BMC Genomics

RESEARCH

Open Access

Generation of quartet reference materials for preimplantation genetic testing for structural rearrangements

Nan Sun[1†], Wenxin Zhang[1†], Dandan Yu[2†], Mi Zhang[1†], Yang Zou[3], Ximei Dong[2], Yiping Wang[1], Nana Song[2], Yaya Guo[2], Chen Jiang[2], Weifeng Wang[2], Lingyin Kong[4], Bo Liang[4,5], Jia Fei[6], Sijia Lu[7], Zhiyu Peng[2], Ting Yu[1*], Jiale Xiang[2*] and Shoufang Qu[1*]

Abstract

Background Preimplantation genetic testing for structural rearrangements (PGT-SR) facilitates the selection of embryos with balanced karyotype prior to implantation. Currently, there is a critical need for reference materials (RMs) to ensure PGT-SR development.

Results This study designed a novel and comprehensive strategy to produce renewable RMs using quartet families of structural rearrangements. A panel of quartet RMs for PGT-SR was successfully developed from 13 families, including 8 reciprocal translocations, 2 Robertsonian translocations, 2 inversions, and 1 normal family. Each quartet RMs comprised 3 tubes of DNA and 1 tube of sorted cells mimicking embryonic cells, all passed stringent quality assessments. The panel of quartet RMs was characterized by three volunteer laboratories with 100% accuracy and 100% specificity, using different sequencing platforms to assess the biomimetics. Additionally, while both short-read and long-read genome sequencing could detect exact breakpoints of the reciprocal translocations and standard inversion, long-read genome sequencing outperformed short-read sequencing for complex structural rearrangements.

Conclusions This panel of quartet RMs for PGT-SR is well-characterized, renewable, and publicly accessible, playing a crucial role in proficiency test and quality assurance for structural rearrangement detection. The methods developed in this study are adaptable for other preimplantation genetic tests, thereby enhancing the advancement of assisted reproduction techniques.

Keywords PGT-SR, Reference materials, Quartet family, Immortalized cell lines, Nanopore sequencing, WGS

[†]Nan Sun, Wenxin Zhang, Dandan Yu and Mi Zhang contributed equally to this work.

*Correspondence:
Ting Yu
yuting@nifdc.org.cn
Jiale Xiang
xiangjiale@bgi.com
Shoufang Qu
qushoufang@126.com
[1]National Institutes for Food and Drug Control, Dongcheng District tiantanxili 2#, Beijing 100050, China
[2]BGI Genomics, BGI Center No.9, Yunhua Road, Yantian District, 518053 Shenzhen, China
[3]Institute of Women, Children and Reproductive Health, Shandong University, Jinan, China
[4]Basecare Medical Device Co., Ltd, Suzhou, China
[5]State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China
[6]Peking Jabrehoo Med Tech Co., Ltd, Beijing, China
[7]Yikon Genomics Company, Ltd, Suzhou, China

Background

Structural rearrangement (SR) are abnormalities that do not alter the overall quantity of genetic materials but can lead to the production of unbalanced gametes, potentially causing adverse effects on offspring [1]. It primarily encompasses inversions and balanced translocations. In the general population, the prevalence of balanced translocation stands at approximately 0.25% [2]. More specifically, balanced translocations consist of reciprocal translocations and Robertsonian translocations. Reciprocal translocation occurs when two terminal segments from non-homologous chromosomes exchange places, resulting in the formation of two derivative chromosomes. It is estimated that the incidence of reciprocal translocation is 1 in 600 live births [2]. In contrast, Robertsonian translocation occurs when two long arms from two acrocentric chromosomes (typically chromosomes 13, 14, 15, 21, 22) break near the centromeres or peri-centromeric regions and joint together at the centromeres to form a derivative chromosome, reducing the number of chromosomes to 45 in the karyotype. The reported incidence of Robertsonian translocation is 1 in 1085 live births [3, 4]. As for inversions, they occur simultaneous breaks in a chromosome lead to the segments reattaching after a 180-degree flip. Inversion can be further categorized into two types based on whether the breaks occur on the same arm of the chromosome (paracentric inversion) or not (pericentric inversion) [5]. The incidence of inversions is estimated to be 1 in 833 live births [6].

Individuals carrying SRs commonly face elevated risks of infertility, recurrent miscarriage, or having offspring with birth defects, as they produce over 60% unbalanced gametes due to improper chromosome pairing and unequal exchanges during meiosis [7–9]. Studies indicate that patients with reciprocal translocations, Robertsonian translocations and inversions experience frequencies of implantation failure or miscarriages at 68.1%, 36.4% and 42.9% respectively [10]. Consequently, couples with at least one carrier of balanced translocation have a reduced chance of conceiving a normal fetus or balanced carrier, ranging from 34.5% to 50.5% [11–13], significantly lower than that of unaffected couples. Therefore, the carrier rate of SRs in couples experiencing recurrent miscarriages or infertility is significantly higher than in general population. Among infertility couples, approximately 2.26% carry balanced translocation, with 1.35% having reciprocal translocations, 0.77% having Robertsonian translocations, and 0.14% having inversions [14]. Among couples experiencing recurrent miscarriages, approximate 4.5%-5% are carriers of balanced translocation [15–17]. Importantly, the rate of inherited balanced SRs in fetuses is 0.27% compared to 0.076–0.096% for de novo balanced SRs, suggesting that most of the balanced SR are inherited [18]. Therefore, the use of preimplantation genetic testing (PGT) technology is crucial to assist reproduction for individuals carrying SRs.

The techniques of PGT include detection of aneuploidy (PGT-A), chromosomal structural rearrangement (PGT-SR), and monogenic disease (PGT-M). PGT-SR has been widely adopted to select normal or balanced embryos prior to implantation for couples with at least one carrier of balanced SR, thereby reducing the chance of termination of pregnancy or recurrent miscarriages [11]. It is reported that before PGT-SR testing, 83.8% of pregnancies were miscarriages and 13.3% had congenital disabilities or induced labor, with only 2.9% gave birth to normal infant in 592 pregnancies. With the help of PGT-SR, the rate of normal newborns increased to 85.6% in 118 pregnancies, showing great improvement in reproductive medicine [19].

There are several methods for PGT-SR, such as fluorescence in situ hybridization (FISH), single nucleotide polymorphism (SNP) array, comparative genomic hybridization (CGH) and massive parallel sequencing (MPS), each with its specific pros and cons. FISH is the original diagnostic approach yet limited by the complicated requirements of probe preparing and testing [20]. The other three methods are all based on the detection of SNPs. At a frequency of every 1 in 1000 nucleotides, SNPs are the most common variant that have no effect on development or health. SNP-based PGT-SR is a target test performed when parental genomes have known chromosomal abnormalities. It requires the construction of a core family model for haplotype analysis with parents and a proband or reference embryos. Then it conducts the haplotype analysis on the testing embryos in the region adjacent to SR breakpoints on the chromosomes [1]. Using informative SNP for linkage analysis, PGT-SR can differentiate between unbalanced, balanced and normal embryos for chromosomal translocation or inversion. It is essential when the translocation affects the chromosome X due to the random inactivation in female embryos [21]. Overall, PGT-SR provides effective testing for carrier couples to avoid implantation failure, abortion or having congenitally affected newborns.

Reference materials (RMs), including genomic DNA (gDNA) and cells, are recommended by professional guidelines for assay development, quality control and validation [22–24]. However, there has been a lack of characterized and renewable RM specifically tailored for PGT-SR, hindering the monitoring of test performance and consistency. Additionally, existing RMs have primarily focused on specific genes or disorders, such as BRCA1/BRCA2, MTHFR, HLA loci, cystic fibrosis, Rett syndrome, Huntington disease [25–29], neglecting the comprehensive needs of PGT-SR. In the context of PGT-SR, the development of renewable cell RMs alongside gDNA RMs is crucial. While typically 5 to 7

trophectoderm cells are dissected from the embryos for whole genome amplification (WGA) during PGT-SR, their limited quantity renders them unsuitable as direct cell RMs. Furthermore, gDNA from parents and a relative is required for SNP-based haplotype analysis, necessitating the use of synthetic samples without kinship as RMs for PGT-SR. Moreover, to ensure clinical relevance, the panel of RMs should encompass various types of SRs across different chromosomes. Last, efficient large-scale production and comprehensive characterization of RMs across different laboratories are essential to ensure uniformity and reliability.

To address these challenges, we were determined to designed a new diagram and comprehensive strategy for the development of quartet RMs tailored specifically for PGT-SR. This initiative involved the creation of both cell RMs and gDNA RMs from quartet families. This panel of quartet RMs would undergo rigorous characterization by three independent laboratories to assess their biomimetic properties.

Methods

Family recruited and samples collection

This study was approved by the Institutional Review Board of Beijing Genomics Institute (BGI, No. BGI-IRB 22010) and the Ethics Committee of the First People's Hospital of Chenzhou (No. 2022003 H), China. 13 quartet families with a couple and two offspring were recruited at the First People's Hospital of Chenzhou, China. All participants gave informed consent. Peripheral bloods of all participants were harvested into two EDTA vacutainers with at least 5mL per tube and shipped to BGI genomics.

Cell immortalization and analysis of short tandem repeats (STR)

Lymphocytes were isolated from peripheral blood and infected by Epstein-Barr viruses (EBVs) to establish immortalized cell lines according to previous study [30]. To prevent EBV contamination, the handling of EBV was restricted in class II biosafety cabinets (BSL-2). All viral supernatants and equipment must be inactivated and sterilized thoroughly before disposal. Immortalized cells were also harvested to perform STR analysis to prevent cell cross-contamination using Human DNA Typing Kit (BGI, China) according to the manufacturer's instruction.

Preparation of DNA RMs and quality control

Immortalized cells were harvested to extract gDNA using the MagPure Buffy Coat DNA Midi KF Kit (MAGEN, UK) following the manufacturer's instructions. DNA concentration was measured by a Qubit® 3.0 Fluorometer (Life Technologies, USA). DNA purity was monitored by detecting the optical density (OD) using a NanoDrop™ (Agilent Technologies, USA) since OD260/280 and OD260/230 were calculated to see whether they were polluted. DNA integrity was evaluated with agarose gel electrophoresis to see whether they were degraded. DNA sample was aliquoted into 15 μL per tube at a concentration of 20 ng/μL.

Preparation of cell RMs and quality control

Immortalized cells were filtered to perform cell sorting on FACS Aria™ II flow cytometer (Becton, Dickinson and Company, USA). Briefly, immortalized cells were resuspended in phosphate buffered saline (PBS) supplemented with 0.1% bovine serum albumin and 1mM EDTA at a concentration of approximately 1 × 10^7 cells/mL. After filtering through a 40 μm nylon mesh, these unstained cells were loaded into the flow cytometer with a 100 μm nozzle and gated in a dop plot of forward scatter area (FSC-A) versus side scatter area (SSC-A). Pre-sort testing was performed with Accudrop beads following the manufacturers' instruction. "Single-cell" sort mode at "7 cells" parameters setting was selected since the number of cells were consistently 5–10 when sorted onto a glass slide and observed under a light microscope. Subsequently, each cell lines were sorted into 100 PCR tubes (0.2 mL) with 5 μL PBS. 3 tubes from each cell line were randomly picked to perform WGA using the REPLI-g Single Cell Kit (QIAGEN, Germany). The concentration of WGA products was measured and compared with the clinical embryonic samples to see whether there was significant difference.

Characterization of quartet RMs by three laboratories

Three sets of quartet RMs were separately sent to three laboratories for characterization. PGT-SR were performed in three laboratories as described previously, including Basecare Medical laboratory (laboratory 1) [31], Jabrehoo laboratory (laboratory 2) [32] and Yikon Genomics laboratory (laboratory 3) [33, 34]. Briefly, laboratory 1 adopted REPLI-g Single Cell Kit (QIAGEN, German) to perform WGA using cell RMs. The libraries were constructed as previous stated [31] from gDNA RMs and WGA product, followed by sequencing on DA5000 platform by pair-end 100 bp. Genotype were analyzed at 1 Mb around the breakpoints. Laboratory 2 used SurePlex DNA Amplification System (Illumina, USA) for WGA using cell RMs and HumanCytoSNP-12 BeadChip (Illumina, USA) for SNP array for gDNA RMs and WGA product. The libraries constructed were sequenced using Miseq DX platform (Illumina, USA) by pair-end 150 bp. Genotype were analyzed at 2 Mb around the breakpoints. For laboratory 3, Prenatal Abnormality Detection Kit for Chromosomal Structural (Combined Probe Anchored Sequencing Method) (Yikon, China) was used for WGA and library construction according to the instructions, followed by sequencing on DNBSEQ-T7 (MGI, China) by

pair-end 150 bp. Genotype were analyzed at 5 Mb around the breakpoints.

Breakpoint determination

The breakpoints of SR were detected by short-read whole genome sequencing (WGS) and long-read WGS sequencing (Nanopore) to compared the performance of these two techniques. As a gold-standard method, the breakpoints of SR were further verified and precisely determined by Sanger sequencing.

For short-read WGS, gDNA was sheared to fragments of around 400 bp using Custom 5X DNA shearing master mix (Enzymatics, USA) and was then size-selected with VAHTS™ DNA Clean Beads (VAZYME, China). Libraries were prepared using WGS Library Preparation Kit (BGI, China) and sequenced on DNBSEQ-T7 (MGI, Shenzhen) sequencer for 100 bp at paired end. Standard WGS analysis was performed using the LUSH Toolkit software, with a required sequencing depth of over 10x.

For Nanopore sequencing, samples were size-selected with Pippin HT (Sage Science, USA) to remove DNA fragments ≤ 10 kb. Genomic libraries were prepared using the Ligation Sequencing Kit SQK-LSK110 (Oxford Nanopore, UK) and Native Barcoding Expansion kit NBD196 (Oxford Nanopore, UK). The libraries were sequenced using PromethION48 (Oxford Nanopore, UK). The Fastp software is used for quality control (length = 150, Q value = 7), and the filtered data were analyzed using the Minimap2 software.

For Sanger sequencing, primers of polymerase chain reaction (PCR) were designed using Primer 5.0 software. PCR was performed using 2x Super PCR Mix (with dye) (BGI, China), and the products were electrophoresed on a 1.0% agarose gel. Sanger sequencing was run on ABI 3730XL sequencer (Applied Biosystems, USA).

Statistical analysis

Circos plot was drawn using shinyCircos-V2.0 [35] according to the guidance. The Chi-square test was used to assess differences between groups and a value of $P < 0.05$ was considered significantly different.

Results

Design the flowchart for the development of quartet RMs

To create a comprehensive panel of RMs mirroring real-world clinical scenarios, we devised a flowchart for establishing quartet RMs specifically tailored for PGT-SR (Fig. 1). Briefly, we recruited and collected the blood samples from quartet families, comprising a couple and two offspring, wherein at least one parent carried a SR.

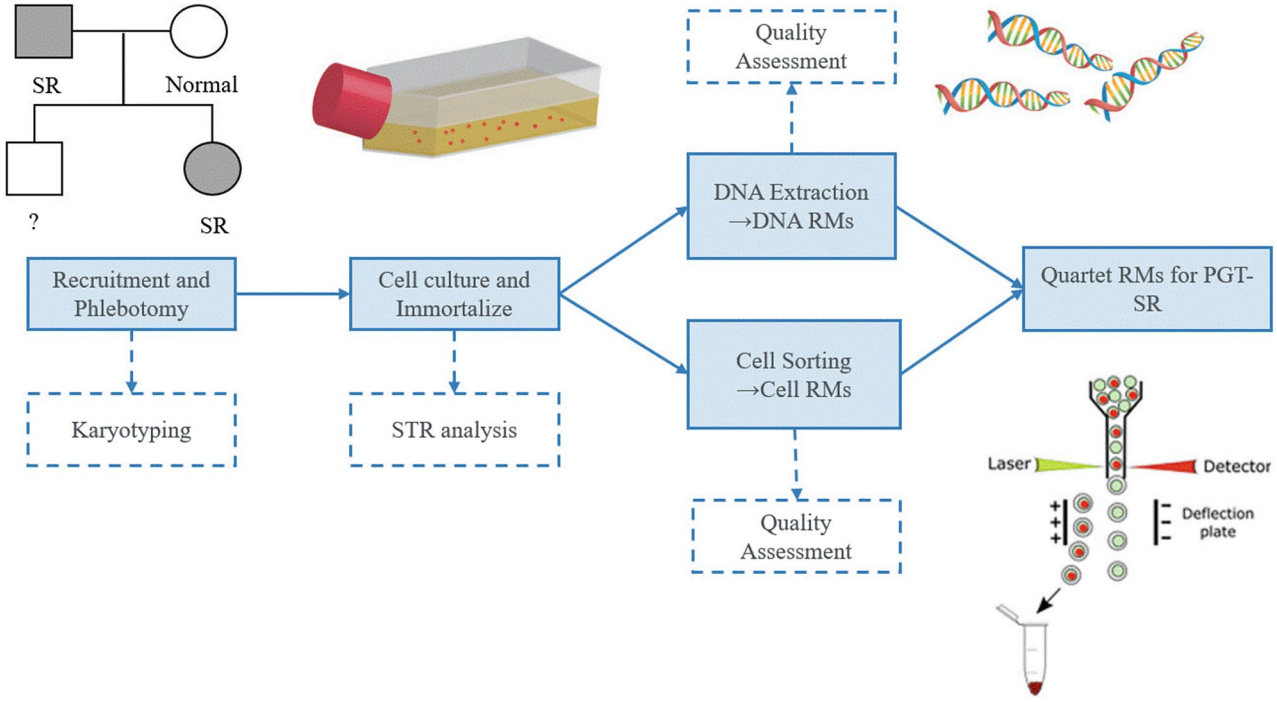

Fig. 1 Schematic flowchart for the preparation of quartet RMs for PGT-SR. The work flow of the generation of quartet RMs in vitro. The quartet families were recruited and phlebotomized, with their blood samples were sent for karyotyping as well as for establishment of immortalized cell lines. Short tandem repeats (STR) analysis was adopted to preclude contamination. All four cell lines in the families were used for the preparation of DNA RMs by DNA extraction, while only one cell line mimicking embryonic cells in the family were used for the preparation of cell RMs by cell sorting. Quality assessment was performed for both DNA RMs and Cell RMs using various methods. RMs, reference materials. STR, short tandem repeats. PGT-SR, preimplantation genetic testing for structural rearrangements

Within each family, one offspring served as reference for haplotype analysis (referred as "reference") while the other was used to simulate an embryo with the investigated karyotype (referred as "embryo"). Then, we isolated and immortalized the lymphocytes from peripheral blood samples obtained from all family members. Genomic DNA (gDNA) was then extracted from immortalized cell lines of parents and "reference" samples to prepare the DNA RMs. Meanwhile, we sorted immortalized cells from the "embryo" samples into 5–10 cells per tube, replicating the clinical scenario of trophectoderm cells used in PGT-SR. Consequently, the quartet RMs from each family comprised 3 tubes of DNA RMs and 1 tube of cell RM, while multiple families harboring different SR constituted a comprehensive panel of RMs for PGT-SR. As illustrated in Fig. 1, a series of quality testing protocols were implemented at each key point, including genotyping and experimental assessments, to ensure the high quality and reliability of the RMs.

Recruitment of quartet family and determination of karyotype

A total of 13 quartet families were recruited with informed consent. The carrier status of SR in each family were determined by karyotyping. Various types of balanced structural abnormalities were covered in this cohort including reciprocal translocations (8 families), Robertsonian translocations (2 families) and inversions (2 families) (Table 1). Within 8 families of reciprocal translocations, 50% (4/8) were inherited from fathers, while the rest (50%, 4/8) were inherited from mothers. All the "reference" were carrier of reciprocal translocations (87.5%, 7/8) or partial trisomy + partial monosomy (12.5%, 1/8). 87.5% (7/8) of the "embryo" were carrier of reciprocal translocations, while the one left was normal. Interestingly, pedigree SR5 contained 2 reciprocal translocations and each was inherited by "reference" or "embryo" separately. Within 2 families of Robertsonian translocations (SR12/SR14), either "reference" or "embryo" was carrier of translocation that both inherited from their fathers. For 2 families of inversion (SR15/SR16), both "reference" and "embryo" were carrier inherited from the father in SR15, while only the "reference" was carrier inherited from the mother in SR16. One normal family (SR19) was included to act as negative control (Table 1). In total, our panel of quartet RMs for PGT-SR covered 3 types of SR and 15 different chromosomes (Supplemental Fig. 1), demonstrating a representative of chromosomal structural abnormalities.

Table 1 The genetic information of recruited families for the construction of quartet RMs for PGT-SR

Pedigree	Type	Karyotype of Father	Karyotype of Mother	Karyotype of Reference	Karyotype of Embryo	Chromosomes
SR1	Reciprocal translocation	46,XY, t(1;11)(q31;q25)	46,XX	46,XY, der(11)t(1;11)(q31;q25); 1q31q44×3,11q25×1	46,XX, t(1;11)(q31;q25)	1;11
SR2	Reciprocal translocation	46,XY	46,XX, t(8;18)(p21;p11.2)	46,XY, t(8;18)(p21;p11.2)	46,XX, t(8;18)(p21;p11.2)	8;18
SR3	Reciprocal translocation	46,XY	46,XX, t(5;7)(p15.1;q32)	46,XX, t(5;7)(p15.1;q32)	46,XX, t(5;7)(p15.1;q32)	5;7
SR4	Reciprocal translocation	46,XY, t(6;11)(p23;p13)	46,XX	46,XX, t(6;11)(p23;p13)	46,XX	6;11
SR5	Reciprocal translocation	46,XY, t(4;21)(p10;q10), t(5;12)(p15.1;q22)	46,XX	46,XY, t(4;21)(p10;q10)	46,XX, t(5;12)(p15.1;q22)	4;5;12;21
SR6	Reciprocal translocation	46,XY	46,XX, t(3;18)(q23;q22)	46,XY, t(3;18)(q23;q22)	46,XX, t(3;18)(q23;q22)	3;18
SR8	Reciprocal translocation	46,XY, t(11;13)(p13;q12)	46,XX	46,XY, t(11;13)(p13;q12)	46,XY, t(11;13)(p13;q12)	11;13
SR9	Reciprocal translocation	46,XY	46,XX, t(16;18)(q22;q21)	46,XX, t(16;18)(q22;q21)	46,XX, t(16;18)(q22;q21)	16;18
SR12	Robertsonian translocations	45,XY, der(13;14)(q10;q10)	46,XX	45,XY, der(13;14)(q10;q10)	46,XX	13;14
SR14	Robertsonian translocations	45,XY, der(14;21)(q10;q10)	46,XX	46,XX	45,XX, der(14;21)(q10;q10)	14;21
SR15	Inversion	46,XY	46,X, inv(X)(p11.23q22.3)	46,X, inv(X)(p11.23q22.3)	46,Y, inv(X)(p11.23q22.3)	X
SR16	Inversion	46,XY, inv(1)(q25.3q32.3)	46,XX	46,XX, inv(1)(q25.3q32.3)	46,XY	1
SR19	Negative	46,XY	46,XX	46,XX	46,XX	NA

NA Not applicable

Preparation and quality control of DNA RMs

After confirming the recruited families' karyotypes, immortalized cell lines were established by isolating lymphocytes from peripheral blood and infecting the cells using Epstein-Barr virus (EBV). The success rate for EBV-mediated immortalization is 94% (49/52) since 3 cell lines failed at the first time and succeed after retaking the blood test. To ensure the integrity of cell culture and prevent potential contamination, short tandem repeats (STR) analysis was employed to confirm the consistent kinship within each family (Supplemental Table 1). For haplotype analysis, gDNA were extracted from cell lines of all members and the concentration were diluted into 15ng/μL in 20μL water per tube, resulting in 39 tubes of DNA RMs in each panel of quartet RMs. Electrophoresis and detection of optical density (OD) were performed to assess the quantity, intactness and purity of the gDNA, confirming the sufficient quantity and satisfactory quality of all DNA RMs (Supplemental Table 2).

Preparation and quality control of cell RMs

To generate cell RMs mimicking embryonic cells, we developed a robust cell sorting method based on unstained flow-cytometry. This approach efficiently separated 5–10 cells into one tube without any adverse effects from staining. The parameter settings were adjusted and confirmed by counting the number of cells on glass slides under the light microscope to ensure the accuracy of cell numbers. After sorting of "embryo" cell lines, three tubes were randomly selected from each cell lines. Whole genome amplification (WGA) was conducted to verify the consistency of cell sorting. The concentrations of WGA (mean: 636 ng/μL, SD: 0.06) were found to be consistent across cell RMs and aligned with those of clinical samples (mean: 658 ng/μL, SD: 0.05, Supplemental Fig. 2), demonstrating the stability and reliability of cell-sorting process.

Characterization of quartet RMs by three laboratories

To characterize our panel of quartet RMs for PGT-SR, we collaborated with three volunteer laboratories in China, each employing distinct techniques and platforms (Supplemental Table 3). In brief, two common methods were utilized in PGT-SR, genome-wide SNP analysis (laboratories 1 and 3 with different platforms) and SNP array (laboratory 2), both based on haplotype analysis. Each laboratories received and tested one set of RMs, including 39 tubes of DNA RMs and 13 tubes of cell RMs. Since normal family (SR19) could not be detected by PGT-SR, these three laboratories actually detected and analyzed 12 pedigrees with 36 tubes of DNA and 12 tubes of cell RMs. As shown in Table 2, the concentration of libraries for all samples met the requirement across all three laboratories. Additionally, although every laboratory employed unique quality control matrixes, our RMs successfully met all criteria for high quality (data not shown). More importantly, the carrier status of SR of all cell RM mimicking embryonic cells were successfully detected in all three laboratories of DNA RMs with 100% accuracy and100% specificity (Table 2).

Determine the breakpoints of quartet RMs

Although the karyotype of all RMs were initially identified, various sequencing methods were employed to precisely determine the exact breakpoints of SRs. For balanced reciprocal translocation, the breakpoints were accurately detected through WGS (Supplemental Table 4). The number of reads spanning the breakpoints varied from 13 to 47 (median 35). Using Sanger sequencing, 87.5% (14/16) of the breakpoints of reciprocal translocation were successfully verified except for two breakpoints of pedigree SR6, which failed the polymerase chain reaction (PCR) process probably due to the poly-A structure around the breakpoints. For inversion in SR15, the

Table 2 The characterization results from three volunteer laboratories that performed the characterization of RMs for PGT-SR

		Laboratory 1		Laboratory 2		Laboratory 3	
Pedigree	Karyotype	Library concentration (ng/μl)	Detection	Library concentration (ng/μl)	Detection	Library concentration (ng/μl)	Detection
SR1	46,XN, t(1;11)(q31.2;q25)	24.2	√	76.5	√	19.5	√
SR2	46,XN, t(8;18)(p21;p11.2)	27.8	√	34.8	√	20	√
SR3	46,XN, t(5;7)(p15.1;q32)	22	√	55.3	√	20.2	√
SR4	46,XN, t(6;11)(p23;p13)	21	√	49.9	√	23.6	√
SR5	46,XN, t(5;12)(p15.1;q22)	18.1	√	48.7	√	23.4	√
SR6	46,XN, t(3;18)(q23;q22)	21.2	√	42.1	√	22.6	√
SR8	46,XN, t(11;13)(p13;q12)	27.8	√	71.1	√	19.2	√
SR9	46,XN, t(16;18)(q22;q21)	28.4	√	78	√	20.6	√
SR12	45,XN, der(13;14)(q10;q10)	23	√	31	√	24.2	√
SR14	45,XN, der(14;21)(q10;q10)	25.8	√	78.9	√	18	√
SR15	46,N, inv(X)(p11.23q22.3)	27.8	√	69.1	√	17.1	√
SR16	46,XN, inv(1)(q25.3q32.3)	24.8	√	61.6	√	18.3	√

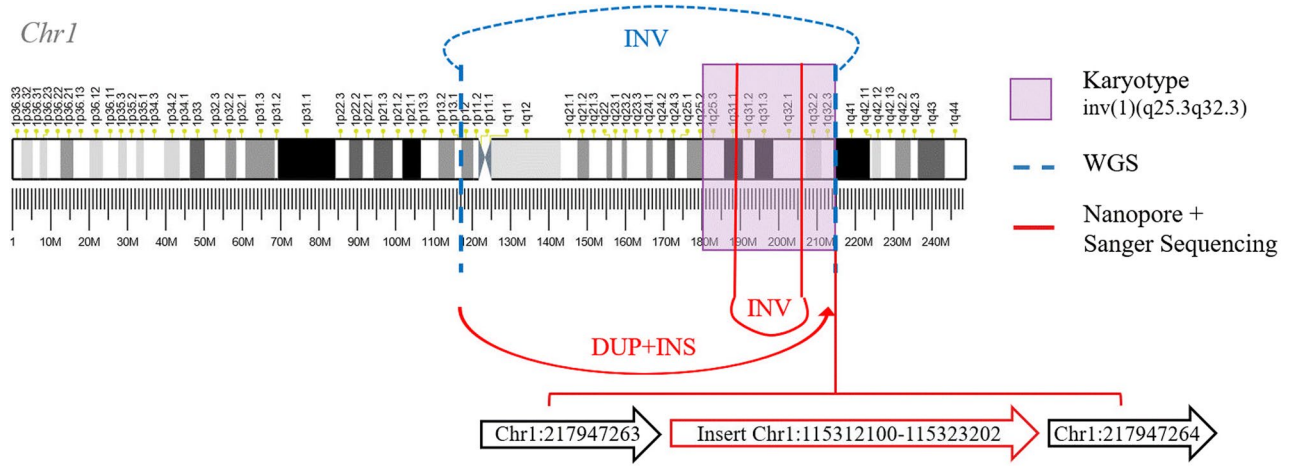

Fig. 2 The information of breakpoints detected in the pedigree SR16 with complex structural recombination. Here depicted a comprehensive picture of cytoband and chromosome size for reference genome of chromosome 1. The abnormalities detected by different methods in pedigree SR16 were labeled on the cytoband of correspondent positions. The karyotype analysis detected an inversion of q25.3 and q32.3 as shown in purple block. WGS, however, detected an inversion of p13.2 and q41, as indicated by blue lines. Nanopore sequencing detected an inversion of q31.1 and q32.1, confirming the result of karyotype. Meanwhile, it detected an insertion on q41 of a segment of micro-duplication on p13.2, explaining why WGS would give the false result. The results of Nanopore sequencing (red lines) were verified by Sanger sequencing

breakpoints were also successfully determined by WGS and were validated by Sanger sequencing.

However, for Robertsonian translocations, neither short-read based genome sequencing nor long-read based genome sequencing via nanopore sequencing could ascertain the breakpoints, primarily due to tandemly repeated sequences around centromeres. Moreover, in pedigree with complex structural rearrangements (SR16) involving combinations of inversion and insertion, WGS failed to accurately identify the breakpoints due to the limitation of short reads. In contrary, nanopore sequencing utilizing long reads successfully identified the exact breakpoints of both insertion and inversion (Fig. 2) as confirmed by Sanger sequencing. Therefore, while long-read based nanopore sequencing demonstrated comparable performance with short-read based WGS in detecting standard reciprocal translocation and inversion, it exhibited superior capability in identifying complicated structural rearrangements.

Discussion

In this study, we designed a methodology and developed a panel of quartet RMs for PGT-SR, consisting of 39 DNA RMs and 13 sorted cell RMs from 13 quartet families, which have 4 obvious advantages as following:

(1) Firstly, our samples were representative and effectively simulated the clinical scenario.13 quartet family recruited from the clinic were used to create cell lines that were not available from the common cell repositories. Compared with existing unrelated cell lines that lack familial context, quartet family-based RMs enabled linkage-based haplotyping by analyzing Mendelian inheritance patterns without introducing unrelated variants. SRs included in this panel of RMs were selected by clinical experts and test developers to make sure that this panel was as representative as possible. For instance, Robertsonian translocations and inversion were both included in two families of our panel despite the fact that they were comparatively rare in the population. The most common type of Robertsonian translocations-der(13;14) -was successfully recruited, accounting for approximately 75% of all Robertsonian translocations [36]. The most common inversion reported was a pericentric inversion of the q and p arm of the chromosome 9, with an incident of 0.25%-3.5% in the population [37]. However, whether this inversion was clinically significant was still under debate, so it was not included in our RM panel of PGT-SR. In total, this panel was a large and comprehensive RMs panel for PGT-SR which could be extended to cover all chromosomes in the future.

(2) Moreover, the methodologies and the quality control procedures we established effectively replicated the inherent characteristics of clinical samples. Clinical embryonic cells were nearly impossible to generate reference materials since they were hardly available due to ethical barriers, extremely unstable due to pluripotency and uneven due to mosaicism. Therefore, immortalized cell lines were constructed and sorted into 5–10 cells per tube to mimic embryonic cell specimens in the clinic. By contrast, synthetic gDNA commonly used for the quality control (QC) of PGT assay were normally used in bulk with over-simplified genomic architecture instead of in single-cell WGA with authentic familial genetic architecture, making them unable to monitor the important process of WGA and haplotyping analysis. Therefore, these quartet RMs proved biological fidelity and biomimetic

advantages when compared with synthetic gDNA and unrelated cell lines. Prior to and following cell immortalization using the Epstein-Barr virus (EBV), we confirmed genotypes, with no discernible impact observed. The process of unstained-cell sorting was developed and well-adjusted to produce cell RMs. Compared with conventional staining methods, unstained-cell sorting avoided any potential harm on downstream analysis. We continuously adjusted the modes and parameters until consistently achieving 5–10 cells per tube. The reliability of cell sorting was affirmed through the WGA concentrations across three randomly-selected tubes, indicating that our methodology not only offered efficient renewal but also boasted high reliability.

(3) More importantly, this panel of RMs were well-characterized with stable and high quality. Most of the RMs developed nowadays and available in the market have never been characterized by various laboratories or platform, bringing all kinds of troubles when used in the clinic or product development. To address these problems, 3 laboratories using SNP array and random MPS were selected to characterized this panel based on their representative technologies and willingness to characterize the samples. The testing results were sent to us along with the results of quality assessment evaluated by different quality matrix, all demonstrating the high quality of this panel of RMs for PGT-SR.

(4) Since the exact breakpoints were clarified, our panel of quartet RMs could not only be used for PGT-SR but also for many other tests of SR. Our findings were concordant to previous studies proving that long-read sequencing was superior than short-read sequencing in terms of complex chromosome rearrangements [38–40], yet our case was novel and rare as a combination of duplication, insertion and inversion in one sample in a quartet family of PGT-SR. This study also gave us a clear picture about why short-read sequencing was inaccurate enough to identify complex SR, thus broadening our knowledge in terms of the limitation of technologies. Although long-read sequencing is not yet cost-effective globally for routine PGT-SR, the development of targeted sequencing and multiplexing, as well as the constant reduction of cost and requirement would gradually improve the cost-effectiveness and feasibility in routing application.

Moreover, the superior accuracy of long-read sequencing on haplotype phasing and SV detection in complex cases would ensure higher diagnostic yield and fewer confirmatory testing, thus reducing the total cost of diagnostic workflow. Our results prove that a comprehensive detection methods using normal PGT-SR for routine screening and long-read sequencing for unsolved cases is the most cost-effective and feasible solution for clinical workflow at present.

Still, our panel of quartet RMs for PGT-SR have some drawbacks that needed to be noticed. On one hand, our immortalized cell lines would never be exactly identical to clinical embryonic cells for PGT-SR. Since the collection of trophectoderm cells from the embryos was practically and ethically challenging, we used immortalized lymphocytes instead in order to produce stable, standardized and renewable RMs. The process of immortalization using EB virus and longstanding cell culture might trigger some minor genomic alterations (duplication or deletion) on the chromosomes. Since they didn't affect the flanking region near breakpoints, these minor genomic alterations were unnoticed and negligible for haplotyping-based PGT-SR in our study. However, they might be discriminated by new methods like long-read whole genome sequencing. It is critical to perform confirmation test for each batch of RMs and comparation with pre-immortalized results to filter out artifacts and non-inherited alterations during the quality control process. In conclusion, RMs made from immortalized lymphocytes were necessary compromise and the best solution so far for the standardization of PGT-SR, until embryonic cells become more accessible.

On the other hand, our panel of RMs could never cover all the clinical scenario. For instance, many embryos were actually unbalanced in the clinic. Sine the pregnancy of unbalanced embryos would terminate spontaneously, the incidence of unbalanced SR in the population was extremely low. Therefore, we could not recruit lots of unbalanced SR in our panel. Mosaicism, complex and *de novo* SRs were also not recruited in this panel due to the low frequencies in the population. Since they could not be detected by haplotype-based PGT-SR, the clinical application of this panel of PGT-SR would not be affected for most of the detection platform. Mosaicism, complex and de novo SRs would be recruited or engineered when new methods emerge or the problem of allele dropout is solved.

In the future, we would adapt this novel methodology to produce more quartet RMs for PGT-M and comprehensive PGT workflows, which use linkage-based haplotyping for analysis. These expanding quartet RMs for PGT-M better include some single-gene diseases with relatively high frequency, like Duchenne muscular dystrophy (DMD) and spinal muscular atrophy (SMA). Actually, we have recruited several quartet families with SMA and generated a panel of RMs for PGT-M of SMA. As for PGT-A, quartet families are not necessary since they directly detect chromosomal abnormalities without haplotyping, which make unrelated cell lines work well as RMs. Still, our workflow of internal quality control and external characteristic could be integrated in the generation of RMs for PGT-A.

Conclusions

In summary, we have developed a comprehensive strategy and successfully generated a panel of quartet RMs tailored for PGT-SR. Since publicly available in large quantities, this well-characterized panel serves as standardized quality control for diverse applications, including test development, performance validation, quality control or external quality assessment for PGT-SR as well as other tests for SR. More importantly, the methodology designed in the project lays the groundwork for producing additional RMs for PGT-M or novel techniques capable of simultaneously detecting PGT-A, PGT-M and PGT-SR. Overall, this panel of RMs for PGT-SR represents a significant stride towards the rapid development and clinical introduction of PGT-SR.

Abbreviations

PGT-SR	Preimplantation genetic testing for structural rearrangements
RMs	Reference materials
SR	Structural rearrangement
PGT-A	Preimplantation genetic testing for aneuploidy
PGT-M	Preimplantation genetic testing for monogenic disease
FISH	Fluorescence in situ hybridization
SNP	Single nucleotide polymorphism
CGH	Comparative genomic hybridization
MPS	Massive parallel sequencing
gDNA	Genomic DNA
WGA	Whole genome amplification
EBVs	Epstein-Barr viruses
OD	Optical density
WGS	Whole genome sequencing
STR	Short tandem repeats
PCR	Polymerase chain reaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12556-7.

Supplementary Material 1.

Acknowledgements

We appreciate all the participants enrolled who provided peripheral blood. We also acknowledge BGI Research and Clin Lab, BGI Genomics for providing the service of cell culture as well as the equipment for cell sorting. This work is supported by grants from the National Key Research and Development Program of China (2021YFC2700500 and 2022YFF1201903).

Authors' contributions

S. Qu, J. Xiang, N. Sun, W. Zhang, M. Zhang, T. Yu and Z. Peng designed and supervised the study. D. Yu wrote the manuscript. S. Qu and J. Xiang reviewed the manuscript. N. Song enrolled the participants. X. Dong, Y. Guo and D. Yu conducted the experiments and data analysis. C. Jiang and W. Wang performed the bioinformatic analysis. Y. Wang, L. Kong, B. Liang, J. Fei, S. Lu, Y. Zou, X. Dong, D. Yu, M. Zhang, T. Yu, J. Xiang and S. Qu accomplished and analyzed the characterization of RMs.

Funding

This study received funding from the National Key R&D Program of China (2021YFC2700500 and 2022YFF1201903).

Data availability

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA009786) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human.

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of Beijing Genomics Institute (BGI, No. BGI-IRB 22010) and the Ethics Committee of the First People's Hospital of Chenzhou (No. 2022003 H), China. All participants gave informed consent.

Consent for publication

Not applicable.

Competing interests

D. Yu, X. Dong, N. Song, Y. Guo, C. Jiang, W. Wang, Z. Peng and J. Xiang are employed by the BGI Genomics. L. Kong and B. Liang are employed by the Basecare Medical Device Co., Ltd. J. Fei are employed by the Peking Jabrehoo Med Tech Co., Ltd. S. Lu are employed by the Yikon Genomics Company, Ltd. The remaining authors declare no competing interests.

Received: 18 September 2025 / Accepted: 13 January 2026
Published online: 10 February 2026

References

1. Viotti M. Preimplantation genetic testing for chromosomal abnormalities: aneuploidy, mosaicism, and structural rearrangements. Genes (Basel). 2020;11(6):602.
2. Jacobs PA, Browne C, Gregson N, Joyce C, White H. Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. J Med Genet. 1992;29(2):103.
3. Braekeleer MD, Dao TN. Cytogenetic studies in couples experiencing repeated pregnancy losses. Hum Reprod. 1990;5(5):519–28.
4. De Braekeleer M, Dao TN. Cytogenetic studies in male infertility. Hum Reprod. 1991;6(2):245–50.
5. Pettenati M, Rao P, Phelan M, Grass F, Rao K, Cosper P, Carroll A, Elder F, Smith J, Higgins M. Paracentric inversions in humans: a review of 446 paracentric inversions with presentation of 120 new cases. Am J Med Genet. 1995;55(2):171–87.
6. Van Dyke DL, Weiss L, Roberson JR, Babu VR. The frequency and mutation rate of balanced autosomal rearrangements in man estimated from prenatal genetic studies for advanced maternal age. Am J Hum Genet. 1983;35(2):301–8.
7. Vozdová M, Kasikova K, Oracova E, Prinosilova P, Rybar R, Horinova V, Gaillyová R, Rubes J. The effect of the swim-up and hyaluronan-binding methods on the frequency of abnormal spermatozoa detected by FISH and SCSA in carriers of balanced chromosomal translocations. Hum Reprod. 2012;27(3):930–7.
8. Ishikawa T, Shiotani M, Izumi Y, Hashimoto H, Kokeguchi S, Goto S, Fujisawa M. Fertilization and pregnancy using cryopreserved testicular sperm for intracytoplasmic sperm injection with azoospermia. Fertil Steril. 2009;92(1):174–9.
9. Morin SJ, Eccles J, Iturriaga A, Zimmerman RS. Translocations, inversions and other chromosome rearrangements. Fertil Steril. 2017;107(1):19–26.
10. Alfarawati S, Fragouli E, Colls P, Wells D. Embryos of Robertsonian translocation carriers exhibit a mitotic interchromosomal effect that enhances genetic instability during early development. PLoS Genet. 2012;8(10):e1003025.
11. Shetty S, Nair J, Johnson J, Shetty N, Kumar A, Thondehalmath N, Ganesh D, Bhat VR, Sajana M, Anjana R. Preimplantation genetic testing for couples with balanced chromosomal rearrangements. J Reprod Infertil. 2022;23(3):213.
12. Dang T, Xie P, Zhang Z, Hu L, Tang Y, Tan Y, Luo K, Gong F, Lu G, Lin G. The effect of carrier characteristics and female age on preimplantation genetic testing results of blastocysts from Robertsonian translocation carriers. J Assist Reprod Genet. 2023;40(8) :1995-2002.
13. Tong J, Niu Y, Wan A, Zhang T. Effect of parental origin and predictors for obtaining a euploid embryo in balanced translocation carriers. Reprod Biomed Online. 2022;44(1):72–9.
14. Clementini E, Palka C, Iezzi I, Stuppia L, Guanciali-Franchi P, Tiboni G. Prevalence of chromosomal abnormalities in 2078 infertile couples referred for assisted reproductive techniques. Hum Reprod. 2005;20(2):437–42.

15. Sugiura-Ogasawara M, Ozaki Y, Sato T, Suzumori N, Suzumori K. Poor prognosis of recurrent aborters with either maternal or paternal reciprocal translocations. Fertil Steril. 2004;81(2):367–73.
16. Campana M, Serra A, Neri G, Reynolds J. Role of chromosome aberrations in recurrent abortion: a study of 269 balanced translocations. Am J Med Genet. 1986;24(2):341–56.
17. Dong Z, Yan J, Xu F, Yuan J, Jiang H, Wang H, Chen H, Zhang L, Ye L, Xu J. Genome sequencing explores complexity of chromosomal abnormalities in recurrent miscarriage. Am J Hum Genet. 2019;105(6):1102–11.
18. Hook E, Schreinemachers D, Willey A, Cross P. Inherited structural cytogenetic abnormalities detected incidentally in fetuses diagnosed prenatally: frequency, parental-age associations, sex-ratio trends, and comparisons with rates of mutants. Am J Hum Genet. 1984;36(2):422.
19. Huang C, Jiang W, Zhu Y, Li H, Lu J, Yan J, Chen Z. Pregnancy outcomes of reciprocal translocation carriers with two or more unfavorable pregnancy histories: before and after preimplantation genetic testing. J Assist Reprod Genet. 2019;36:2325–31.
20. Griffin DK, Ogur C. PGT-SR: A comprehensive overview and a requiem for the interchromosomal effect. DNA. 2023;3(1):41–64.
21. Ferfouri F, Bernicot I, Schneider A, Haquet E, Hédon B, Anahory T. Is the resulting phenotype of an embryo with balanced X-autosome translocation, obtained by means of preimplantation genetic diagnosis, linked to the X inactivation pattern? Fertil Steril. 2016;105(4):1035–46.
22. Richardson H. Medical laboratories—requirements for quality and competence: an ISO perspective. Vox Sang. 2002;83:333–5.
23. Chen B, O'Connell CD, Boone DJ, Amos JA, Beck JC, Chan MM, Farkas DH, Lebo RV, Richards CS, Roa B. Developing a sustainable process to provide quality control materials for genetic testing. Genet Med. 2005;7(8):534–49.
24. Chen B, Gagnon M, Shahangian S, Anderson NL, Howerton DA, Boone D. Good laboratory practices for molecular genetic testing for heritable diseases and conditions. MMWR Recomm Rep. 2009;58(6):1–CE.
25. Barker SD, Bale S, Booker J, Buller A, Das S, Friedman K, Godwin AK, Grody WW, Highsmith E, Kant J. Development and characterization of reference materials for MTHFR, SERPINA1, RET, BRCA1, and BRCA2 genetic testing. J Mol Diagn. 2009;11(6):553–61.
26. Bettinotti MP, Ferriola D, Duke JL, Mosbruger TL, Tairis N, Jennings L, Kalman LV, Monos D. Characterization of 108 genomic DNA reference materials for 11 human leukocyte antigen loci: a GeT-RM collaborative project. J Mol Diagn. 2018;20(5):703–15.
27. Pratt VM, Caggana M, Bridges C, Buller AM, DiAntonio L, Highsmith WE, Holtegaard LM, Muralidharan K, Rohlfs EM, Tarleton J. Development of genomic reference materials for cystic fibrosis genetic testing. J Mol Diagn. 2009;11(3):186–93.
28. Kalman LV, Tarleton JC, Percy AK, Aradhya S, Bale S, Barker SD, Bayrak-Toydemir P, Bridges C, Buller-Burckle AM, Das S. Development of a genomic DNA reference material panel for Rett syndrome (MECP2-related disorders) genetic testing. J Mol Diagn. 2014;16(2):273–9.
29. Kalman L, Johnson MA, Beck J, Berry-Kravis E, Buller A, Casey B, Feldman GL, Handsfield J, Jakupciak JP, Maragh S. Development of genomic reference materials for huntington disease genetic testing. Genet Med. 2007;9(10):719–23.
30. Ling P, Huls H. Isolation and immortalization of lymphocytes. Curr Protoc Mol Biol. 2005;70(1):28. 22.21–28.22.25.
31. Zhang S, Zhao D, Zhang J, Mao Y, Kong L, Zhang Y, Liang B, Sun X, Xu C. Base-Phasing: a highly efficient approach for preimplantation genetic haplotyping in clinical application of balanced translocation carriers. BMC Med Genom. 2019;12(1):52.
32. Hu L, Cheng D, Gong F, Lu C, Tan Y, Luo K, Wu X, He W, Xie P, Feng T. Reciprocal translocation carrier diagnosis in preimplantation human embryos. EBioMedicine. 2016;14:139–47.
33. Ye Y, Ma J, Cui L, Lu S, Jin F. A rapid NGS-based preimplantation genetic testing for chromosomal abnormalities in day-3 blastomere biopsy allows embryo transfer within the same treatment cycle. Front Genet. 2021;12:636370.
34. Ren J, Keqie Y, Li Y, Li L, Luo M, Gao M, Peng C, Chen H, Hu T, Chen X. Case report: optical genome mapping revealed double rearrangements in a male undergoing preimplantation genetic testing. Front Genet. 2023;14:1132404.
35. Wang Y, Jia L, Tian G, Dong Y, Zhang X, Zhou Z, Luo X, Li Y, Yao W. shinyCircos-V2. 0: leveraging the creation of circos plot with enhanced usability and advanced features. Imeta. 2023;2(2):e109.
36. Du R, Zhao D, Kang K, Wang F, Xu R, Chi C, Kong L, Liang B. A review of preimplantation genetic testing technologies and applications. Reproductive Dev Med. 2023;7(01):20–31.
37. Abdi A, Bagherizadeh I, Bahman I, Shafaghti Y, Behjati F. Prevalence of chromosome inversions (pericentric and paracentric) in patients with recurrent abortions. Sarem J Reproductive Med. 2018;3(1):45–50.
38. Thibodeau ML, O'Neill K, Dixon K, Reisle C, Mungall KL, Krzywinski M, Shen Y, Lim HJ, Cheng D, Tse K, et al. Improved structural variant interpretation for hereditary cancer susceptibility using long-read sequencing. Genet Med. 2020;22(11):1892–7.
39. Damián A, Núñez-Moreno G, Jubin C, Tamayo A, de Alba MR, Villaverde C, Fund C, Delépine M, Leduc A, Deleuze JF, et al. Long-read genome sequencing identifies cryptic structural variants in congenital aniridia cases. Hum Genomics. 2023;17(1):45.
40. Mariya T, Shichiri Y, Sugimoto T, Kawamura R, Miyai S, Inagaki H, Sugihara E, Ikeda K, Baba T, Ishikawa A, et al. Clinical application of long-read nanopore sequencing in a preimplantation genetic testing pre-clinical workup to identify the junction for complex Xq chromosome rearrangement-related disease. Prenat Diagn. 2023;43(3):304–13.

Publisher's note