], Yasmine El Ayeb[1,†], Tatjana Jatsenko[1,†], Joke Allemeersch[2], Cindy Melotte[2], Jia Ding[2], Sophie Debrock[3], Karen Peeraer[3], Arne Vanhie[3], Anne De Leener[4], Céline Pirard[5], Candice Kluyskens[5], Ellen Denayer[2], Eric Legius[2], Joris Robert Vermeesch [1,2,*,‡], Hilde Brems [2,‡,*], and Eftychia Dimitriadou [2,‡,*]**

[1]Laboratory for Cytogenetics and Genome Research, Department of Human Genetics, KU Leuven, Leuven, Belgium [2]Centre for Human Genetics, University Hospitals Leuven, Leuven, Belgium [3]Leuven University Fertility Center, University Hospitals Leuven, Leuven, Belgium [4]Centre for Human Genetics, Cliniques Universitaires Saint Luc, UCLouvain, Brussels, Belgium [5]Department of Gynaecology, Cliniques Universitaires Saint Luc, UCLouvain, Brussels, Belgium

*Correspondence address. E-mail: eftychia.dimitriadou@uzleuven.be (E.D.) https://orcid.org/0000-0003-2845-0057; hilde.brems@kuleuven.be (H.B.) https://orcid.org/0000-0002-0325-4060; joris.versmeesch@uzleuven.be (J.R.V.) https://orcid.org/0000-0002-3071-1191

*Submitted on July 11, 2022; resubmitted on November 16, 2022; editorial decision on December 15, 2022*

**STUDY QUESTION:** Can long-read amplicon sequencing be beneficial for preclinical preimplantation genetic testing (PGT) workup in couples with a *de novo* pathogenic variant in one of the prospective parents?

**SUMMARY ANSWER:** Long-read amplicon sequencing represents a simple, rapid and cost-effective preclinical PGT workup strategy that provides couples with *de novo* pathogenic variants access to universal genome-wide haplotyping-based PGT programs.

**WHAT IS KNOWN ALREADY:** Universal PGT combines genome-wide haplotyping and copy number profiling to select embryos devoid of both familial pathogenic variants and aneuploidies. However, it cannot be directly applied in couples with a *de novo* pathogenic variant in one of the partners due to the absence of affected family members required for phasing the disease-associated haplotype.

**STUDY DESIGN, SIZE, DURATION:** This is a prospective study, which includes 32 families that were enrolled in the universal PGT program at the University Hospital of Leuven between 2018 and 2022. We implemented long-read amplicon sequencing during the preclinical PGT workup to deduce the parental origin of the disease-associated allele in the affected partner, which can then be traced in embryos during clinical universal PGT cycles.

**PARTICIPANTS/MATERIALS, SETTING, METHODS:** To identify the parental origin of the disease-associated allele, genomic DNA from the carrier of the *de novo* pathogenic variant and his/her parent(s) was used for preclinical PGT workup. Primers flanking the *de novo* variant upstream and downstream were designed for each family. Following long-range PCR, amplicons that ranged 5–10 kb in size, were sequenced using Pacific Bioscience and/or Oxford Nanopore platforms. Next, targeted variant calling and haplotyping were performed to identify parental informative single-nucleotide variants (iSNVs) linked to the *de novo* mutation. Following the preclinical PGT workup, universal PGT via genome-wide haplotyping was performed for couples who proceeded with clinical PGT cycle. In parallel, 13 trophectoderm (TE) biopsies from three families that were analyzed by universal PGT, were also used for long-read amplicon sequencing to explore this approach for embryo direct mutation detection coupled with targeted long-read haplotyping.

**MAIN RESULTS AND THE ROLE OF CHANCE:** The parental origin of the mutant allele was identified in 24/32 affected individuals during the preclinical PGT workup stage, resulting in a 75% success rate. On average, 5.95 iSNVs (SD = 4.5) were detected per locus of

[†]These authors are co-first authorship.
[‡]These authors are co-senior authorship.

interest, and the average distance of closest iSNV to the *de novo* variant was ~1750 bp. In 75% of those cases (18/24), the *de novo* mutation occurred on the paternal allele. In the remaining eight families, the risk haplotype could not be established due to the absence of iSNVs linked to the mutation or inability to successfully target the region of interest. During the time of the study, 12/24 successfully analyzed couples entered the universal PGT program, and three disease-free children have been born. In parallel to universal PGT analysis, long-read amplicon sequencing of 13 TE biopsies was also performed, confirming the segregation of parental alleles in the embryo and the results of the universal PGT.

**LIMITATIONS, REASONS FOR CAUTION:** The main limitation of this approach is that it remains targeted with the need to design locus-specific primers. Because of the restricted size of target amplicons, the region of interest may also remain non-informative in the absence of iSNVs.

**WIDER IMPLICATIONS OF THE FINDINGS:** Targeted haplotyping via long-read amplicon sequencing, particularly using Oxford Nanopore Technologies, provides a valuable alternative for couples with *de novo* pathogenic variants that allows access to universal PGT. Moreover, the same approach can be used for direct mutation analysis in embryos, as a second line confirmation of the preclinical PGT result or as a potential alternative PGT procedure in couples, where additional family members are not available.

**STUDY FUNDING/COMPETING INTEREST(S):** This work was supported by KU Leuven funding (no. C1/018 to J.R.V.) and Fonds Wetenschappelijk Onderzoek (1241121N to O.T.). J.R.V. is co-inventor of a patent 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' licensed to Agilent Technologies. All other authors have no conflict of interest to declare.

**TRIAL REGISTRATION NUMBER:** N/A.

**Key words:** long-read sequencing / preclinical PGT workup / universal PGT / ONT / oxford nanopore / *de novo* PGT

# Introduction

Preimplantation genetic testing for monogenic disorders (PGT-M) enables embryo selection to avoid the transmission of a heritable genetic condition from parents to offspring. The fundamental pillar of PGT-M is to differentiate the disease-associated allele from the normal allele in the embryo using discriminating genetic markers. Currently, targeted sequencing of the mutated locus, combined with the analysis of linked polymorphic single tandem repeats (STRs) or single-nucleotide polymorphisms (SNPs) remains the gold standard for PGT-M (De Rycke *et al.*, 2017; Coonen *et al.*, 2020). In addition, various generic genome-wide haplotyping-based PGT methods have been developed and implemented into routine clinical practice (Handyside *et al.*, 2010; Zamani Esteki *et al.*, 2015; Backenroth *et al.*, 2019; Masset *et al.*, 2019, 2022; De Witte *et al.*, 2022). By using genotype information of the couple (or only a single prospective parent (De Witte *et al.*, 2022)) and additional family members (e.g. offspring or parent(s) of the affected prospective parent, or parental siblings (Ding *et al.*, 2020)), informative loci for each of the parental haplotypes can be identified across all chromosomes, and haplotypes present in each embryo biopsy can be further reconstructed. Consequently, the inheritance of any Mendelian disease variants in the embryo can be inferred across the whole genome. Since genome-wide analyses also allow for the detection of whole chromosome aberrations and sub-chromosomal copy number variations, additional screening of the same single biopsy for aneuploidy (PGT-A) and structural rearrangements (PGT-SR) became feasible. This concurrent PGT-M, PGT-SR and PGT-A using the same biopsy and the same protocol is termed as universal PGT.

Currently, universal PGT methods require affected family members to trace the transmission of the disease-associated haplotype in the family. Hence, they cannot be offered directly to families, in which one of the partners or an offspring carries a *de novo* pathogenic variant. For such cases, different strategies have been implemented clinically, which include direct mutation detection and polymorphic marker evaluation (Rechitsky *et al.*, 2015; ESHRE PGT-M Working Group *et al.*, 2020). Depending on whether the male or female partner is a carrier of the *de novo* variant, single sperm cell or polar body (PB) analysis can be performed to determine healthy and mutant haplotypes, thus establish a phase. If the origin of a *de novo* variant is unknown (i.e. in the offspring), STR analysis coupled with direct mutation detection in embryos can be performed during the PGT cycle. For *de novo* cases, it is advisable to analyze multiple single sperm cells or oocytes/PBs to exclude germline mosaicism and to determine the risk-associated haplotype (Rechitsky *et al.*, 2011). Similarly, for direct embryo analysis, at least one affected and one unaffected embryo is needed to establish the correct phase and/or detect recombination events. When too few embryos are available, it is recommended to analyze unfertilized oocytes (in case of maternal origin of the *de novo* variant) or embryos that are not suitable for biopsy; otherwise analysis of embryos from the next cycle might be warranted (ESHRE PGT-M Working Group *et al.*, 2020). Thus, while these strategies are highly accurate, high- and low-risk haplotypes may be determined only during PGT cycle(s) and may be associated with certain complexities.

In this study, we propose an alternative strategy that utilizes third-generation long-read sequencing (LRS) for preclinical PGT workup of families with *de novo* variants. The advantage of LRS technologies, such as Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio), is that they can generate long (>10 kb) and even ultra-long (>100 kb) reads, which in turn allow detailed characterization of the genome and direct phasing of variants multiple kilobases apart (Mantere *et al.*, 2019). Here, we have performed targeted long-read amplicon sequencing in the affected partner and his/her parents as part of preclinical PGT workup to determine the parental origin of the disease-associated allele. During this analysis, the *de novo* variant in the affected partner was linked to the corresponding parental allele by tracking the inheritance of informative single-nucleotide variants (iSNVs). Consequently, by using LRS it was possible to deduce the

haplotype on which the *de novo* pathogenic variant arose in the affected individual, allowing the couple to proceed with universal PGT cycle using genome-wide haplotyping. Thus, we demonstrate that the application of long-read amplicon sequencing enables a rapid and inexpensive preclinical PGT workup for haplotype imputation in couples with *de novo* mutations.

# Materials and methods

## Ethical approval and study design

This study was approved by the Ethical Committee of UZ/KU Leuven (S59364). All data used in this study were generated in the framework of PGT-M program at the Center of Human Genetics, UZ Leuven. The data were coded, and the study was conducted in compliance with the principles of Declaration of Helsinki, good clinical practice and all applicable regulatory requirements.

There were 32 families, where one of the partners carries a *de novo* pathogenic variant causing a Mendelian disease, enrolled in the study (Supplementary Table SI). Primers flanking the *de novo* variant upstream and downstream were designed, and long-range PCR was performed for the affected partner and his/her parents using DNA extracted from whole blood. The exception was family PGD743, where two affected children and their parents were used for quartet analysis to determine the parental origin of mutated alleles in both children. Following long-read amplicon sequencing, the sequences were screened for iSNVs linked to the mutation. For families that proceeded with clinical PGT cycles, universal PGT was performed using siCHILD/haplarithmisis (Zamani Esteki *et al.*, 2015) and embryo classifications were performed according to previously published guidelines (Dimitriadou *et al.*, 2017). In three families that underwent universal PGT, the same long-read amplicon sequencing workflow was also used on trophectoderm (TE) biopsies for direct mutation analysis coupled with targeted haplotyping.

## Preclinical PGT workup

### *Primer design and long-range PCR amplification*

Primer pairs, flanking the variant of interest, and ranging from approximately 5–10 kb, were designed for each family using Primer3 software (Supplementary Table SII). For some families, partially overlapping amplicons, flanking the *de novo* variant, were generated to (i) increase the chance of iSNV detection and (ii) confirm the segregation pattern of detected iSNVs in the overlapping regions. Long-range PCR was performed using Takara LA Taq DNA polymerase (Takara Bio Inc., Japan) with the following conditions: 98°C for 20 s, followed by 30 cycles of denaturation at 98°C for 20 s, and annealing and extension at 68°C for 10 min. The presence of PCR products was evaluated by agarose gel electrophoresis. Successfully amplified long-range amplicons were purified using the Agencourt AMPure XP PCR purification system (Beckman Coulter Inc., USA) and quantified using Qubit™ dsDNA BR Assay Kit (Thermo Fisher Inc., USA).

### *Library preparation and long-read sequencing*

For PacBio sequencing, 100 ng from each amplicon were used for library preparation, according to publicly available SMRTbell™ library preparation protocol using PacBio Barcoded Adapters for Multiplex SMRT® Sequencing (PN 100-538-700-02). Briefly, one-step end-repair and ligation reaction was performed for each sample using NEBNext® FFPE DNA Repair Mix (New England Biolabs, NEB, USA), blunt end adaptors provided by PacBio and NEB T4 DNA ligase. Samples were then pooled together, and subsequent size-selection and purification were performed using PacBio's AMPure® PB beads. Next, DNA damage repair was performed for each library pool (with a total input per pool of 2 µg), followed by exonuclease treatment using ExoIII and ExoVII to remove failed ligation products. Finally, a second round of purification by AMPure® PB beads was performed. The quality and yield of the prepared library was checked by Fragment Analyzer™ (Agilent Technologies Inc., USA) and Qubit™ dsDNA HS Assay Kit. The generated libraries were sequenced on PacBio Sequel I platform using 1 M SMRTCells with Chemistry V3.0 with running time up to 20 h.

For Oxford Nanopore Technologies (ONT) sequencing, 150 ng of purified PCR product was used for library preparation following the manufacturer's SQK-LSK109 protocol. Each sample was first end-repaired and dA-tailed using NEBNext® FFPE DNA Repair Mix and NEBNext® Ultra™ II End Repair/dA-Tailing Module. Next, samples were purified using Agencourt AMPure XP beads, and barcoding was performed using ONT native barcoding kit EXP-NBD196. Barcoded samples were purified and pooled equimolarly to 20–30 ng. Next, sequencing adapter ligation was performed using the provided ligation sequencing kit 1D (SQK-LSK109) or NEBNext Quick Ligation Module, which includes T4 ligase. The final sequencing library was purified and quantified using Qubit™ dsDNA HS Assay Kit. Sequencing was performed on ONT MinION instrument according to Flongle™ flow cells loading instructions with the running time up to 20 h.

### *Variant calling and targeted haplotype phasing*

For PacBio data, the consensus reads were aligned to hg38 with Minimap2 (v2.17). The BAM files were merged per sample, sorted and indexed with samtools (v1.9). Variant calling was done with GATK HaplotypeCaller (v4.1.9.0), and the gvcf files were combined per trio with GATK CombineGVCFs and GenotypeGVCFs. The trio vcf was filtered for QualByDepth (QD $\geq$ 8 for indels and QC $\geq$ 2 for SNVs). With whatshap (v0.18), the vcf file is phased (considering the pedigree information) and a tag is added to the reads in the index BAM file to indicate the haplotype block for each read.

For ONT, MinKNOW Guppy (v4.2.2 or v5.0.11) was used for the basecalling and demultiplexing. Sequencing data quality was assessed using Nanoplot (v1.32.1). Trimming, quality and read length filtering were performed using Nanofilt (v2.7.1). Filtered fastq files were aligned to hg38 with Minimap2 (v2.17). Samtools (v1.11) was used to sort the BAM files and to remove unmapped reads and supplementary alignments. Candidate variant and haplotype consensus calling were performed using longshot (Edge and Bansal, 2019) (v0.4.1) and medaka (v1.2.1), released by ONT.

Haplotype-separated BAM files for both PacBio and ONT data were visualized using Integrative Genomics Viewer (IGV). For each family, the detected iSNVs were selected for validation by Sanger sequencing, using standard protocols (Crossley *et al.*, 2020).

## Clinical PGT cycles

### *Embryo culture and biopsy*

For families that proceeded with clinical PGT cycle, embryo culture and biopsy procedure was performed by standard clinical workflow of UZ Leuven/UCLouvain, described in detail elsewhere (Debrock *et al.*, 2010; Tšuiko *et al.*, 2021). Briefly, oocytes were fertilized by ICSI and were cultured at 37°C in 5–6% $CO_2$ and 5% $O_2$ in single Global Total® or Global Total LP® medium (CooperSurgical, USA) under mineral oil. Cleavage-stage biopsy was performed on Day 3 embryos with ≥6 blastomeres using Saturn 5TM Laser system (CooperSurgical, USA). For TE biopsies, laser-assisted zona opening was first performed on Day 3. Depending on blastocyst development, the biopsy was done on Days 5–7 post-insemination, aspirating approximately 5–10 cells with laser pulses (Saturn 5™ Laser system) in combination with mechanical detachment (flicking). All embryo biopsies were washed in 1% PVP/PBS droplets before transferring them into PCR tubes filled with 2 µl PBS. All embryo biopsies were stored at −20°C until further processing.

### *Embryo biopsy processing*

All biopsies were whole-genome amplified using multiple displacement amplification (REPLI-g Single Cell kit, Qiagen, Germany). All steps were done according to the manufacturer's protocol, except with a reduced incubation time of 2 h and inactivation of the enzyme at 65°C for 3 min. Amplified biopsies were quantified using Qubit™ dsDNA BR Assay Kit, and further processed for genotyping by Illumina HumanCytoSNP-12 BeadChip (Illumina Inc, USA). Parental and first-degree relatives (i.e. parents of prospective parents or couple's offspring) bulk DNA, extracted from blood, was also genotyped for subsequent universal PGT analysis.

### *siCHILD/haplarithmisis-based universal PGT*

Discrete genotypes, logR and B-allele frequency values, derived from SNP array data were used for genome-wide haplotyping-based PGT by siCHILD/haplarithmisis, the detailed concept of which has been described elsewhere (Zamani Esteki *et al.*, 2015). The output of universal PGT analysis includes (i) haplotyping information, defining the presence or absence of the disease-associated allele in the embryo, and (ii) genome-wide aneuploidy information, including parental and mechanistic origin (i.e. meiotic/mitotic origin of aneuploidy). Embryo classification was done according to internal guidelines (Dimitriadou *et al.*, 2017), and included five diagnostic categories: (i) 'unaffected', if the embryo is not a carrier of a disease-linked haplotype; (ii) 'affected', or carrier of the disease-linked haplotype; (iii) 'inconclusive', if the locus of interest lies <150 consecutive SNPs from a homologous recombination site; (iv) 'abnormal', if embryo has a meiotic aneuploidy or >3 chromosomes are affected; or (v) 'no result', if the sample failed quality control metrics (e.g. high cumulative standard deviation of SNP intensities or low SNP call rate).

### *TE biopsy direct mutation analysis using long-read amplicon sequencing*

For a subset of TE biopsies that were processed using universal PGT, direct mutation analysis coupled with targeted LRS haplotyping was also performed. Whole-genome amplified embryo biopsies, as well as unaffected partner, were subjected to long-range PCR, ONT long-read amplicon sequencing and data analysis using the preclinical PGT workflow, described above.

# Results

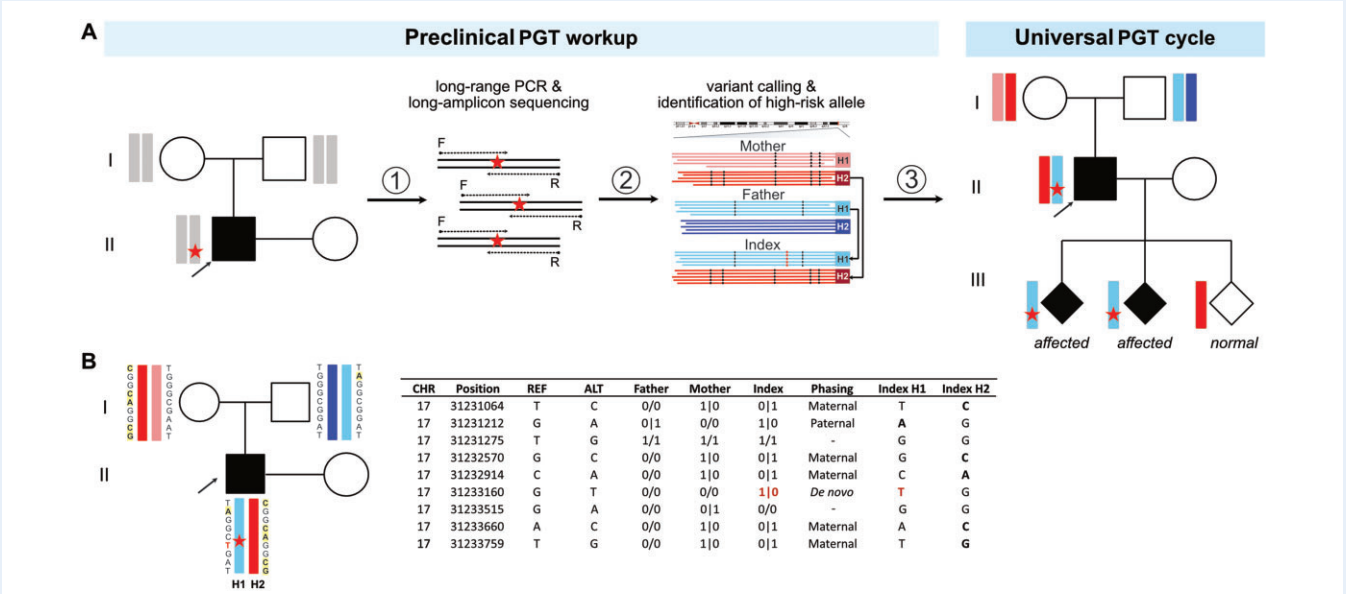
## LRS-based preclinical PGT workup identifies parental origin of the mutated allele

Targeted 5–10 kb amplicons flanking the *de novo* pathogenic variant were sequenced for 32 families with various PGT indications (Supplementary Table SI). LRS-based analysis allows the detection of iSNVs linked to the mutation in proximity, to deduce the parental origin of the disease-associated allele (Fig. 1). First 17 families with *de novo* pathogenic variants were sequenced using the PacBio RSI platform. From those 17 families, five were also selected as a validation set for ONT sequencing using MinION Flongle™ flow cells, which yielded a mean on target mapping rate of 99.8% (Supplementary Table SIII). While indels can also represent an informative source of genetic variations, we mainly focused on SNV detection, as LRS is known to have a higher rate of false positive indels, particularly when using ONT platform (Ahsan *et al.*, 2021). We then compared the performance of ONT variant calling in these five families to heterozygous and homozygous SNVs (n = 73), that were previously derived from PacBio data. This comparison demonstrated an overlap of 95.8% (70/73) and 97.2% (71/73) using longshot and medaka, respectively. When assigning the parental origin of the mutant allele, a 100% concordance rate was achieved between ONT and PacBio data. Consequently, because ONT provides more flexibility and is more cost-effective, compared to PacBio platform, the following 15 families were further processed by ONT sequencing only.

From all analyzed families, *de novo* pathogenic variants could be linked to iSNVs present on the same parental allele in 75% of cases (n = 24/32). On average, 5.95 iSNVs (SD = 4.5) were detected per analyzed family, and the average distance from the *de novo* variant to the closest flanking iSNV was 1754 bp (min = 126 bp, max = 6047 bp). All iSNVs were independently confirmed by Sanger sequencing. Apart from *de novo* point mutations, this approach has also been applied in a patient with ~2.4 kb *de novo* microduplication in NF1, spanning from exon 24 to exon 27. From successfully analyzed families, the pathogenic variant occurred on paternal and maternal alleles in 75% (n = 18/24) and 25% (n = 6/24) of cases, respectively. For one case, maternal germline mosaicism was confirmed, following the consecutive birth of two children with the same seemingly *de novo* pathogenic variant, and negative semen screening in the father (Fig. 2A, Supplementary Table SIV). In the remaining eight families, the risk-haplotype could not be established either due to the complete absence of iSNVs (n = 5/8) in the targeted regions or due to an inability to successfully amplify and/or sequence the *PKD1/PKD2* loci (n = 3/8), known for their genomic complexity, GC-rich content and extensive allelic heterogeneity (Audrézet *et al.*, 2012).

## Outcome of clinical universal PGT cycles

Based on the results of the preclinical PGT workup, 12/24 successfully analyzed couples underwent 28 IVF-PGT cycles (with an average of




| CHR | Position | REF | ALT | Father | Mother | Index | Phasing | Index H1 | Index H2 |
|---|---|---|---|---|---|---|---|---|---|
| 17 | 31231064 | T | C | 0/0 | 1\|0 | 0\|1 | Maternal | T | **C** |
| 17 | 31231212 | G | A | 0\|1 | 0/0 | 1\|0 | Paternal | **A** | G |
| 17 | 31231275 | T | G | 1/1 | 1/1 | 1/1 | - | G | G |
| 17 | 31232570 | G | C | 0/0 | 1\|0 | 0\|1 | Maternal | G | **C** |
| 17 | 31232914 | C | A | 0/0 | 1\|0 | 0\|1 | Maternal | C | **A** |
| 17 | 31233160 | G | T | 0/0 | 0/0 | 1\|0 | *De novo* | T | G |
| 17 | 31233515 | G | A | 0/0 | 0\|1 | 0/0 | - | G | G |
| 17 | 31233660 | A | C | 0/0 | 1\|0 | 0\|1 | Maternal | A | **C** |
| 17 | 31233759 | T | G | 0/0 | 1\|0 | 0\|1 | Maternal | T | **G** |

**Figure 1. Long-read amplicon sequencing-based preclinical PGT workup.** (**A**) Schematic workflow of the preclinical preimplantation genetic testing (PGT) workup for a family with a *de novo* pathogenic variant. (1) First, 5–10 kb amplicons, flanking the *de novo* variant, are generated using long-range PCR and long-read amplicon sequencing is performed for the affected individual (Index) and his/her parents (Father and Mother, respectively). (2) Next, sequenced long reads mapped to the region of interest were used for targeted variant and haplotype calling. The inheritance of informative single-nucleotide variants (iSNVs; black dots) is then linked to the *de novo* pathogenic variant (red stars), allowing to determine the parental origin of the disease-associated allele. In the provided example, the Index inherited H1 from the Father and H2 from the Mother, with the *de novo* variant linked to the paternal allele. (3) Following the preclinical PGT workup, the couple can proceed to universal PGT, using the parents of the affected partner (i.e. 'grandparents') as the phasing reference. (**B**) Example of a family (PGD292) with a *de novo* pathogenic variant in the NF1 gene in the affected male partner (Index). Based on the inheritance of maternal and paternal iSNVs (marked in bold), the *de novo* variant was linked to the paternal allele. CHR: chromosome; H1: haplotype 1; H2: haplotype 2; REF: reference allele; ALT: alternative allele; 0/0: homozygous reference variant; 0|1 or 1|0: heterozygous variant; 1/1: homozygous alternative variant.

2.3 cycles per couple and 3.4 embryos per cycle). All embryos were analyzed by siCHILD/haplarithmisis-based universal PGT (Zamani Esteki *et al.*, 2015), using 'grandparents' (i.e. parents of the affected partner) as the phasing reference. For family PGD743, affected children were used as the phasing reference. In total, 97 embryos were analyzed, resulting in 30 disease-free (unaffected) embryos, 47 affected (or 'at risk' in case of the germline mosaicism) embryos and 16 abnormal embryos due genome-wide aneuploidy (Supplementary Table SV). Based on the genome-wide information, unaffected embryos were further ranked for transfer, according to internal guidelines (Dimitriadou *et al.*, 2017). Four embryos were categorized as 'inconclusive' due to homologous recombination in the biopsied sample in the region of interest. Finally, in cycles with a known clinical outcome at the time of study, 12 unaffected embryos have been transferred, resulting in six implantation failures, one late miscarriage, two ongoing pregnancies and three deliveries. All born children were disease-free.

## Embryo direct mutation analysis coupled with long-read amplicon haplotyping

In parallel to universal PGT, we aimed to evaluate the performance of direct mutation analysis via the same long-read amplicon sequencing workflow used for preclinical PGT workup. To do this, we have selected 13 whole-genome amplified TE biopsies from three families (PGD656, PGD669 and PGD743) that have already been processed by universal PGT using siCHILD/haplarithmisis. For families PGD656 and PGD743, the corresponding unaffected partner was also included in the analysis; for PGD669, only embryos were processed further, because the female partner carried a *de novo* 3 kb mosaic deletion in *COL4A5* gene, the inheritance of which is easily detectable by sequencing. Since primer pairs had already been designed and validated during the preclinical PGT step, no additional workup was required. For PGD669, direct mutation analysis was in line with the results of universal PGT cycle (Supplementary Fig. S1A and Table SV). For PGD743 with identified maternal germline mosaicism, all embryos inherited a risk allele from the mother, but did not carry the pathogenic variant itself (Fig. 2B and C, Supplementary Table SIV). Although these embryos were unaffected by the mutation in question, none of them would have been chosen for transfer due to the presence of various aneuploidies, including meiotic trisomy 22 (Supplementary Table SV). This in turn highlights the value of genome-wide information in universal PGT cycle, as embryos with meiotic aneuploidy are immediately discarded from further use. Finally, for PGD656, one embryo was classified as 'inconclusive' by universal PGT due to homologous

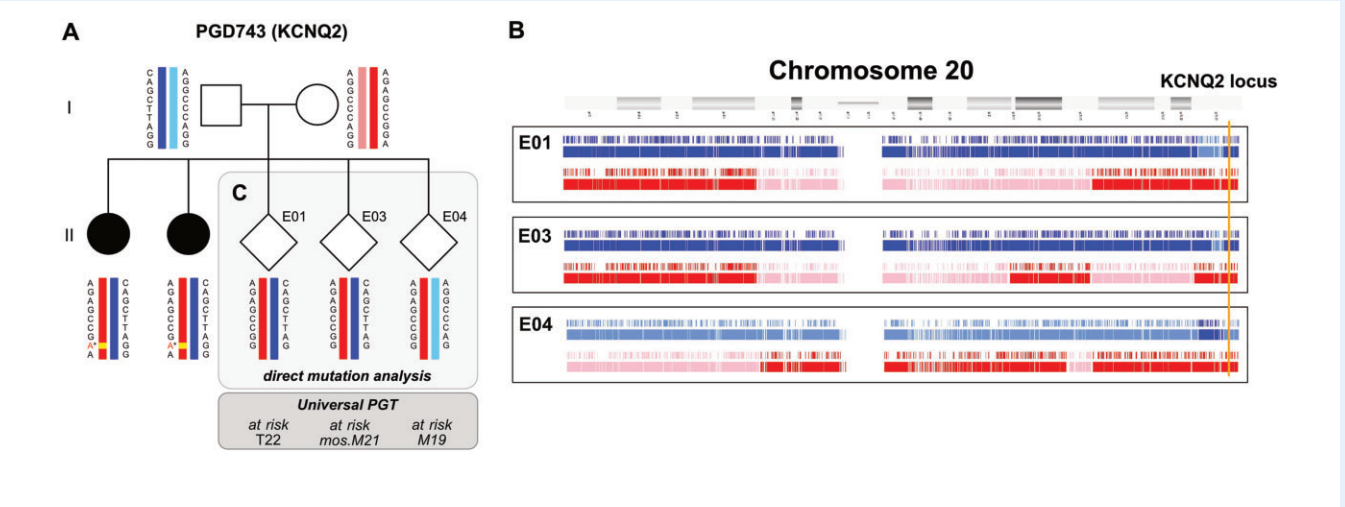



**Figure 2. Preclinical PGT workup and clinical PGT cycle in a family with suspected germline mosaicism.** (**A**) Preclinical preimplantation genetic testing (PGT) workup is shown for family (PGD743) with a seemingly *de novo* pathogenic variant in *KCNQ2* gene in two children. Long-read amplicon analysis showed that pathogenic variants in both children (indicated in yellow) are located on the same maternal allele (dark red), confirming the suspicion of maternal germline mosaicism. (**B**) Results of universal PGT using affected siblings as the phasing reference. The orange line indicates the region of interest on chromosome 20. For each embryo, parental haplotype blocks are depicted: light/dark blue and pink/red haplotypes represent paternal and maternal alleles, respectively. The switch between haplotype blocks indicates homologous recombination. All three embryos inherited the risk allele from the mother (dark red haplotype). (**C**) Targeted long-read sequencing of TE biopsies from three embryos. Based on inheritance of iSNVs (also Supplementary Table SV), all three embryos inherited the risk allele (red) from the mother but were deprived of the mutation. Embryo (E) 04 also inherited a different paternal allele, compared to embryos 1 and 3, confirming the results of universal PGT analysis. iSNVs: informative single-nucleotide variants; M: monosomy; T: trisomy; mos.: mosaic.

recombination in the region of interest, but turned out to be a mutation carrier, although in a homozygous state potentially due to allele drop out (ADO) (Supplementary Fig. S1B).

## Discussion

Improvements in accuracy and throughput have the potential to put LRS technologies centerstage in medical genetics (as reviewed by Mantere *et al.* (2019)). LRS in combination with linkage analysis has also been successfully applied in PGT-SR for reciprocal translocations (Chow *et al.*, 2020; Madjunkova *et al.*, 2020; Liu *et al.*, 2021; Yanfei *et al.*, 2021) and pathogenic inversions (Zhang *et al.*, 2019) with the possibility to differentiate carriers from non-carrier embryos. The use of LRS in a few PGT-M cases has also been reported (Liu *et al.*, 2021; Yanfei *et al.*, 2021). In two studies, LRS analysis for PGT-SR and PGT-M for alpha thalassemia was performed directly on embryos (Madjunkova *et al.*, 2020; Liu *et al.*, 2021). In other studies, LRS was first used in prospective parents to characterize the breakpoint regions or map SNPs in region of interest in cases of PGT-SR and PGT-M, respectively. The embryos were then analyzed either by targeted (Yanfei *et al.*, 2021) or genome-wide PGT approaches (Zhang *et al.*, 2019).

In this study, we implemented long-read amplicon sequencing as part of preclinical PGT workup for couples, in which one of the partners carries a *de novo* pathogenic variant. The *de novo* variant was successfully linked to corresponding inherited parental iSNVs in 75% of cases. From successfully analyzed families, the majority of *de novo* mutations occurred on the paternal allele. Given that in at least half of our cases the indication involves NF1, this can be a disease-specific effect: *NF1* has one of the highest mutation rates (Clementi *et al.*, 1990) and most *de novo* mutations are paternal in origin (Jadayel *et al.*, 1990; Stephens *et al.*, 1992). This is in line with our observation, as the paternal allele was affected in 76.4% of all *NF1* cases (n = 13/17). At the same time, it has been shown that overall around 80% of *de novo* variants are transmitted from the father, and the number of germline *de novo* mutations is associated with increased paternal age (Wong *et al.*, 2016), likely as a result of mutation accumulation during the continuous cell division in spermatogenesis.

The benefit of LRS-based preclinical PGT workup strategy is that it only requires DNA samples from the affected individual and his/her parents. Moreover, unlike germ cell or embryo STR analysis, which traces haplotype inheritance and relies on multiple flanking STR markers to rule out ADO and homologous recombination (ESHRE PGT-M Working Group *et al.*, 2020), the current workup does not require the presence of iSNVs on each side of the *de novo* variant. Instead, by following iSNV segregation, merely parental origin of the mutated allele is investigated. Even if a recombination would occur in an affected individual in the locus of interest, the parental origin of inherited iSNVs linked with the *de novo* variant would remain the same. To increase diagnostic accuracy, the presence of at least two iSNVs might be warranted, however the detection of one iSNV can be sufficient to link the pathogenic variant to the inherited paternal or maternal allele, given that it is validated by Sanger sequencing and/or by an independent long-range amplicon sequencing experiment that

confirms the iSNV segregation pattern. Moreover, this preclinical PGT workup consequently allows genome-wide haplotyping-based universal PGT to be performed using parents of the affected individual as the phasing reference. During universal PGT only embryos deprived of both the disease-associated haplotype and aneuploidy are selected, thus avoiding transfer of embryos with low developmental potential, and likely increasing the clinical pregnancy rate per transfer (Rechitsky *et al.*, 2015; Minasi *et al.*, 2017).

Targeted long-read amplicon sequencing can also be applied to deduce the parental origin of *de novo* pathogenic variants in children. Here, we have used this approach to confirm the germline mosaicism in the mother (age 36 at the time of the first offspring conception), after the birth of a second child with the same, at first seemingly *de novo* pathogenic variant in the *KCNQ2* gene, and negative semen screening in the father. Following this analysis, the couple underwent universal PGT cycle. Hence, deducing parental origin of the mutated allele can also provide reproductive options for couples, who have an affected child and/or are at risk of gonadal mosaicism. From a diagnostic perspective, considering that most *de novo* mutations occur on the paternal allele, sperm analysis could also be offered in routine genetic screening for germline mosaicism in fathers of patients with *de novo* mutations.

The limitation of this approach is that it remains targeted with the need to design locus-specific primers. Although possible, preclinical workup for *de novo* microdeletions/microduplications or complex genomic regions (such as PKD1/PKD2 loci) remains challenging for primer design, and alternatives may be explored. The region of interest can also be non-informative in the absence of iSNVs. This accounted for one in four of all couples in our cohort. Hence, patients should be counseled about the current limitations of the technology and further developments are necessary to be able to offer universal PGT for couples with a high risk of a non-informative result. The second limitation is the need to have DNA from the parent(s) of the affected partner for the analysis, which may not always be available. However, as shown by our data, embryo direct mutation analysis on TE biopsy, coupled with targeted haplotyping, can offer an alternative PGT approach for couples where no other family members are available. The presence of iSNVs near the mutation site largely excludes homologous recombination events in the region of interest, reducing the chance of misdiagnosis. Importantly, in our embryo data, the direct mutation detection was supported by linkage of flanking parental SNVs, as pathogenic variants always segregated with the mutant haplotype of one of the parents. The exception was the germline mosaic case, where embryos inherited the risk allele but were deprived of the mutation. However, since the effect of ADO cannot be excluded, caution must be applied when interpreting the results of embryo direct mutation analysis in germline mosaicism cases. Finally, embryos diagnosed as mutation carriers by targeted long-read haplotyping can be used as the phasing reference for universal PGT, providing reliable haplotype and copy number information, as was also recently demonstrated (De Witte *et al.*, 2022).

In conclusion, we show the reliable accuracy of haplotype phasing using long-read amplicon sequencing, particularly ONT. In comparison to PacBio, which requires high capital equipment costs, the ONT MinION platform used in this study is a portable benchtop sequencer that can be run in real time from an average laptop. In addition, due to the low cost per sample, improved quality of the data and a relatively straightforward workflow, similar to any other NGS techniques, ONT long-read amplicon sequencing represents a valuable preclinical PGT workup strategy for couples with *de novo* pathogenic variants. With all the advances we are witnessing in the field of LRS, faster and more accurate methods will arise and will likely simplify the whole PGT process ultimately eliminating the preclinical PGT step or the need for additional family members in PGT patients.

## Supplementary data

Supplementary data are available at *Human Reproduction* online.

## Data availability

In compliance with the EU General Data Protection Regulation (GDPR), the patient dataset used in the study is not publicly available. Raw long-read sequencing data are available to academic users upon request.

## Acknowledgments

We would like to thank the couples who participated in the study. We especially want to thank Silke Hollants for the help in setting up the initial protocol and Greet Peeters for her support with all wet lab-related procedures.

## Authors' roles

Y.E.A. and O.T. wrote the original manuscript; E.L. and H.B. developed the original long-read amplicon sequencing protocol; T.J., J.A. and J.D. performed the computational analysis; Y.E.A., O.T., T.J., H.B., C.M. and Ef.D. performed long-read data interpretation; O.T., C.M. and Ef.D. performed embryo PGT data interpretation; S.D. provided the clinical PGT cycle data; S.D., K.P., A.V., A.D.L., C.P. and C.K. were responsible for the IVF-PGT associated procedures; K.P., A.V., E.L. and El.D. counseled the patients and provided patient data. J.R.V., E.L., H.B. and Ef.D. conceived the study. O.T., J.R.V. and Ef.D. supervised the study. All authors read and approved the final version of the manuscript.

## Funding

This work was supported by KU Leuven funding (no. C1/018 to J.R.V.) and Fonds Wetenschappelijk Onderzoek (1241121N to O.T.).

## Conflict of interest

J.R.V. is co-inventor of a patent 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' licensed to Agilent Technologies. All other authors have no conflict of interest to declare.

## References


Ahsan MU, Liu Q, Fang L, Wang K. NanoCaller for accurate detection of SNPs and indels in difficult-to-map regions from long-read sequencing by haplotype-aware deep neural networks. *Genome Biol* 2021;**22**:1–33.

Audrézet MP, Cornec-Le Gall E, Chen JM, Redon S, Quéré I, Creff J, Bénech C, Maestri S, Meur YL, Férec C. Autosomal dominant polycystic kidney disease: comprehensive mutation analysis of PKD1 and PKD2 in 700 unrelated patients. *Hum Mutat* 2012;**33**:1239–1250.

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.

Chow JFC, Cheng HHY, Lau EYL, Yeung WSB, Ng EHY. Distinguishing between carrier and noncarrier embryos with the use of long-read sequencing in preimplantation genetic testing for reciprocal translocations. *Genomics* 2020;**112**:494–500.

Clementi M, Barbujani G, Turolla L, Tenconi R. Neurofibromatosis-1: a maximum likelihood estimation of mutation rate. *Hum Genet* 1990;**84**:116–118.

Coonen E, van Montfoort A, Carvalho F, Kokkali G, Moutou C, Rubio C, De Rycke M, Goossens V. ESHRE PGT Consortium data collection XVI–XVIII: cycles from 2013 to 2015. *Hum Reprod Open* 2020;**2020**:hoaa043.

Crossley BM, Bai J, Glaser A, Maes R, Porter E, Killian ML, Clement T, Toohey-Kurth K. Guidelines for Sanger sequencing and molecular assay monitoring. *J Vet Diagn Invest* 2020;**32**:767–775.

De Rycke M, Goossens V, Kokkali G, Meijer-Hoogeveen M, Coonen E, Moutou C. ESHRE PGD Consortium data collection XIV-XV: cycles from January 2011 to December 2012 with pregnancy follow-up to October 2013. *Hum Reprod* 2017;**32**:1974–1994.

De Witte L, Raman L, Baetens M, De Koker A, Callewaert N, Symoens S, Tilleman K, Vanden MF, 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.

Debrock S, Melotte C, Spiessens C, Peeraer K, Vanneste E, Meeuwis L, Meuleman C, Frijns J-P, Vermeesch JR, D'Hooghe TM. Preimplantation genetic screening for aneuploidy of embryos after *in vitro* fertilization in women aged at least 35 years: a prospective randomized trial. *Fertil Steril* 2010;**93**:364–373.

Dimitriadou E, Melotte C, Debrock S, Esteki MZ, Dierickx K, Voet T, Devriendt K, de Ravel T, Legius E, Peeraer K *et al.* Principles guiding embryo selection following genome-wide haplotyping of preimplantation embryos. *Hum Reprod* 2017;**32**:687–697.

Ding J, Dimitriadou E, Tšuiko O, Destouni A, Melotte C, Bogaert K, Van den Debrock S, Jatsenko T, Zamani Esteki M, Voet T *et al.* Identity-by-state-based haplotyping expands the application of comprehensive preimplantation genetic testing. *Hum Reprod* 2020;**35**:718–726.

Edge P, Bansal V. Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat Commun* 2019;**10**:1–10.

ESHRE PGT-M Working Group, F Carvalho, C Moutou, E Dimitriadou, J Dreesen, C Giménez, V Goossens, G Kakourou, N Vermeulen, D Zuccarello *et al.* ESHRE PGT Consortium good practice recommendations for the detection of monogenic disorders. *Hum Reprod Open* 2020;**2020**:hoaa018.

Handyside AH, Harton GL, Mariani B, Thornhill AR, Affara N, Shaw M-A, 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.

Jadayel D, Fain P, Upadhyaya M, Ponder MA, Huson SM, Carey J, Fryer A, Mathew CGP, Barker DF, Ponder BAJ. Paternal origin of new mutations in von Recklinghausen neurofibromatosis. *Nature* 1990;**343**:558–559.

Liu S, Wang H, Leigh D, Cram DS, Wang L, Yao Y. Third-generation sequencing: any future opportunities for PGT? *J Assist Reprod Genet* 2021;**38**:357–364.

Madjunkova S, Sundaravadanam Y, Antes R, Abramov R, Chen S, Yin Y, Zuzarte PC, Moskovtsev SI, Jorgensen LGT, Baratz A *et al.* Detection of structural rearrangements in embryos. *N Engl J Med* 2020;**382**:2472–2474.

Mantere T, Kersten S, Hoischen A. Long-read sequencing emerging in medical genetics. *Front Genet* 2019;**10**:426.

Masset H, Ding J, Dimitriadou E, Debrock S, Tšuiko 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, Drüsedau M, Meekels J *et al.* Multi-centre evaluation of a comprehensive preimplantation genetic test through haplotyping-by-sequencing. *Hum Reprod* 2019;**34**:1608–1619.

Minasi MG, Fiorentino F, Ruberti A, Biricik A, Cursio E, Cotroneo E, Varricchio MT, Surdo M, Spinella F, Greco E. Genetic diseases and aneuploidies can be detected with a single blastocyst biopsy: a successful clinical approach. *Hum Reprod* 2017;**32**:1770–1777.

Rechitsky S, Pakhalchuk T, San Ramos G, Goodman A, Zlatopolsky Z, Kuliev A. First systematic experience of preimplantation genetic diagnosis for single-gene disorders, and/or preimplantation human leukocyte antigen typing, combined with 24-chromosome aneuploidy testing. *Fertil Steril* 2015;**103**:503–512.

Rechitsky S, Pomerantseva E, Pakhalchuk T, Pauling D, Verlinsky O, Kuliev A. First systematic experience of preimplantation genetic diagnosis for de-novo mutations. *Reprod Biomed Online* 2011;**22**:350–361.

Stephens K, Kayes L, Riccardi VM, Rising M, Sybert VP, Pagon RA. Preferential mutation of the neurofibromatosis type 1 gene in paternally derived chromosomes. *Hum Genet* 1992;**88**:279–282.

Tšuiko O, Vanneste M, Melotte C, Ding J, Debrock S, Masset H, Peters M, Salumets A, De Leener A, Pirard C *et al.* Haplotyping-based preimplantation genetic testing reveals parent-of-origin specific mechanisms of aneuploidy formation. *NPJ Genom Med* 2021;**6**:81.

Wong WSW, Solomon BD, Bodian DL, Kothiyal P, Eley G, Huddleston KC, Baker R, Thach DC, Iyer RK, Vockley JG *et al.* New observations on maternal age effect on germline *de novo* mutations. *Nat Commun* 2016;**7**:1–10.

Yanfei YC, Yu Q, Ma M, Wang H, Tian S, Zhang W, Jinning Zhang MM, Liu Y, Yang Q, Pan X *et al*. Variant haplophasing by long-read sequencing: a new approach to preimplantation genetic testing workups. *Fertil Steril* 2021;**116**:774–783.

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.

Zhang S, Liang F, Lei C, Wu J, Fu J, Yang Q, Luo X, Yu G, Wang D, Zhang Y *et al*. Long-read sequencing and haplotype linkage analysis enabled preimplantation genetic testing for patients carrying pathogenic inversions. *J Med Genet* 2019;**56**:741–749.

1 STEP AHEAD

SAGE 1-Step™ GM-CSF medium

# Put your patients one step ahead

- SAGE 1-Step™ GM-CSF medium is the first single-step culture and transfer medium containing the GM-CSF cytokine and Hyaluronan.
- GM-CSF supports embryo-endometrial communication for improved implantation and chances of pregnancy.[1]
- Provides an additional treatment option for patients with a previous failed treatment cycle.[2]


**Upgrade your lab with innovation that matters.**

Click to explore

CooperSurgical®

1. Hardy, K. & Spanos, S. (2002) Growth factor expression and function in the human and mouse preimplantation embryo. J Endocrinol 172: 221-236.
2. Sipahi, M., Mümüşoğlu, S. et al. (2021). The impact of using culture media containing granulocyte-macrophage colony-stimulating factor on live birth rates in patients with a history of embryonic developmental arrest in previous in vitro fertilization cycles. Journal of the Turkish German Gynecological Association, 22(3), 181–186.