Heliyon 9 (2023) e18868

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Analysis of potential copy-number variations and genes associated with first-trimester missed abortion

Wen Zeng [a], Hong Qi [a], Yang Du [b], Lirong Cai [a], Xiaohui Wen [a], Qian Wan [b], Yao Luo [a], Jianjiang Zhu [a,*]

[a] *Prenatal Diagnosis Center, Haidian District Maternal and Child Health Care Hospital, No.53 Suzhou Street, Haidian District, Beijing 100080, PR China*

[b] *Annoroad Gene Technology Co., Ltd, Beijing 100176, PR China*

A R T I C L E I N F O

Keywords:
Copy number variation sequencing
Missed abortion
MYOM2
SDHA
TPPP

A B S T R A C T

Background: Copy number variation sequencing (CNV-seq) was proven to be a highly effective tool in studying of chromosomal copy number variations (CNVs) in prenatal diagnosis and postnatal cases with developmental abnormalities. However, the overall characteristics of missed abortion (MA) CNVs were largely unexplored.

Methods: We retrospectively analyzed the results of CNV-seq in first-trimester MA. The samples included were single pregnancy loss before 13 gestational weeks, and other potential factors affecting embryonic implantation and development had been excluded. Gene ontology and KEGG enrichment analysis was performed on the smallest overlapping regions (SORs) of high-frequency deletion/duplication.

Result: On the basis of strict inclusion and exclusion criteria, only 152 samples were included in our study. 77 (50.7%) samples displayed chromosome number abnormalities, 32 (21%) showed isolated CNVs, and 43 (28.3%) showed no CNVs. A total of 45 CNVs, ranging in size between 300 Kb and 126.56 Mb were identified, comprising 13 segmental aneuploidies CNVs, and 32 submicroscopic CNVs. Among these CNVs, we screened out four SORs (5q31.3, 5p15.33-p15.2, 8p23.3-p23.2, and 8q22.2–24.3), which were potentially associated with first-term MA. 16 genes were identified as potential miscarriage candidate genes through gene-prioritization analysis, including three genes (*MYOM2*, *SDHA* and *TPPP*) critical for embryonic heart or brain development.

Conclusion: We identified some potential candidate CNVs and genes associated with first-trimester MA. 5q31.3 duplications, 5p15.33-p15.2 deletions, 8p23.3-p23.2 deletions and 8p22.2-p24.3 duplications are four potential candidate CNVs. Additionally, *MYOM2*, *SDHA* and *TPPP* are potential genes associated with first-trimester MA.

1. Introduction

Missed abortion (MA), also known as continued abortion, is a common type of pregnant outcome during early pregnancy [1]. Above 50% of first-trimester miscarriages are attributed to chromosomal abnormalities [2–4], such as aneuploidies, polyploidies,

* Corresponding author.
E-mail address: 99237898@qq.com (J. Zhu).

https://doi.org/10.1016/j.heliyon.2023.e18868
Received 16 February 2023; Received in revised form 16 July 2023; Accepted 1 August 2023
Available online 1 August 2023

unbalanced structural rearrangements and copy number variants (CNVs) [5]. Other factors associated with miscarriage also include endocrine disorders, autoimmune diseases, viral infection, and gene mutations [6,7]. Identification of the cause for MA may be invaluable, as it may streamline the diagnostic testing process, providing good recurrence risk estimation for the couple involved, and reassurance, guidance and preparation for the patient's next pregnancy [8].

Recently, massive parallel sequencing has begun to offer genome-wide detection of CNVs. Compared with chromosomal microarray analysis (CMA), low coverage/pass whole-genome CNV sequencing (CNV-seq) can identify additional and clinically significant CNVs with enhanced resolution and increased sensitivity, as well as better limit of detection for mosaicism events [9]. In addition, by changing the method of library construction and using different sequencing chips, CNV-seq could define and discover cryptic chromosomal rearrangements due to its agnostic advantage to size and banding pattern of chromosomal segments [10]. Therefore, this cost-effective strategy could supersede microarray in terms of quality and cost, and its versatile adaptability in clinical diagnosis [11,12].

Several CNVs, genes, and signaling pathways are found to be potentially associated with pregnancy loss [13–15]. With advances in molecular diagnostic technologies, an increasing number of CNVs have been observed in samples of miscarriage, most of these CNVs are unique; thus, further investigation is required to determine their potential clinical significance. The overall characteristics of miscarriage CNVs (size, gene content and function) remain largely unexplored [16]. Many studies on miscarriage did not elucidate the pathogenicity of the detected CNVs [17]. Similarly, *de novo* rare miscarriages CNVs appear to be infrequent in euploid miscarriages as the vast majority of rare CNVs are parentally inherited [3]. As the cost of sequencing decreases and sequencing depth increases, CNV-seq will become a powerful and affordable technology for high-resolution of CNVs detection, which can uncover more small variants of uncertain significance (VOUS). Therefore, the refined characteristics of these CNVs and gene content in MA samples must need to be further explored. Only a relatively small set of the entire genome is expressed in each type of tissue, and the expression of genes depends on the stage of development [18]. Therefore, gene expression in eukaryotes is specific to each tissue [19]. Analysis the etiology of missed abortion at specific stage of embryonic development is warranted.

In this study, we retrospectively analyzed the results of CNVs in first-trimester MA samples detected by CNV-seq. We aimed to investigate the presence and prevalence of CNVs which could be causative for MA. Moreover, we sought to identify potential miscarriage candidate genes from critical regions of miscarriage-associated CNVs using gene-prioritization analysis.

2. Materials and methods

2.1. Declarations and ethics statement

All methods were performed in accordance with relevant guidelines and regulations. All experimental protocols were approved by prenatal diagnosis center and ethics committee of Haidian District Maternal and child health care hospital (Maternal and Child Development Special Program of Haidian District Maternal and child health care hospital, NO.2017–04, China).

2.2. Subjects

The results of first-trimester MA samples (confirmed by ultrasonic diagnosis) were collected from Beijing Haidian Maternal and Child Health Hospital (Beijing, China) from August 2015 to January 2020. All samples must meet all of the inclusion criteria (single pregnancy loss before 13 gestational weeks) and exclusion criteria (including endometriosis uterine myoma and other factors affecting embryonic implantation and development). Principal component analysis was used to group the samples according to age ($\geq$35 years group and $<$35 years group), times of gestation loss (recurrent miscarriage (RM) group (history of spontaneous abortions$\geq$2), and mode of conception (natural conception and assisted reproduction), respectively. All the patients had signed informed consent forms.

2.3. CNV sequencing

After surgical abortion, the samples were taken from chorionic sac and varied in size from 10 to 100 mg. Abortion tissue samples were washed several times with saline, and then the chorionic villis were carefully separated from maternal deciduas and blood clots to minimize any maternal cell contamination. We performed multi-point sampling to ensure the accuracy of the results. Detection of aneuploidies and CNVs were performed as previously described [11], which has the capacity to detect aneuploidies, unbalanced structural rearrangements and CNVs ($>$100 Kb). In brief, genomic DNA extraction was performed using the Amp Genomic DNA Kit (TIANGEN, China) according to the protocols. 2.5 ng of fragmented genomic DNA was used for the construction of sequencing libraries. Purified libraries were sequenced using the Nextseq 550AR platform. Uniquely aligned reads were kept and counted for each 100 Kb window, equally divided along the chromosomes. The sequencing results are compared with the human reference genome, using bioinformatic analysis to discover possible chromosomal abnormalities in the sample.

Large CNVs ($\geq$10 Mb) were defined as segmental aneuploidies, whereas submicroscopic CNVs ($<$10 Mb) were classified as microdeletions and microduplications [15]. The conventional genomic and phenotype public databases were used for retrieval and interpretation of the identified CNVs, such as UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway), ClinGene (http://www.clinicalgenome.org/), OMIM (http://omim.org), DECIPHER (https://decipher.sanger.ac.uk), DGV (http://dgv.tcag.ca/dgv/app/home) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed). We used a quantitative, evidence-based scoring framework to reappraise the pathogenicity of detected CNVs according to the ACMG technical standards [20]. Benign CNVs were considered as normal chromosomal variants.

2.4. Gene expression ranking among various tissue types

Samples with numerical chromosomal abnormalities were excluded from CNV analysis. For each CNV event, we first retrieved official HGNC gene symbols within the CNV region. On the other hand, we downloaded processed microarray expression data (Human U133A/GNF1H Gene Atlas) from bioGPS [21] and also RNA-seq data from GEO (GSE69360) [22], both of which contained a total of nine expression profiles from fetal tissues and placenta, as well as other 93 normal adult tissue types. The data were individually normalized within each group. Then for each gene, individually for each platform, which is called between Tissue rankings and is calculated as the average percentile of the gene expression in all fetal specific samples among all other samples. We also calculated the within Tissue ranking, which is calculated as the average percentile of the gene expression within each fetal sample among other genes. The two ranking scores were then summed up as one for each gene within each platform (namely gps score for the bioGPS arrayset and seq_score for the RNA-seq dataset in the supplement table), an averaged fetal specific gene expression ranking, for which the higher the numeric score the more specific and the higher the expression lever of the gene during fetal development. Large CNVs were found in at least three cases, whereas submicroscopic CNVs were found in two or more cases, and such CNVs were evaluated for the smallest overlapping regions (SORs). Gene ontology (GO) and KEGG enrichment analysis were performed on SORs using the DAVID bioinformatics database (https://david.ncifcrf.gov/) [23–25].

2.5. Short tandem repeat (STR) analysis

Sex chromosome-specific marker (AMEL) and 15 STR polymorphisms localized on 13 autosomes (PowerPlex® 16 HS, Promega, American) were analyzed on the basis of the operating procedure to exclude potential maternal contamination. Among the 16 STR loci, at least three loci had more than two alleles, which can be used to identify maternal contamination.

2.6. Data and statistical analysis

The chi-squared test was utilized to compare the frequency of chromosomal abnormal between the groups by using SPSS statistics (version 22.0) [26]. $P < 0.05$ was considered statistically significant.

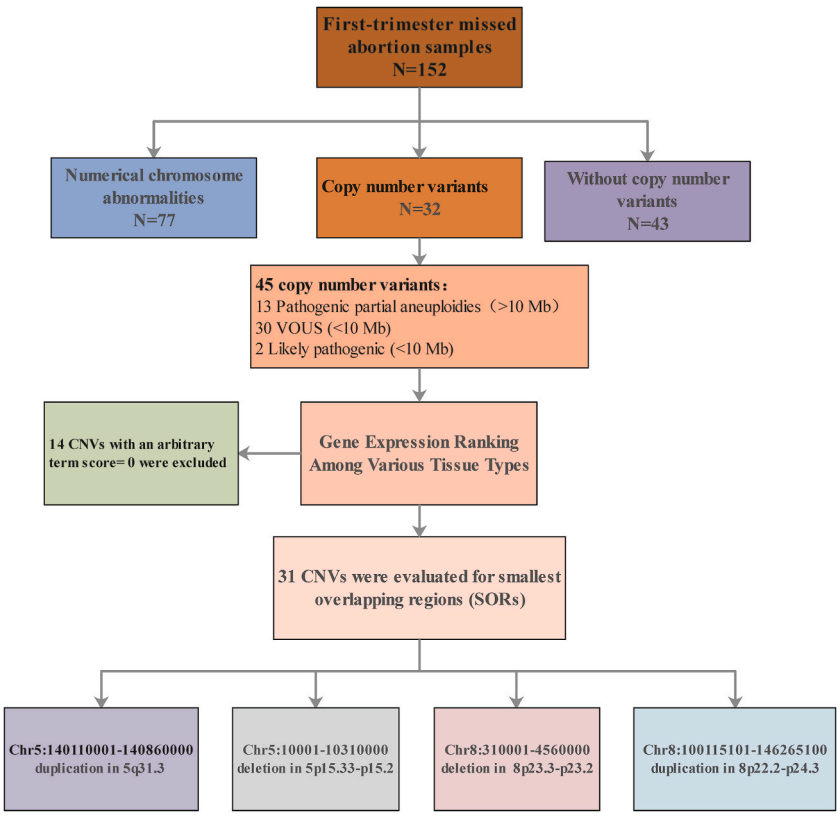


Fig. 1. Flowchart depicting the details of the testing strategy for the characterized chromosomal copy number variations (CNVs) in this study.

3. Results

The testing strategies were summarized in Fig. 1. A total of 152 samples were included in our study. Average gestational weeks and maternal age were 8.1 ± 2.3weeks and 32.9 ± 4.3 years (range 25~45), respectively. The average times of gestation, and miscarriage per woman were 2.3 ± 1.3 and 1.6 ± 0.8, respectively. In addition, normal results were identified in 43 samples (28.3%), numerical chromosomal abnormalities were identified in 77 (50.7%) samples, and 45 CNVs were identified in 32 (21%) samples (Table 1 and Supplemental Table 1, Supplemental Table 2).

Numerical chromosomal abnormalities were higher in the ≥35 year group but no statistically significant difference was found ($\chi^2 = 0.13$, $P > 0.05$). In addition, no statistically significant difference was found between the RM (≥2) group and spontaneous abortions≤1 group ($\chi^2 = 0.04$, $P > 0.05$), as well as between natural conception and assisted reproduction ($\chi^2 = 0.00$, $P > 0.05$).

3.1. Numerical chromosomal abnormalities are common causes of first-trimester MA

50.7% of samples had numerical chromosomal abnormalities. The frequently observed numerical abnormal chromosomes were chromosomes X, 16, 15, 21, 22, 2, 7 and 13, which was arranged in a descending order (Supplemental Table 3). A total of eight samples exhibited mosaicisms with different proportions. The results showed that the proportion of mosaicism ranged from 7% to 79%. We identified one sample with the lowest proportion of 7% with double-trisomy (T5, T7) and mixed sex chromosome (ET0995). The results of CNV-seq showed that abnormalities were found in 5, 7, and sex chromosomes. The relative dosage of chr5 versus other autosomes was 1.08:1, whereas that of chr7 versus other autosomes was 1.07:1. Furthermore, the ratio between chromosomes X and Y was 0.58:0.47 (Fig. 2).

3.2. Segmental aneuploidies as a contributor to miscarriage had a prevalence of 5.3% in our study, and only two likely pathogenic (LP) submicroscopic CNVs (<10 Mb) were observed

A total of 13 pathogenic segmental aneuploidies (>10 Mb) were identified in eight samples (5.3%). In addition, 32 submicroscopic CNVs (30 VOUS, 2 LP) were identified in 26 samples, ranging in size from 300 Kb to 6.55 Mb. Seven samples presented deletion combined with duplication (ET0628, ET1010 had been confirmed to arise from a balanced parental rearrangement). CNVs occurred most frequently on chromosome 8 (9 times in 5 samples), followed by chromosomes 5, 3, and 10.

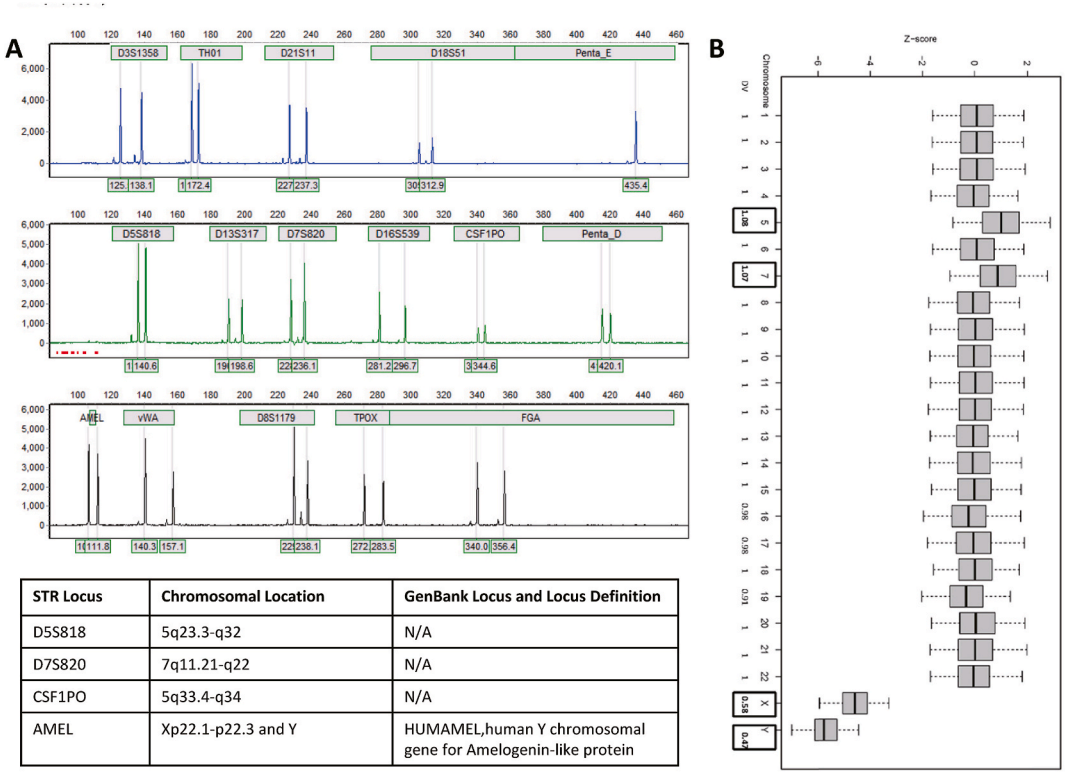
3.3. 4 SORs and 16 genes were potentially associated with first-trimester MA

Gene (highly expressed in fetal development) expression ranking among various tissue types is shown in Supplemental Table 4. A total of 14 CNVs did not contain genes specific and highly expressed in fetal development; thus, such CNVs were excluded from further analysis. We screened out four SORs from high-frequency deletion/duplication regions, which might be potentially associated with first-trimester MA, including duplication in 5q31.3 (chr5:140110001- 140860000, 750 Kb), deletion in 5p15.33-p15.2 (chr5:10001-10310000, 10.3 Mb), deletion in 8p23.3-p23.2 (chr8:310001-4560000, 4.25 Mb), and duplication in 8p22.2-p24.3 (chr8:100115101-146265100, 46.15 Mb) (Fig. 3). Genes from 8p22.2-p24.3 duplication were not significantly enriched in any specific pathways, and they did not result in any significant GO and KEGG terms. A total of 16 genes from regions 5q31.3, 5p15.33-p15.2 and 8p23.3-p23.2 were enriched in the following functional categories: basal transcription factors, membrane adhesion molecules, transmembrane transporter activity, symporter activity, membrane adhesion molecules, and signaling (Table 2, Fig. 3, and Supplemental Table 5).

Table 1
Summary of chromosomal abnormalities detected by CNV-seq.

	Total	Age of pregnant woman		Number of Miscarriages (include this time)		Conception way	
	152	<35	≥35	1	≥2(RM group)	natural onception	assisted reproductive
Normal	**43(28.3%)**	**29**	**14**	**23**	**20**	**35**	**8**
Numerical abnormalities	**77(50.7%)**	**48**	**29**	**44**	**33**	**62**	**15**
Polyploidy *n* (%)	7	5	2	6	1	6	1
Trisomy *n* (%)	42	21	21	20	22	36	6
Double trisomy *n* (%)	5	3	2	2	3	2	3
Triple trisomy *n* (%)	1	0	1	1	0	0	1
Monosomy X *n* (%)	12	11[b,c]	1	9[b,c]	3	8[b]	4[c]
Mosaicism *n* (%)	8	6[a]	2	4	4[a]	8[a]	0
Monosomy X + trisomy *n* (%)	2	2	0	2	0	2	0
Insolated CNVs	**32(21%)**	**30**	**2**	**19**	**13**	**29**	**3**

[a] ET1105:seq[GRCh37]Xp22.33-Xq28 × 1/Xp22.33-Xq28 × 1,dup(X)(p11.22q12).
[b] ET0837:seq[GRCh37]Xp22.33-Xq28 × 1,dup(10)(q11.22q11.23).
[c] ET0510:seq[GRCh37]Xp22.33-Xq28 × 1,dup(3)(p26.3p26.1).



STR Locus	Chromosomal Location	GenBank Locus and Locus Definition
D5S818	5q23.3-q32	N/A
D7S820	7q11.21-q22	N/A
CSF1PO	5q33.4-q34	N/A
AMEL	Xp22.1-p22.3 and Y	HUMAMEL,human Y chromosomal gene for Amelogenin-like protein

Fig. 2. STR and CNV-seq results of ET0995 with the lowest proportion of 7% mosaicisms. (A) STR analysis showed no evident maternal contamination. (B) Chromosome Z-score boxplots and DV showed mosaicisms of 7, 8 and sex chromosomes.

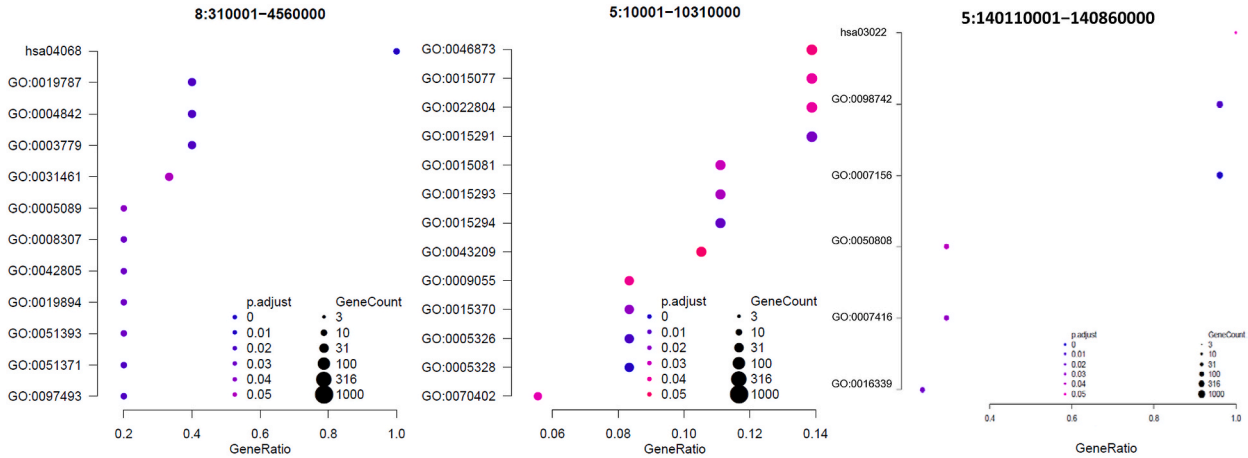


Fig. 3. GO and KEGG enrichment of the smallest overlapping regions (SORs).

4. Discussion

In this study, we applied CNV-seq to investigate the incidence and distribution of chromosomal abnormalities in first-trimester MA samples. As expected, numerical abnormality was the main cause of MA, which was correlated with maternal age, but not with the times of gestation loss and the mode of conception. In addition, no significant correlation was observed between maternal age and risk of Turner's syndrome, which was consistent with previous reports [27,28]. We also confirmed that CNV-seq could be sensitive and specific for the detection of low-level mosaicism, but it cannot accurately deduce the exact composition of the mixture (particularly sex chromosomal mosaicism/chimerism).

Segmental aneuploidies as a contributor to miscarriage were identified with a prevalence of 5.3% in our study, which was higher

Table 2
4 smallest overlapping regions (SORs) selected from high-frequency deletion/duplication regions and genes listed in each enriched GO and KEGG term.

SORs	Chr	del/dup	Clinical significance	Number of cases affected	location	Size (Kb)	Genes enriched (P < 0.05)
1	5q31.3	dup	LP	3	140,110,001–140,860,000	750	TAF7
2	5p15.33-p15.2	del	P	2	10,001–10,310,000	10,300	SLC9A3, SLC12A7, SLC6A19, SLC6A18, SLC6A3, SRD5A1, MTRR, SDHA, NDUFS6, TPPP, CCT5
3	8p23.3-p23.2	del	VOUS	4	310,001–4,560,000	4250	FBXO25,MYOM2, KBYBD11, ARHGEF10
4	8p22.2-p24.3	dup	P	3	100,115,101–146,265,100	46,150	–

than that reported in previous reports [15,29]. Consistent with previous studies [3,15], segmental aneuploidies occurred most frequently on chromosome 8, especially around 8p23 (4 of 5 samples). The high frequent occurrence of segmental aneuploidies on chromosome 8p with gross deletion or duplication occurred in our study, which may reflect the bias of limited sample size. However, the reasons underlying this bias also cannot be ignored: in particular, the 8p23 inversion, which is the most common genomic polymorphism on autosomes, can lead to the formation of chromosome 8 rearrangements. In addition, the populational minor allele frequency of 8p23 inversion was estimated to be 27% in the general Japanese population [30,31]. These critical regions identified in miscarriage cases provide a unique source for prioritizing miscarriage candidate genes [15]. We identified two SORs encompassing this high-frequency deletion/duplication interval, segmental aneuploidy duplication in 8p22.2-p24.3 and microdeletion in 8p23.3-p23.2. Genes from 8p22.2-p24.3 duplication were not significantly enriched in any specific KEGG pathway or GO term. A total of four genes (*FBXO25*, *MYOM2*, *KBYBD11*, and *ARHGEF10*) in 8p23.3-p23.2 were associated with early fetal development. Enriched biological process terms included the FoxO signaling pathway, actinin binding, transferase activity, and Rho guanyl-nucleotide exchange factor activity. *MYOM2* has high expression in the heart (RPKM 193.6) [32], and it not only plays a critical role in maintaining robust heart function, but also serves as a candidate gene for several heart related diseases, as it is clearly involved in the development of the heart [33]. Accordingly, we suggest that *MYOM2* is a potential candidate gene, and early embryonic death could be due to major heart malformations resulting from 8p23.3–23.2 microdeletion.

Another SOR of deletion in 5p15.33-p15.2 was associated with Cri du Chat syndrome (CdCs). A total of 11 genes in the region were functional enriched for several GO terms and KEGG pathways. Five of these genes were solute carrier (SLCs) family members, the largest family of transmembrane transporters that determine the exchange of various substances, including nutrients, ions, metabolites, and drugs across biological membranes. Among these genes, *SLC6A3* was considered dose-sensitive or conditionally haploinsufficient [34]. Interaction between *SLC6A3*, *TPPP* and *CCT5* are reported to be related to neuronal development and function in CdCs [35]. *TPPP* has relatively high expression in the brain (RPKM 45.7) and lung (RPKM 7.7) [32]. It plays an important role in the pathogenesis of the brain through the co-enrichment and co-localization of *TPPP* and α-synuclein in human brain inclusions. Furthermore, the *SDHA* gene has ubiquitous expression in the heart (RPKM 63.7) and kidney (RPKM 34.3) [32], which is a crucial contributor to the mitochondrial respiratory chain. Based on the non-human mutagenesis study, no reported differences was observed in behavioral, biochemical, or molecular evaluations between WT and *SDHA* ±rats at 6 weeks or 6 months of age, but 100% of *SDHA*−/− rats died prior to birth [36]. Furthermore, homozygous *SDHA* mutations or compound heterozygous *SDHA* mutations often result in early death [37]. Moreover, 75% of the registered deaths in 5p-patients occurred during the first months after birth and 90% during the first year of life. Therefore, we suggest that defects in these genes within the critical regions of CdCs may contribute to the first-trimester MA.

The role of submicroscopic CNVs remains unclear whether these abnormalities contribute to miscarriage. A systematic review of 36 studies on early pregnancy loss (up to 20 gestational weeks) suggested that the common pathogenic CNVs reported were 22q11.21 and 1p36.33 deletion, size ranging between 100 Kb and 10 Mb [38]. Compared with healthy controls, three statistically significant recurrent pathogenic submicroscopic CNVs (microdeletions in 22q11.21, 2q37.3 and 9p24.3p24.2, size ranging between 400 Kb and 8 Mb) were considered to be associated with early pregnancy loss before 13 gestational weeks, whereas no VOUS were statistically prevalent in miscarriage cases [15]. In this study, none of above mentioned submicroscopic pathogenic CNVs has been detected, which might be related to the differences in population and sample inclusion criteria, but this finding is more likely a reflection of non-canonical functional relevancy of such microscopic VOUS in first trimester MA. We screened out submicroscopic SORs of 5q31.3 duplication, which were rarely reported comparing with 5q31.3 microdeletion syndrome. The TAF7 gene was enriched in a pathway that is essential for embryonic development. Homozygous deletion of TAF7 was embryonic lethal, but heterozygous TAF7 ± mice were not haploinsufficient [39]. In addition, one published case showed a similar duplication in the DGV database (esv2758019, 671.7 Kb). Thus, although 5q31.3 duplication was identified in three MA samples in our study, it might not be functionally associated with the early abortion.

Cataloguing of all CNVs and detailed description of their characteristics (e.g. gene content, genomic breakpoints) is desirable in the future, for better understanding of their relevance in pregnancy loss [16]. Nevertheless a large proportion of the novel CNVs of Asians were not catalogued in DGV [40], making the interpretation and genetic counseling of VOUS difficult and challenging for laboratory technician and clinical geneticist [41]. We found 30 (93.7%) VOUS in submicroscopic CNVs, 25 of them were smaller than 1 Mb. Most VOUS may be inherited [42]. Therefore additional examination of parental samples, and the comprehensive family-based analysis are

helpful for the interpretation and genetic counseling of reports with VOUS [43]. Interestingly, medical complications arise in the cases when CNVs inherited from a healthy carrier parent could potentially lead to miscarriage or recurrent pregnancy loss [5,44,45]. Therefore, parental verification is necessary to provide additional evidence for clinical diagnosis and support further interventional options in re-pregnancy. However, as regard to VOUS, even if the families spend more money on parental verification, it still cannot provide additional useful information for clinical management and re-pregnancy, but rather caused anxiety for the couple when CNVs were inherited. This emphasizes that the clinical utility of applying CNV-seq analysis would not only rely on the accuracy of variant detection but also build on the accumulated data sets for interpretation. Retrospective analysis supports a more clinically appropriate reporting threshold of $\geq$1 Mb for duplications of uncertain clinical significance CNVs detected by any methodology, including CMA and exome/genome sequencing [42]. Therefore, reporting only pathogenic or highly suspected clinically significant CNVs (even $<$1 Mb) and CNVs $>$1 Mb in clinical report of first-trimester MA may be a good option, which may significantly reduce the difficulty of clinical consultation and anxiety of parents. This emphasizes the need to extend clinical utility of refined screening technologies like CNV-seq to provide comprehensive profiles of the human genomic variants, and also more importantly to build an accumulated knowledge base for better clinical variant interpretation [46].

The contribution of CNVs to miscarriages is complex, and even CNVs (such as 22q11 microdeletions) with the same fragment size may lead to different pregnancy outcomes, from first-trimester MA to mid-term ultrasound structural congenital anomalies. Therefore, the impact of CNVs on pregnancy is determined not only by their sizes, locations and genetic contents, but also by other co-existing variants such as single-nucleotide variants and deleterious gene mutations [7]. At present, no data were found for supporting the clinical use of exome sequencing (ES) for some reproductive indications, such as a history of recurrent unexplained pregnancy loss [47]. Even if further genetic results are obtained, appropriate medical management strategy for such couples is still lacking. Therefore, the application of ES technology in miscarriage is still limited to scientific research rather than clinical application. Furthermore, multi-level interactions between the genome, epigenome and environmental factors might occur [48]. Furthermore, numerous lines of evidence suggest the influence of epigenome variation on health and production [49,50]. The correlation between epigenome and missed abortion is also worth studying.

Beyond all doubt, there are limitations to our study. The small sample size and with limited number of samples with parental verification, our experience shows phenotypically normal carrier parents are quite common in such VOUS cases. As a result of such complication in practice, we rarely encourage parental verification of VOUS to avoid additional social-economic burden to the families, but more relying on verified clinical evidences provided in public knowledge databases to confirm its symptom-causing potential. This further remind us the improved clinical utility of CNV-seq would not only rely on the accuracy of variant detection but more importantly depend on the accumulated medical knowledge for genetic interpretation. Therefore, continuous data accumulation effort still needed, which is also the direction of our work in the future. In future work, we are adapting more customized reporting and scoring procedures for VOUS such as obtaining additional parental samples to access the origin of the variants, and applying cWES for those samples carrying no detectable variant in CNV-seq and karyotyping. In addition, with the merit of its highly automated and multiplexing experimental process comparing to traditional karyotyping, the routine application of CNV-seq in clinical diagnosis still limited by its cost, which should be further relaxed as technology progresses. With a better covered patient population using such innovative diagnostic tools, we should be able to better understand the development of early trimester MA.

5. Conclusion

We identified four potential first-trimester MA candidate CNVs. A total of 16 genes were identified as potential candidate genes through gene-prioritization analysis, including three genes (*MYOM2*, *SDHA* and *TPPP*) critical for embryonic heart or brain development. Our findings extended the CNVs spectrum of first-trimester MA and provided a promising source for future functional validation of MA-related genes.

Ethics approval and consent to participate

Samples-related data analyzed for this manuscript were entirely retrospective with no patient or patient-related identifiers included in the analysis. The written informed consent was obtained from every participants or the participants' family. There was no participant under the age of 16. The participants gave informed consent to donate the remaining samples and use their data to scientific research, technical innovation and publish scientific papers after the identifiable personal information was removed.

Author contribution statement

Wen Zeng: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.
Hong Qi: Conceived and designed the experiments; Analyzed and interpreted the data.
Yang Du: Performed the experiments; Analyzed and interpreted the data.
Lirong Cai, Xiaohui Wen, Qian Wan, Yao Luo: Performed the experiments; Contributed reagents, materials, analysis tools or data.
Jianjiang Zhu: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank all the participants for their cooperation in this study and Annoroad (Beijing) Gene Technology for their expert laboratory work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e18868.

References

[1] S.G. Petersen, A.R. Perkins, K.S. Gibbons, J.I. Bertolone, K. Mahomed, The medical management of missed miscarriage: outcomes from a prospective, single-centre, Australian cohort, Med. J. Aust. 199 (2013) 341–346.
[2] P.A. Jacobs, T.J. Hassold, The origin of numerical chromosome abnormalities, Adv. Genet. 33 (1995) 101–133.
[3] B. Levy, S. Sigurjonsson, B. Pettersen, M.K. Maisenbacher, M.P. Hall, Z. Demko, R.B. Lathi, R. Tao, V. Aggarwal, M. Rabinowitz, Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis, Obstet. Gynecol. 124 (2014) 202–209.
[4] T. Segawa, T. Kuroda, K. Kato, M. Kuroda, K. Omi, O. Miyauchi, Y. Watanabe, T. Okubo, H. Osada, S. Teramoto, Cytogenetic analysis of the retained products of conception after missed abortion following blastocyst transfer: a retrospective, large-scale, single-centre study, Reprod. Biomed. Online 34 (2017) 203–210.
[5] C.D. Viaggi, S. Cavani, M. Malacarne, F. Floriddia, G. Zerega, C. Baldo, M. Mogni, M. Castagnetta, G. Piombo, D.A. Coviello, F. Camandona, D. Lijoi, W. Insegno, M. Traversa, M. Pierluigi, First-trimester euploid miscarriages analysed by array-CGH, J. Appl. Genet. 54 (2013) 353–359.
[6] M. Fu, S. Mu, C. Wen, S. Jiang, L. Li, Y. Meng, H. Peng, Wholeexome sequencing analysis of products of conception identifies novel mutations associated with missed abortion, Mol. Med. Rep. 18 (2018) 2027–2032.
[7] Y. Qiao, J. Wen, F. Tang, S. Martell, N. Shomer, P.C. Leung, M.D. Stephenson, E. Rajcan-Separovic, Whole exome sequencing in recurrent early pregnancy loss, Mol. Hum. Reprod. 22 (2016) 364–372.
[8] R.B. Lathi, F.K. Gray Hazard, A. Heerema-McKenney, J. Taylor, J.T. Chueh, First trimester miscarriage evaluation, Semin. Reprod. Med. 29 (2011) 463–469.
[9] H. Wang, Z. Dong, R. Zhang, M.H.K. Chau, Z. Yang, K.Y.C. Tsang, H.K. Wong, B. Gui, Z. Meng, K. Xiao, X. Zhu, Y. Wang, S. Chen, T.Y. Leung, S.W. Cheung, Y. K. Kwok, C.C. Morton, Y. Zhu, K.W. Choy, Low-pass Genome Sequencing versus Chromosomal Microarray Analysis: Implementation in Prenatal Diagnosis, Genetics in Medicine, official journal of the American College of Medical Genetics, 2019.
[10] Z. Dong, J. Yan, F. Xu, J. Yuan, H. Jiang, H. Wang, H. Chen, L. Zhang, L. Ye, J. Xu, Y. Shi, Z. Yang, Y. Cao, L. Chen, Q. Li, X. Zhao, J. Li, A. Chen, W. Zhang, H. G. Wong, Y. Qin, H. Zhao, Y. Chen, P. Li, T. Ma, W.J. Wang, Y.K. Kwok, Y. Jiang, A.N. Pursley, J.P.W. Chung, Y. Hong, K. Kristiansen, H. Yang, R.E. Pina-Aguilar, T.Y. Leung, S.W. Cheung, C.C. Morton, K.W. Choy, Z.J. Chen, Genome sequencing explores complexity of chromosomal abnormalities in recurrent miscarriage, Am. J. Hum. Genet. 105 (2019) 1102–1111.
[11] H. Qi, Z.L. Xuan, Y. Du, L.R. Cai, H. Zhang, X.H. Wen, X.D. Kong, K. Yang, Y. Mi, X.X. Fu, S.B. Cao, J. Wang, C.J. Chen, J.B. Liang, High resolution global chromosomal aberrations from spontaneous miscarriages revealed by low coverage whole genome sequencing, Eur. J. Obstet. Gynecol. Reprod. Biol. 224 (2018) 21–28.
[12] Z. Dong, J. Zhang, P. Hu, H. Chen, J. Xu, Q. Tian, L. Meng, Y. Ye, J. Wang, M. Zhang, Y. Li, H. Wang, S. Yu, F. Chen, J. Xie, H. Jiang, W. Wang, K.W. Choy, Z. Xu, Low-pass whole-genome sequencing in clinical cytogenetics: a validated approach, Genetics in medicine, Off. J. Am. Coll. Med. Gen. 18 (2016) 940–948.
[13] Y. Chen, J. Bartanus, D. Liang, H. Zhu, A.M. Breman, J.L. Smith, H. Wang, Z. Ren, A. Patel, P. Stankiewicz, D.S. Cram, S.W. Cheung, L. Wu, F. Yu, Characterization of chromosomal abnormalities in pregnancy losses reveals critical genes and loci for human early development, Hum. Mutat. 38 (2017) 669–677.
[14] Y.R. Sheng, S.Y. Hou, W.T. Hu, C.Y. Wei, Y.K. Liu, Y.Y. Liu, L. Jiang, J.J. Xiang, X.X. Sun, C.X. Lei, H.L. Wang, X.Y. Zhu, Characterization of copy-number variations and possible candidate genes in recurrent pregnancy losses, Genes 12 (2021).
[15] Y. Wang, Y. Li, Y. Chen, R. Zhou, Z. Sang, L. Meng, J. Tan, F. Qiao, Q. Bao, D. Luo, C. Peng, Y.S. Wang, C. Luo, P. Hu, Z. Xu, Systematic analysis of copy-number variations associated with early pregnancy loss, Ultrasound Obstet. Gynecol. 55 (2020) 96–104.
[16] H. Bagheri, E. Mercier, Y. Qiao, M.D. Stephenson, E. Rajcan-Separovic, Genomic characteristics of miscarriage copy number variants, Mol. Hum. Reprod. 21 (2015) 655–661.
[17] S. Bug, B. Solfrank, F. Schmitz, J. Pricelius, M. Stecher, A. Craig, M. Botcherby, C. Nevinny-Stickel-Hinzpeter, Diagnostic utility of novel combined arrays for genome-wide simultaneous detection of aneuploidy and uniparental isodisomy in losses of pregnancy, Mol. Cytogenet. 7 (2014) 43.
[18] M. Mohammadabadi, S.H. Masoudzadeh, A. Khezri, O. Kalashnyk, R.V. Stavetska, N.I. Klopenko, V.P. Oleshko, S.V. Tkachenko, Fennel (Foeniculum vulgare) seed powder increases Delta-Like Non-Canonical Notch Ligand 1 gene expression in testis, liver, and humeral muscle tissues of growing lambs, Heliyon 7 (2021), e08542.
[19] F. Mohammadinejad, M. Mohammadabadi, Z. Roudbari, T. Sadkowski, Identification of Key Genes and Biological Pathways Associated with Skeletal Muscle Maturation and Hypertrophy in Bos taurus, Ovis aries, and Sus scrofa, 2022, p. 12. Animals (Basel).
[20] E.R. Riggs, E.F. Andersen, A.M. Cherry, S. Kantarci, H. Kearney, A. Patel, G. Raca, D.I. Ritter, S.T. South, E.C. Thorland, D. Pineda-Alvarez, S. Aradhya, C. L. Martin, Technical Standards for the Interpretation and Reporting of Constitutional Copy-Number Variants: a Joint Consensus Recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen), Genetics in Medicine, official journal of the American College of Medical Genetics, 2019.
[21] A.I. Su, T. Wiltshire, S. Batalov, H. Lapp, K.A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, M.P. Cooke, J.R. Walker, J.B. Hogenesch, A gene atlas of the mouse and human protein-encoding transcriptomes, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 6062–6067.
[22] J.Y. Choy, P.L. Boon, N. Bertin, M.J. Fullwood, A resource of ribosomal RNA-depleted RNA-Seq data from different normal adult and fetal human tissues, Sci. Data 2 (2015), 150063.
[23] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30.
[24] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nat. Genet. 25 (2000) 25–29.
[25] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat. Protoc. 4 (2009) 44–57.

[26] IBM Corp. Released, IBM SPSS Statistics for Windows, IBM Corp, NY, 2013.
[27] C.H. Gravholt, S. Juul, R.W. Naeraa, J. Hansen, Prenatal and postnatal prevalence of Turner's syndrome: a registry study, BMJ 312 (1996) 16–21.
[28] C.H. Gravholt, S. Juul, R.W. Naeraa, J. Hansen, [Prenatal and postnatal prevalence of Turner syndrome. A registry-based study], Ugeskr Laeger 159 (1997) 3160–3166.
[29] T. Sahoo, N. Dzidic, M.N. Strecker, S. Commander, M.K. Travis, C. Doherty, R.W. Tyson, A.E. Mendoza, M. Stephenson, C.A. Dise, C.W. Benito, M.S. Ziadie, K. Hovanes, Comprehensive genetic analysis of pregnancy loss by chromosomal microarrays: outcomes, benefits, and challenges, Genetics in medicine, Off. J. Am. Coll. Med. Gen. 19 (2017) 83–89.
[30] A. Soler, A. Sanchez, A. Carrio, C. Badenas, M. Mila, A. Borrell, Fetoplacental discrepancy involving structural abnormalities of chromosome 8 detected by prenatal diagnosis, Prenat. Diagn. 23 (2003) 319–322.
[31] H. Sugawara, N. Harada, T. Ida, T. Ishida, D.H. Ledbetter, K. Yoshiura, T. Ohta, T. Kishino, N. Niikawa, N. Matsumoto, Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23, Genomics 82 (2003) 238–244.
[32] L. Fagerberg, B.M. Hallstrom, P. Oksvold, C. Kampf, D. Djureinovic, J. Odeberg, M. Habuka, S. Tahmasebpoor, A. Danielsson, K. Edlund, A. Asplund, E. Sjostedt, E. Lundberg, C.A. Szigyarto, M. Skogs, J.O. Takanen, H. Berling, H. Tegel, J. Mulder, P. Nilsson, J.M. Schwenk, C. Lindskog, F. Danielsson, A. Mardinoglu, A. Sivertsson, K. von Feilitzen, M. Forsberg, M. Zwahlen, I. Olsson, S. Navani, M. Huss, J. Nielsen, F. Ponten, M. Uhlen, Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics, Mol. Cell. Proteomics 13 (2014) 397–406.
[33] E. Auxerre-Plantie, T. Nielsen, M. Grunert, O. Olejniczak, A. Perrot, C. Ozcelik, D. Harries, F. Matinmehr, C. Dos Remedios, C. Muhlfeld, T. Kraft, R. Bodmer, G. Vogler, S.R. Sperling, Identification of MYOM2 as a candidate gene in hypertrophic cardiomyopathy and Tetralogy of Fallot, and its functional evaluation in the Drosophila heart, Dis. Model Mech. 13 (2020).
[34] J.M. Nguyen, K.J. Qualmann, R. Okashah, A. Reilly, M.F. Alexeyev, D.J. Campbell, 5p deletions: current knowledge and future directions, Am. J. Med. Gen. C Semin. Med. Gen. 169 (2015) 224–238.
[35] T. Correa, B.C. Feltes, M. Riegel, Integrated analysis of the critical region 5p15.3-p15.2 associated with cri-du-chat syndrome, Genet. Mol. Biol. 42 (2019) 186–196.
[36] E.M. Siebers, M.J. Choi, J.A. Tinklenberg, M.J. Beatka, S. Ayres, H. Meng, D.C. Helbling, A. Takizawa, B. Bennett, A.M. Garces, L.G. Dias Duarte Machado, D. Dimmock, M.R. Dwinell, A.M. Geurts, M.W. Lawlor, Sdha+/- rats display minimal muscle pathology without significant behavioral or biochemical abnormalities, J. Neuropathol. Exp. Neurol. 77 (2018) 665–672.
[37] A.S. Hoekstra, J.P. Bayley, The role of complex II in disease, Biochim. Biophys. Acta 1827 (2013) 543–551.
[38] M. Pauta, M. Grande, L. Rodriguez-Revenga, E. Kolomietz, A. Borrell, Added value of chromosomal microarray analysis over karyotyping in early pregnancy loss: systematic review and meta-analysis, Ultrasound in obstetrics & gynecology, Off. J. Int. Soci. Ultrasound Obs. Gynec. 51 (2018) 453–462.
[39] A. Gegonne, X. Tai, J. Zhang, G. Wu, J. Zhu, A. Yoshimoto, J. Hanson, C. Cultraro, Q.R. Chen, T. Guinter, Z. Yang, K. Hathcock, A. Singer, J. Rodriguez-Canales, L. Tessarollo, S. Mackem, D. Meerzaman, K. Buetow, D.S. Singer, The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells, Mol. Cell Biol. 32 (2012) 1984–1997.
[40] H. Xu, W.T. Poh, X. Sim, R.T. Ong, C. Suo, W.T. Tay, C.C. Khor, M. Seielstad, J. Liu, T. Aung, E.S. Tai, T.Y. Wong, K.S. Chia, Y.Y. Teo, SgD-CNV, a database for common and rare copy number variants in three Asian populations, Hum. Mutat. 32 (2011) 1341–1349.
[41] L.A. Kiedrowski, K.M. Owens, B.M. Yashar, J.L. Schuette, Parents' perspectives on variants of uncertain significance from chromosome microarray analysis, J. Genet. Counsel. 25 (2016) 101–111.
[42] C.A. Marcou, B. Pitel, C.E. Hagen, N.J. Boczek, R.A. Rowsey, L.B. Baughn, N.L. Hoppman, E.C. Thorland, H.M. Kearney, Limited Diagnostic Impact of Duplications <1 Mb of Uncertain Clinical Significance: a 10-year Retrospective Analysis of Reporting Practices at the Mayo Clinic, Genetics in Medicine, official journal of the American College of Medical Genetics, 2020.
[43] P. Shi, R. Li, C. Wang, X. Kong, Influence of validating the parental origin on the clinical interpretation of fetal copy number variations in 141 core family cases, Mol. Gen. Genom. Med. 7 (2019), e00944.
[44] E. Colley, S. Hamilton, P. Smith, N.V. Morgan, A. Coomarasamy, S. Allen, Potential genetic causes of miscarriage in euploid pregnancies: a systematic review, Hum. Reprod. Update 25 (2019) 452–472.
[45] E. Rajcan-Separovic, D. Diego-Alvarez, W.P. Robinson, C. Tyson, Y. Qiao, C. Harvard, C. Fawcett, D. Kalousek, T. Philipp, M.J. Somerville, M.D. Stephenson, Identification of copy number variants in miscarriages from couples with idiopathic recurrent pregnancy loss, Hum. Reprod. 25 (2010) 2913–2922.
[46] Z. Dong, W. Xie, H. Chen, J. Xu, H. Wang, Y. Li, J. Wang, F. Chen, K.W. Choy, H. Jiang, Copy-number variants detection by low-pass whole-genome sequencing, Curr. Prot. Human Gen. 94 (2017) 8 17 11–18 17 16.
[47] K.G. Monaghan, N.T. Leach, D. Pekarek, P. Prasad, N.C. Rose, The use of fetal exome sequencing in prenatal diagnosis: a points to consider document of the American College of Medical Genetics and Genomics (ACMG), Genetics in medicine, Off. J. Am. Coll. Med. Gen. 22 (2020) 675–680.
[48] L. Mohamadipoor Saadatabadi, M. Mohammadabadi, Z. Amiri Ghanatsaman, O. Babenko, R. Stavetska, O. Kalashnik, D. Kucher, O. Kochuk-Yashchenko, H. Asadollahpour Nanaei, Signature selection analysis reveals candidate genes associated with production traits in Iranian sheep breeds, BMC Vet. Res. 17 (2021) 369.
[49] F. Bordbar, M. Mohammadabadi, J. Jensen, L. Xu, J. Li, L. Zhang, Identification of Candidate Genes Regulating Carcass Depth and Hind Leg Circumference in Simmental Beef Cattle Using Illumina Bovine Beadchip and Next-Generation Sequencing Analyses, Animals (Basel), 2022, p. 12.
[50] S.M.H. Safaei, M. Dadpasand, M. Mohammadabadi, H. Atashi, R. Stavetska, N. Klopenko, O. Kalashnyk, An Origanum Majorana Leaf Diet Influences Myogenin Gene Expression, Performance, and Carcass Characteristics in Lambs, Animals (Basel), 2022, p. 13.