Abstract. Fetal chromosomal abnormalities are a common cause of spontaneous abortion. The present study investigated the association between fetal chromosomal abnormalities and the frequency of spontaneous abortions to enable clinicians to provide more informed genetic counseling. A total of 182 patients with a history of spontaneous abortions were recruited from July 2015 to August 2017. G-banding cytogenetic analysis and novel high-throughput ligation-dependent probe amplification (HLPA) techniques were performed on conception in all 182 patients to detect chromosomal abnormalities. Low-coverage whole-genome sequencing (WGS) was performed in 74 patients to detect copy number variations (CNVs).

There were no significant differences in the incidence of karyotype abnormalities between patients with sporadic miscarriages (48.0%; SM group) and patients suffering recurrent spontaneous abortions (44.8%; RSA group). The maternal age was markedly higher in patients with 3 miscarriages. WGS indicated that the incidence of pathogenic CNVs in the RSA group was higher than that in the SM group, but the difference was not significant. In conclusion, a high incidence of karyotype abnormalities and pathogenic CNVs was observed in patients with spontaneous abortion. However, no association between fetal chromosomal abnormalities and the number of spontaneous abortions was observed. HLPA assays may be used as an alternative method for fetal karyotype analysis and determination of CNVs in patients with SM and RSA.

Introduction

Spontaneous abortion is a common complication during human pregnancies that occurs in 10-15% of clinically recognized pregnancies (1). Recurrent spontaneous abortion (RSA), defined as two or more constitutive miscarriages, the pathogenesis of which is complex, occurs in 5% of clinically recognized pregnancies (2). Understanding the etiology of miscarriages may improve the risk evaluation in RSA and may enable clinicians to provide more informed genetic counseling. Fetal chromosomal abnormalities are considered a major genetic cause of spontaneous abortions, accounting for 50-60% of spontaneous pregnancy losses. However, the etiology of fetal chromosome abnormalities remains undetermined (3,4). Furthermore, differences between the incidence of sporadic miscarriages (SMs) and RASAs for couples carrying normal karyotypes are complex. A limited number of studies have focused on the association between the number of spontaneous abortions and the prevalence of pathogenic copy number variations (CNVs) in aborted fetuses (5). High-throughput ligation-dependent probe amplification (HLPA) is a novel molecular diagnostic technique for aneuploidy detection that provides a rapid, high-throughput and cost-effective alternative to multiplex ligation-dependent probe amplification (MLPA) methods for the detection of copy numbers in all 24 chromosomes. HLPA may also detect duplication/deletions of genomic regions >10 Mb in a single PCR (6-8). In the present study, a comprehensive genetic analysis of fetal villus tissues of patients with spontaneous abortion was performed using a combination of G-banding karyotyping, HLPA and low-coverage whole-genome sequencing (WGS) (9). Diagnosing the genetic causes of spontaneous abortions is of great importance and allows clinicians to evaluate recurrence risks in pregnancies, which
has a key role in family planning prior to and after subsequent conceptions.

Materials and methods

Study participants. A total of 182 patients who experienced spontaneous abortions were recruited between July 2015 and August 2017 at Wuxi Maternal and Child Health Hospital. The average age of recruited patients was 28.46 years and ranged from 23 to 39 years old. Patients suffering from systemic infections and injury during pregnancy were excluded. Demographic and clinical data were collected from the medical records, including obstetric/gynecological history, as well as maternal age and gestational age at the time of the spontaneous abortion.

Chorionic villus sample preparation. Chorionic villus sampling was performed on the products of conception from all patients according to standard clinical procedures. Conventional G-banding cytogenetic analysis and HLPA detection were then performed. Low-coverage WGS was performed in the chorionic villus samples of low quantity to ensure successful detection rates. G-banded karyotyping assays were performed to analyze chromosomal abnormalities according to the standard protocols of the International System for Human Cytogenetic Nomenclature (ISCN) (10). Fig. S1 provides an overview of this procedure and sample sizes are included.

Cytogenetic analysis. Qualified chorionic villus tissues were collected from the products of conception according to standard procedures. Chromosomes were G-banded and karyotypes were analyzed using the ISCN from 2013 (10).

HLPA analysis. Genomic DNA was extracted from uncultured chorionic villus tissues using QIAamp DNA Mini Kits (Qiagen). DNA concentrations were measured on a Nandrop -1000 (Thermo Fisher Scientific, Inc.). HLPA is a modified version of MLPA designed to measure genomic CNVs in a single multiplex PCR. The copy numbers for the 24 chromosomes were determined using HLPA kits (cat. no. N1002; Genesky). A total of 170 pairs of probes targeting all 24 chromosomes were designed. A total of 3 probes were placed at the end of each chromosome (excluding chromosomes 13, 14, 15, 21 and 22), with nearly 10 Mb of genomic distance between each probe. A single probe at each side of the centromere was also designed. Genomic DNA (100-200 ng) was denatured at 98˚C for 6 min in a 10-µl reaction containing 1X TE and mixed with 10 µl ligation premix containing 2 µl of 10X Taq ligase buffer, 0.5 µl of Taq ligase, 1 µl of 20X Probe Mix (Thermo Fisher Scientific, Inc.) and 7.5 µl Milli-Q water. Ligation reactions were performed on an ABI 2720 thermal cycler (Thermo Fisher Scientific, Inc.) using the following program: A total of 4 cycles of 94˚C for 1 min, 60˚C for 4 h and 94˚C for 2 min, followed by a hold at 72˚C for 2 min until 20 µl of 2X Stop Buffer was added. PCR mixtures (total volume, 20 µl) were prepared to amplify the ligation products, containing 1X Taq Buffer, 0.3 mM of dNTP, 1X Fluorescence Primer Mix, 0.8 U of Hotstart Taq DNA polymerase and 1 µl of the ligation reaction mixture. The PCR program was as follows: Initial denaturation at 95˚C for 2 min; 32 cycles of 94˚C for 20 sec, 57˚C for 40 sec, 72˚C for 90 sec and 68˚C for 60 min, followed by a hold at 4˚C. PCR products were diluted 20-fold prior to loading on an ABI 3730XL sequencer (Thermo Fisher Scientific, Inc.). Raw data were analyzed using GENEMAPPER 4.0 (Thermo Fisher Scientific, Inc.) according to standard analytic procedures. Short tandem repeat (STR) profiling was applied to identify whether chromosomal abnormalities of the samples were genetic mosaics or contaminated with maternal DNA. The cut-off values of significant maternal DNA contamination percentages were ≥30%. Samples were amplified using the Human Personal Identification Detection kit (R1006; Genesky). A total of 17 markers were amplified in a single multiplex reaction and products were analyzed on an ABI 3130XL sequencer. STR alleles were analyzed using GENEMAPPER 4.0 software (Thermo Fisher Scientific, Inc.).

Low-coverage WGS. Genomic DNA was extracted from test samples using a QIAamp DNA Mini kit (Qiagen). Genomic DNA (~3 µg) was sheared to construct the library with insert sizes of 3-8 kb. After circularization and ligation, circular DNA fragments were again sheared to generate fragments of ~500 bp. The fragments were end-repaired and A-tailed in preparation for ligation to Illumina paired-end oligonucleotide adapters. PCRs were used to assess the DNA fragments with adapter molecules at both ends. PCR products of ~626 bp were collected. Constructed libraries were subjected to 50 bp-end multiplex sequencing on the Illumina HiSeq™ 2000 platform (Illumina, Inc.). Qualified paired-end reads were aligned to the NCBI human assembly hg19/GRCh37.1 using SOAP2 and only unique reads were used for the follow-up copy number analysis. To improve detection sensitivity, a 60-kb window with 5-kb sliding was used to allocate and analyze ~5 million sequencing reads in overlapping 60-kb bins. For each chromosome, the mean log2 values of the normalized sequencing reads (y-axis) vs. the number of sequential 5-kb sliding 60-kb sequencing bins (x-axis) were plotted. The mean log2 values were then calculated along the length of each chromosome. The theoretical log2 value for duplications (three copies) is log2(3)=1.58 and that for a deletion (one copy) is log2(0.5)=1.0. Cut-off copy number values >2.8 [log2(1.4)=0.49] were used to detect duplications and values <1.2 [log2(0.6)=−0.74] were used to identify deletions.

Pathogenic CNVs were confirmed by aligning detected CNVs with genetic databases, including Decipher (http://decipher.sanger.ac.uk/syndromes; version 9.31), Database of Genomic Variants (DGV; http://dgv.tcag.ca), Online Mendelian Inheritance in Man (OMIM; http://www.omim.org) and University of California Santa Cruz Genome Browser (UCSC Genome Browser; http://genome.ucsc.edu; Human GRCh37/hg19).

Sanger sequencing validation. Blood samples of selected couples were collected and DNA was extracted. Low-coverage WGS results were analyzed and a total of 170 probes were designed to validate the genetics of micro-deletions and micro-duplications. Genomic DNA of 100-200 ng was denatured at 98˚C for 2 min in a 20-µl reaction containing 1.25 µl of 4X DNA lysis buffer, 2 µl of 10X Taq ligase buffer, 0.2 µl of Taq ligase, 1 µl of 20X probe mix and 10.55 µl of Mili-Q water. Ligation were performed on an ABI 2720 thermal cycler (Thermo Fisher Scientific, Inc.) using the following program: A total of 25 cycles of 94˚C for 20 sec, 57˚C for 40 sec, 72˚C for 90 sec and 60 sec, followed by a hold at 4˚C. PCR products were diluted 20-fold prior to loading on an ABI 3730XL sequencer (Thermo Fisher Scientific, Inc.). Raw data were analyzed using GENEMAPPER 4.0 (Thermo Fisher Scientific, Inc.) according to standard analytic procedures. Short tandem repeat (STR) profiling was applied to identify whether chromosomal abnormalities of the samples were genetic mosaics or contaminated with maternal DNA. The cut-off values of significant maternal DNA contamination percentages were ≥30%. Samples were amplified using the Human Personal Identification Detection kit (R1006; Genesky). A total of 17 markers were amplified in a single multiplex reaction and products were analyzed on an ABI 3130XL sequencer. STR alleles were analyzed using GENEMAPPER 4.0 software (Thermo Fisher Scientific, Inc.).
thermal cycler (Thermo Fisher Scientific, Inc.) using the following parameters: A total of 5 cycles of 95°C for 30 sec; 60°C for 3 h; 94°C for 2 min; hold at 72°C for 2 min until the addition of 20 µl of 2X EDTA. PCRs were performed to amplify the ligation products with the 20-µl reaction mixture containing 10 µl of 2Z PCR Master Mix (Thermo Fisher Scientific, Inc.), 1 µl of Primer Mix (Thermo Fisher Scientific, Inc.), 1 µl of the ligation product and 8 µl Milli-Q water. The PCR parameters were as follows: 95°C for 2 min; 5 cycles of 94°C for 20 sec, 62-1°C/cycle for 40 sec and 72°C for 1 min; 27 cycles of 94°C for 20 sec, 57°C for 40 sec, 72°C for 1 min and 68°C for 20 min, and a final hold at 4°C. PCR products were diluted 5-fold prior to loading on the ABI 3730XL sequencer (Thermo Fisher Scientific, Inc.). Raw data were analyzed using GENEMAPPER 4.0 (Thermo Fisher Scientific, Inc.).

**Statistical analysis.** All statistical analyses were performed using the SPSS software package (version 13.0; SPSS Inc.). Values are expressed as the mean ± standard deviation or n (%). χ² tests were used to compare the positive ratios between groups and Student's t-tests were used to examine the mean age. P≤0.05 was considered to indicate a statistically significant difference.

**Results**

**Karyotyping and HLPA analysis.** A total of 50 patients with SMs with single spontaneous pregnancy loss were assigned to the SM group (group A), whilst 132 patients who had RSAs with >1 pregnancy loss were assigned to the RSA group. Amongst all RSAs, 87 patients with two spontaneous pregnancy losses were assigned to group B, 34 patients who had three spontaneous pregnancy losses were assigned to group C and 11 patients who had four or more pregnancy losses were assigned to group D.

Samples were successfully analyzed for 137/182 patients (75.3%). Failed samples were generated from the cytogenetic analysis process due to inevitable failure of cell culture (7). STR profiling revealed no indication of significant maternal cell contamination. Samples were analyzed using G-banded karyotyping and CNVplex assays. Chromosomal abnormalities were detected in 63 cases, including 24 cases from the SM group (48.0%, 24/50) and 39 cases from the RSA group (44.8%, 39/87). However, no significant differences between the SM group and RSA group were observed (P>0.05; Table I).

Using the G-banded karyotyping and CNVplex assays, a total of 63 CNVs were identified in the SM and RSA groups. Trisomy was the most frequent chromosomal abnormality in the two groups (SM, n=15; RSA, n=29), followed by 45,X syndrome (SM, n=6; RSA, n=7). Other CNVs were detected, including triploidy and tetraploidy. Furthermore, a tetrasomy 48,XXY,+13 was observed (Table II). However, there were no significant differences between the SM and RSA groups in terms of chromosomal abnormalities collectively.

The average age of the patients in the SM and RSA groups was 29.67±5.16 and 29.85±4.62 years, respectively, with no significant differences between the groups (P>0.05). Groups A-D were further sub-divided according to the frequency of spontaneous pregnancy loss (Table III). The average age of patients in groups A-D was 29.67±5.16, 29.39±4.59, 29.56±4.27 and 34.70±3.33 years, respectively. However, there were no significant differences between groups A, B and D, whilst group C had a significant age difference between normal and abnormal karyotype subgroups. Group D had the highest average age. The four groups exhibited no significant differences in chromosomal abnormalities (P>0.05).

**Low-coverage WGS analysis.** Low-coverage WGS was applied to 74 cases with normal karyotypes to further explore whether unknown causes led to spontaneous abortions. A total of 21 patients with CNVs were identified. The distribution of the cases is provided in Table IV; no significant differences between the two groups were identified (P>0.05). To verify whether these CNVs were responsible for spontaneous abortions, specific primers of CNVplex assays were designed according to the detected CNVs, which were further subjected to Sanger sequencing validation to determine whether the CNVs were de novo or genetically inherited from the parents. A total of 10 patients with CNVs were selected and the tissues of the fetuses and blood samples obtained from the parents were jointly analyzed. The results revealed that all CNVs, excluding those for which primers could not be designed, were inherited from the parents (Table SI).

**Discussion**

Pregnancy losses affect 15-25% of couples trying to conceive. The etiologies are complex and include uterine anomalies,
immune and endocrine disturbances, maternal thrombophilic disorders and parental cytogenetic aberrations (11). Previous studies have indicated that fetal chromosomal abnormalities are responsible for the genetic causes of spontaneous abortions, presenting in 50-70% of cases (12,13). Karyotyping and fluorescence in situ hybridization (FISH) are classic diagnostic techniques to detect chromosomal abnormalities in abortion tissue. However, the utility of karyotyping is hindered by its high failure rates in cell culture (14). FISH may be used for rapid diagnosis, but the probes are limited and cannot guarantee full mapping of the chromosomal status. Furthermore, karyotyping and FISH are unable to exclude maternal DNA contamination (15,16). CNVplex is a high-throughput ligation-dependent probe amplification method that provides a rapid and cost-effective method for initial screening for aneuploidy, as described previously (17). In the present study, the percentages of chromosomal abnormalities in the SM group were higher than those in the RSA group, while the occurrence trisomy exhibited the opposite effects, as presented in Table II. Trisomy 16, 18 and 22 and 45,X syndrome were the most common abnormalities in the SM and RSA groups.

In the present study, the average age of patients in group D was ≥4 years higher than that in the other groups, but the differences were not significant. Furthermore, the differences in average age of patients in groups A-C were not significant, indicating that the frequencies of spontaneous abortion were not relevant to the age of these patients. Furthermore, the difference in average age of the patients in the normal chromosome and abnormality groups were not statistically significant, indicating that age was not the major cause of chromosomal abnormalities. However, conflicting results have been reported. Nikitina et al (20) reported that the frequency of abortions with normal karyotypes was significantly higher in maternal patients with recurrent miscarriages as compared to that in patients with SM (53.1 vs. 43.3%; P<0.01), in agreement with the studies by Ogasawara et al (21) (23.7 vs. 48.7%; P<0.01) and Sullivan et al (22) (57.9 vs. 74.6%; P<0.01), whilst Stern et al (23), Marquard et al (24) and Grande et al (25) reported conflicting results. The sample size may be an influencing factor that may have led to the differences observed between the studies, in addition to the criteria for RSAs. In the present study, SM was defined as loss of the first pregnancy, whilst Nikitina et al (20) defined SM as the first miscarriages that occurred in maternal patients with prior normal pregnancies (a live birth or elective termination) Grande et al (25) and Marquard et al (24) defined RSA as three or more miscarriages. The authors considered that a more definite and standardized classification should be used to distinguish SM and RSA patients. G-banded karyotypes were previously used to detect chromosomal anomalies and indicated no advantage for the detection of balance translocation, with more recent techniques including HLPA, next-generation sequencing (NGS) and chromosomal microarray analysis (CMA) demonstrating clear advantages, including low test failure rates and faster detection efficiency. The techniques were compared and it was indicated that HLPA was the most efficient and cost-effective method to detect miscarriages during conception.

The clinical application of karyotyping is hindered by technical limitations. This method only reveals chromosomal abnormalities with a length of ≥8-10 Mb (26-28). CMA has been rapidly applied to detect micro-duplications and deletions due to its high resolution. Low-coverage WGS is a burgeoning technique based on NGS as it frequently used for the diagnosis of genetic diseases (29-32). WGS is a high-throughput method with improved accuracy and lower cost compared to CMA. In the present study, low-coverage WGS was applied

### Table III. Age of maternal females with different numbers of spontaneous abortions (years).

<table>
<thead>
<tr>
<th>Number of spontaneous abortions</th>
<th>Normal karyotype (n=74)</th>
<th>Abnormal karyotype (n=63)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.67±5.16</td>
<td>30.08±5.96</td>
<td>0.55</td>
</tr>
<tr>
<td>2</td>
<td>29.39±4.59</td>
<td>29.44±4.98</td>
<td>0.93</td>
</tr>
<tr>
<td>3</td>
<td>29.56±4.27</td>
<td>28.30±4.98</td>
<td>0.03</td>
</tr>
<tr>
<td>≥4</td>
<td>34.70±3.33</td>
<td>34.50±3.74</td>
<td>0.73</td>
</tr>
</tbody>
</table>

### Table IV. Distribution of CNVs detected in patients with different numbers of spontaneous abortions.

<table>
<thead>
<tr>
<th>Number of spontaneous abortions</th>
<th>CNVs, n/total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6/26 (23.08)</td>
</tr>
<tr>
<td>2</td>
<td>8/23 (34.78)</td>
</tr>
<tr>
<td>3</td>
<td>5/17 (29.41)</td>
</tr>
<tr>
<td>≥4</td>
<td>2/8 (25.00)</td>
</tr>
</tbody>
</table>

CNV, copy number variation.
to further detect samples with negative results derived from karyotyping and CNVplex assays. An additional 21 samples were determined to have CNV variations. The CNV variation rates indicated no significant differences amongst the different frequencies of spontaneous abortion. The occurrence of pathogenic CNVs in RSA was higher than that in the SM group, though the differences were not statistically significant. A total of 19 variants of uncertain significance (VOUS) were detected as positive, which was significantly higher than the 2% reported by Dhillon et al (33) using CMA. This indicates that WGS is able to detect more CNVs than CMA. Sanger sequencing was further applied to 10 cases of VOUS to verify whether they were inherited from the parents or de novo. A mere 10% of CNVs were de novo, and a total of 8 VOUS were verified as being inherited from the parents. Studies on heredity of CNVs in cases of spontaneous abortions are infrequent and may not be the cause of spontaneous abortions. However, this is not consistently the case and further studies should pay more attention to possible incomplete penetrance of the CNVs and possible epigenetic modifications, e.g. by genomic imprinting or X-chromosome inactivation mechanisms. This deepens the requirement for further studies to distinguish the differences between polymorphisms and pathogenic CNVs. It is indicated that recurrent miscarriages with normal karyotypes and unclari fied CNVs may have clinical value for the identification of undiscovered pathologies, including epigenetic abnormalities, telomere DNA deficiency, environmental exposure. CNV is the major type of structure variation caused by genomic rearrangement, which mainly includes deletion and duplication of sub-microscopic genomic segments. CNV has been recognized as one of the major genetic factors underlying human diseases. The mutation rate (per locus) of a CNV is much higher than that of single nucleotide polymorphisms. Various molecular mechanisms are involved in CNV formations, which may be divided into two major categories, DNA recombination-based and DNA replication-based mechanisms. CNVs may be associated with Mendelian diseases, sporadic diseases and susceptibility to complex diseases. CNVs may convey clinical phenotypes by gene dosage, gene disruption, gene fusion and position effects. In the present study, several CNVs with unknown pathogenicity were identified, e.g. case V130 (Table SII), a 334-kb duplication in chromosome 9. The inheritance and phenotype of this CNV and genes involved exhibited inconsistency among different cases and further study should focus on confirming the pathogenicity of these unknown CNVs.

Karyotype abnormalities and pathogenic CNVs have been indicated to be the leading causes of spontaneous abortions (34,35). The present results revealed that chromosomal abnormalities were not relevant to the frequency of spontaneous abortions. Fetal chromosomal screening is generally suggested for patients affected by spontaneous abortion during clinical counseling, regardless of the frequency of miscarriages (36,37). Furthermore, most fetal karyotype abnormalities appeared at random events due to unknown reasons, leading to chromosomal non-disjunction during gamete formation. Following the exclusion of endocrine, immune, infection and normality of chromosomes of parents, natural conception may proceed. If fetal karyotypes are normal but the structures are unbalanced, the chromosomal status of the parents should be first assessed to establish the inheritance of the fetal chromosome imbalance (18,38,39). If the structural abnormality is familial, there may be a high risk of relapse and therefore, pre-implantation genetic screening is suggested to select a normal embryo. If the fetal karyotype is normal but possesses CNVs, the pathogenicity and inheritance of CNVs should be clarified and clinical counseling should be performed according to the classification of CNVs, regardless of a pathogenic, benign or VOUS nature. Finally, if the fetal chromosomes of the parents are normal in RSA, maternal and environmental factors, including immune, endocrine, infections and gene expression abnormalities, should be further explored to better understand the mechanisms of spontaneous abortion.

Acknowledgements

Not applicable.

Funding

This study was supported by a grant from Revitalize and the Defense Project in Science and Education of the Developing Department of Genetic and Reproductive Medicine of Wuxi, Jiangsu Province (grant no. KJQW-YC2017), a key talent's subsidy from the Maternal and Child Health Care Project of Jiangsu Province (grant no. FRC201745) and the Wuxi Science and Technology Development Fund (grant no. CSE31N1511).

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors' contributions

JX and LY designed the study and prepared the manuscript. JX reviewed the manuscript. LY and TT conceived the review and modified the manuscript. XZ, HT and JS collected the relevant literature and performed molecular biology experiments. LY, TT, YS and FQ designed the study and performed data analysis. YT developed the manuscript and interpreted the data. All authors approved the final manuscript for publication.

Ethics approval and consent to participate

This study was performed in accordance with the Declaration of Helsinki and was approved by the Medical Research Ethics Board of the Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University (Wuxi, China). Informed written consent was obtained from all patients.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.
References