

Ministry of Science and Higher Education of the Russian Federation
NATIONAL RESEARCH
TOMSK STATE UNIVERSITY (NR TSU)
SAE "Smart Materials and Technologies" Institute

APPROVED BY
Head of the Main Education
Program, Doctor of Physical and
Mathematical Sciences, Professor


I. A. Kurzina

signature
« 6 » 06 2026

BACHELOR'S THESIS

ANALYSIS OF CHROMOSOMAL ABNORMALITIES IN SPONTANEOUS ABORTION
TISSUES USING MOLECULAR KARYOTYPING (NGS).

Specialization 27.03.05 Innovatics
Autonomous educational program "Tomsk International Science Program"

Kankanam Pathirana Manith Banula

Bachelor's Thesis
Doctor of Physical and
Mathematical Sciences, Professor,
Head of "Smart Materials and
Technologies Institute"


I. A. Kurzina

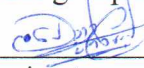
signature
« 6 » 06 2026

Scientific Consultant
Doctor of Biological Sciences,
Deputy Director for Science and
Technology Development of Tomsk
NRMC


S. A. Vasilyev

signature
« 5 » 06 2026

Written by
Student group № 862201


K.P.M. Banula

signature
« 6 » 06 2026

Tomsk – 2026

Ministry of Science and Higher Education of the Russian Federation
NATIONAL RESEARCH
TOMSK STATE UNIVERSITY (NR TSU)
SAE "Smart Materials and Technologies" Institute

APPROVED BY
Head of the Main Education
Program, Doctor of Physical and
Mathematical Sciences, Professor

I. A. Kurzina

« 12 » 01 2026

THE TASK

To perform the final bachelor's qualification work for the student

Kankanam Pathirana Manith Banula

(student full name)

in the field of study 27.03.05 "Innovatics", Tomsk International Science Program

1. The topic of the bachelor's thesis

ANALYSIS OF CHROMOSOMAL ABNORMALITIES IN SPONTANEOUS ABORTION
TISSUES USING MOLECULAR KARYOTYPING (NGS)

2. The deadline for student to submit the completed work:

a) to the Program Education Office

June, 11, 2026

b) to the State Examination Committee

June, 11, 2026

3. Initial data for the work:

Objects of the study – Materials of spontaneous abortuses

Scientific/ applied – Primarily scientific with applied elements

The aim of the study – Evaluation of the method for various chromosomal abnormalities
Statistical analysis of chromosomal pathologies in abortive tissues

Objectives:

Perform whole-genome sequencing with ultra-low coverage for molecular karyotyping of
spontaneous abortuses

Methods of research:

Ultra-low coverage whole-genome sequencing (ulcWGS), bioinformatic linear read-depth
profiling, Fluorescence In Situ Hybridization (FISH) and Real-Time PCR (qPCR)

The organization or industry that the work is being carried out on, –

Laboratory of Instrumental Genomics, Research Institute of Medical Genetics, Tomsk
National Research Medical Center of Russian Academy of Sciences

4. Stages of work

Deadlines

4.1 Topic selection and literature review	15,04,2024
4.2 Methodology design	01,05,2024
4.3 Experimentation and data collection	30,07,2024
4.4 Results analysis and evaluation	01,03,2026
4.5 Writing the thesis	20,15,2026
4.6 Pre-defense	28,05,2026
4.7 Defense	11,06,2026

Bachelor's Thesis Supervisor

Doctor of Physical and Mathematical
Sciences, Professor, Director of "Smart
Materials and Technologies" Institute

(degree, position)



(signature)

/ I.A. Kurzina

(Name, Surname)

Bachelor's Thesis Scientific Consultant

Doctor of Biological Sciences,
Deputy Director for Science and Technology
Development of Tomsk NRMC

(degree, position)

(signature)

/ S.A. Vasilyev

(Name, Surname)

The task was accepted by student

(student)

(signature)

/ K.P.M.Banula

(Name, Surname)

Заключение о форме размещения ВКР в ЭБ ТГУ

В соответствии с п. 3.2. «Регламента размещения текстов выпускных квалификационных работ в Электронной библиотеке Научной библиотеки ТГУ» (Приказ № 450/ОД от 17.04.2024) выпускная квалификационная работа Канканам Патиранге Манит Банула на тему «Analysis of chromosomal abnormalities in spontaneous abortion tissues using molecular karyotyping (NGS).» размещается в ЭБ ТГУ с изъятием раздела 3 (стр. 46–63) по решению правообладателя. Руководитель ООП Курзина И.А.

Руководитель ООП



Курзина И.А.

АННОТАЦИЯ

ведение: Самопроизвольный аборт (выкидыш) осложняет примерно 20% клинически распознаваемых беременностей, причем до двух третей случаев обусловлены хромосомными аномалиями эмбриона. Стандартное цитогенетическое G-окрашивание (кариотипирование) часто затруднено из-за высокой частоты неудач культивирования клеток, вызванных некрозом тканей. Полногеномное секвенирование со сверхнизким покрытием (ulcWGS) представляет собой быструю молекулярную альтернативу, не требующую культивирования клеток, однако имеет математическое «слепое пятно» в отношении выявления полногеномных полиплоидий, что требует его интеграции с ортогональными методами диагностики.

Цель: Выполнить полногеномное секвенирование со сверхнизким покрытием для молекулярного кариотипирования образцов тканей при самопроизвольных абортах.

Методология: В данном исследовании оценивался интегрированный рабочий процесс, сочетающий ulcWGS с интерфазной флуоресцентной гибридизацией *in situ* (FISH) и количественной ПЦР (кПЦР), для анализа 206 образцов тканей, полученных при самопроизвольных абортах. Образцы ворсин хориона и внезародышевой мезодермы тщательно изолировали для минимизации контаминации материнскими клетками. Выделенную ДНК подвергали тагментации с помощью транспозазы Tn5 и высокопроизводительному секвенированию (2×). Секвенированные риды картировали на сборку GRCh38/hg38 и анализировали по статистическим бинам размером 50 килобаз для расчета кариотипов. Интерфазный метод FISH применяли для разрешения полиплоидных конфигураций.

Результаты: Метод ulcWGS продемонстрировал аналитическую чувствительность 88,5% после верификации результатов с помощью FISH и кПЦР. Итоговое подтвержденное распределение кариотипов в когорте составило: 48,1% нормальных диплоидных профилей (n=99), 31,6% анеуплоидий (n=65), 9,8% субхромосомных вариаций числа копий ДНК (CNV) (n=20), 6,3% полиплоидий (n=13), 2,9% сочетаний анеуплоидий и CNV (n=6), 1,0% сочетаний анеуплоидий и полиплоидий (n=2) и 0.5% сочетаний CNV и полиплоидий (n=1). Трисомии 22 и 16 оказались наиболее частыми аутосомными гиперплоидиями, в то время как моносомия X была доминирующей анеуплоидией по половым хромосомам. Сложные мозаичные линии и вариации числа копий (CNV) размером менее 1 миллиона пар нуклеотидов также были точно разрешены с помощью ulcWGS.

Заключение: Метод секвенирования со сверхнизким покрытием, валидированный относительно референсных методов, доказал свою высокую эффективность (88,5%) для клинического скрининга анеуплоидий в когортах пациенток с выкидышами при отсутствии альтернативных вариантов диагностики. Кроме того, практическое применение позволило достичь субмикроскопического разрешения в среднем диапазоне, обеспечив диагностическую значимость субхромосомных вариаций на уровне 4,71%.

Ключевые слова: Самопроизвольный аборт (выкидыш); секвенирование следующего поколения (NGS); полногеномное секвенирование со сверхнизким покрытием (ulcWGS); флуоресцентная гибридизация *in situ* (FISH); хромосомные аномалии.

ABSTRACT

Background: Spontaneous abortion affects approximately 20% of clinically recognized pregnancies, with up to two-thirds driven by embryonic chromosomal anomalies. Standard cytogenetic G-banding is frequently obstructed by high culture failure rates from tissue necrosis. Ultra-low coverage whole-genome sequencing (ulcWGS) provides a fast, culture-free molecular alternative but possesses a mathematical blind spot for detect whole-genome polyploidies, requiring integration with orthogonal diagnostic tools.

Objective: Perform whole-genome sequencing with ultra-low coverage for molecular karyotyping of spontaneous abortuses.

Methodology: This study evaluated an integrated workflow combining ulcWGS with interphase Fluorescence *In Situ* Hybridization (FISH) and quantitative PCR (qPCR) to analyze 206 spontaneous abortuses tissue samples. Chorionic villi and extraembryonic mesoderm tissues were scrupulously isolated to minimize maternal cell contamination. Extracted DNA underwent Tn5 tagmentation and high-throughput sequencing (2×). Sequenced reads were mapped to the GRCh38/ hg38 assembly and analyzed across 50-kilobase statistical bins to calculate the karyotypes. Interphase FISH was deployed to resolve polyploid configurations.

Results: The ulcWGS achieved an analytical sensitivity of 88.5% after verification of the results through FISH and qPCR. The final validated karyotype distribution across the cohort comprised 48.1% normal diploid profiles (n=99), 31.6% aneuploidies (n=65), 9.8% subchromosomal CNVs (n=20), 6.3% polyploidies (n=13), 2.9% aneuploidies combined with CNVs (n=6), 1.0% aneuploidies combined with polyploidies (n=2) and 0.5% CNVs combined with polyploidies (n=1). Trisomies 22 and 16 were the most frequent autosomal hyperploidies, while monosomy X was the dominant sex-chromosome aneuploidy. Complex mosaic lineages and copy number variations (CNV) below 1 million base pairs were also accurately resolved through ulcWGS.

Conclusion: Validated against reference methods, ultra-low-coverage sequencing proved highly effective (88.5%) for clinical aneuploidy screening in miscarriage cohorts where alternative diagnostic options are absent. Furthermore, the practical application achieved average sub-microscopic resolution, establishing a 4.71% diagnostic yield for subchromosomal variations.

Keywords: Spontaneous Abortion (SA); Next-Generation Sequencing; Ultra-Low Coverage Whole-Genome Sequencing (ucl-WGS); Fluorescence *In Situ* Hybridization (FISH); Chromosomal Abnormalities

CONTENTS

GLOSSARY	6
INTRODUCTION	8
1. LITERATURE REVIEW	10
1.1. Pregnancy Loses	10
1.2. Ontogeny of Early Embryonic Tissue and Their Diagnostic Significance	12
1.3. Chromosomal Abnormalities	16
1.4. Aneuploidy	17
1.5. Polyploidy	21
1.6. Copy Number Variations	24
1.7. Mosaic Embryos	25
1.8. Standard Karyotyping Method and Its Limitations	26
1.9. Karyotyping Using Next Generation Sequencing Methods	27
1. METHODOLOGY	32
2.1. Research Approach	32
2.2. Sample Collection	33
2.3. Procedure	34
2.3.1. DNA Extraction	34
2.3.2. Tissue Extraction and Suspension Preparation	34
2.3.3. Real-Time PCR	35
2.3.4. FISH Sample Preparation	41
2.3.5. Ultra-Low Coverage Whole Genome Sequencing	44
2.3.6. Bioinformatic Processing	45
3. RESULTS	46
4. DISCUSSION	64
5. CONCLUSION	67
REFERENCES	

GLOSSARY

Aneuploidy - A state of chromosomal abnormality characterized by an atypical chromosome count that is not an exact multiple of the haploid set (e.g., $2n+1$ or $2n-1$), typically arising from meiotic nondisjunction during gametogenesis. This condition primarily manifests either as monosomy, the absolute loss of a single chromosome from the normal diploid complement ($2n-1$) which is almost universally cell-lethal during early preimplantation windows, or as trisomy, a highly non-random numeric aberration defined by the presence of three homologous copies of a specific chromosome ($2n+1$) that acts as a leading driver of early embryonic arrest.

Chorionic Villi - The vascular, finger-like projections of the chorion that invade the maternal endometrium during early placentation. Sharing a common zygotic progenitor with the embryo proper, they serve as the primary source of fetal genetic material for prenatal diagnostics and products of conception analysis.

Confined Placental Mosaicism (CPM) - A genetic anomaly where a specific chromosomal mutation or numerical aberration is strictly localized to the extraembryonic tissue (placental lineages), while the embryo proper develops with a normal diploid complement.

Copy Number Variation (CNV) - A structural genomic variant involving the deletion, duplication, or translocation of a DNA segment ranging in size from single kilobases to multiple megabases, shifting the baseline gene dosage relative to a standard reference genome.

Extraembryonic Mesoderm - The specific tissue layer lining the interior of the blastocyst cavity and structural components of the chorionic villi. It develops synchronously with early embryonic lineages, making it a high-yield, low-contamination source of fetal DNA.

Fluorescence *In Situ* Hybridization (FISH) - A targeted, slide-based cytogenetic technique that utilizes sequence-specific fluorescent probes to bind complementary chromosomal regions. In interphase cells, it enables the direct visualization and counting of specific centromeres or loci to determine ploidy states.

G-Banded Karyotyping (Giemsa Banding) - The classical cytogenetic benchmark for structural and numerical chromosome evaluation. It relies on chemical cell arrest during mitosis, slide fixation, and proteolytic enzyme treatment (Giemsa staining) to produce macroscopic horizontal bands along chromosomes visible at an optical threshold of ≥ 5 megabases.

Low-Coverage Whole-Genome Sequencing (lcWGS) - A high-throughput sequencing approach operating at low depths of coverage ($0.1\times$ to $1.0\times$) It utilizes specialized statistical filtering and alignment algorithms to map copy number profiles across the genome without requiring massive sequencing depth.

Ultra-Low Coverage Whole-Genome Sequencing (ulcWGS) - A high-throughput molecular karyotyping configuration operating at shallow depths ($0.01\times$ to $0.1\times$). It maps millions of untargeted short sequence reads across reference genomes to digitally isolate aneuploidies and copy number variations without relying on live cell lines.

Meiotic Nondisjunction - An error in gamete formation during which homologous chromosomes fail to separate cleanly in meiosis I, or sister chromatids fail to segregate in meiosis II. This cellular failure results in gametes with an abnormal excess or deficit of chromosomes.

Nullisomy - A lethal genetic state involving the complete, uniform loss of both homologous members of a chromosome pair ($2n-2$).

Polyploidy - A numeric chromosomal aberration where an organism's genome contains whole additional sets of the entire haploid complement, resulting in states such as triploidy ($3n=69$) or tetraploidy ($4n=92$).

Products of Conception (POC) - The biological material expelled or surgically evacuated from the uterus following an early pregnancy loss, encompassing the embryo, fetal membranes, chorionic villi, and associated extraembryonic structures.

Quantitative Real-Time PCR (qPCR) - A molecular laboratory method that tracks the amplification of a targeted DNA molecule in real time via fluorescent dyes. It is used as an orthogonal tool to confirm specific copy number states or clear up technical sequencing boundaries.

Recurrent Pregnancy Loss (RPL) - A clinical condition defined by international reproductive guidelines as the consecutive, spontaneous loss of two or more confirmed pregnancies.

Spontaneous Abortion (SA) - Defined by the World Health Organization as the natural, non-induced expulsion or extraction of an embryo or fetus weighing less than 500 grams, typically corresponding to a developmental window prior to the 22nd week of gestation.

Sporadic Pregnancy Loss (SPL) - An isolated, non-consecutive instance of early pregnancy failure, predominantly driven by spontaneous, non-inheritable de novo meiotic errors.

Statistical Binning - A bioinformatic technique where the linear reference assembly of a human genome is partitioned into discrete, contiguous windows of fixed size (e.g., 50 kilobases). Sequenced read counts are dynamically pooled into these bins to calculate local read-density shifts and deduce copy number changes.

Structural Rearrangement - A broad class of chromosomal aberrations that modify genomic architecture without necessarily altering whole-chromosome numbers. These encompass insertions, deletions, duplications, inversions, and translocations.

Tagmentation - An enzymatic molecular library preparation workflow that utilizes an engineered Transposase (typically Tn5) to simultaneously cleave genomic DNA into fragments and ligate sequencing adaptors in a single reaction step.

INTRODUCTION

Spontaneous abortions (SA) is a pathological condition that occurs in 20% of all clinically diagnosed pregnancies, and in 80% of cases, occurs in the first trimester [1]. Up to two-thirds of SA may be due to abnormalities in the embryo's chromosomal complement [2]. Chromosomal abnormalities identified in early embryonic tissue manifest as either numerical alterations or complex structural rearrangements. While the numerical instability involves deviations from the standard human diploid complement, subclassified into aneuploidies and polyploidies and structural rearrangements, encompass deletions, duplications, inversions, and subchromosomal copy number variations (CNVs).

Determining the fetal karyotype in SA is necessary to understand the cause of the abortion, as well as to overcome inherited chromosomal structural rearrangements that are significant when planning a subsequent pregnancy. The gold standard of karyotyping SA is G-staining of chromosomes, first proposed by Marina Seabright in latter half of the twentieth century. This method, which consists of the stages of cell selection, cell culturing, cell division arrest, followed by fixation, staining, and microscopy, enables the detection of aneuploidies, polyploidies and chromosomal mutations which are larger than 5 million bp. However, in the absence of viable cells or insufficient mitotic activity, the method is inefficient. Although methods based on massive parallel sequencing, which do not require *in vitro* cell culturing, may offer an alternative.

A prominent approach is low-coverage whole genome sequencing (lcWGS) which uses specialized filtering methods and statistical analysis to compensate for low coverage, lcWGS allows for the identification of the presence or absence of additional chromosomal material ranging in size from hundreds of kilobases to millions of basepairs [3]. This resolution allows for the use of this approach for diagnosing pathologies associated with copy number variations (CNVs), detecting aneuploidies, and determining karyotype changes in cultured cell lines [4]. Despite the limitations of this method, related to the complexity of bioinformatics processing, limited statistical power per individual, and interpretation of results, it has considerable advantages, especially in the context of large-scale genomic and population studies, which is already being used for diagnostic purposes.

However, approaches based on lower coverage, such as ulcWGS, can be used for targeted aneuploidy and CNV diagnostics. This method, while having slightly greater limitations, retains the advantages of the previous method and is more cost-effective, bringing it closer to screening methods. To overcome these limitations and comprehensively evaluate the sensitivity and effectiveness of ulcWGS, this research utilizes cross-referencing with fluorescence *in situ* hybridization (FISH) and real-time PCR (qPCR) methods. Specifically, FISH was employed to detect polyploid samples, which cannot be accurately identified by ulcWGS.

Karyotype determination is an important and necessary step in investigating the causes of miscarriage. In this study, products of conception were utilized to isolate chorionic tissue specifically targeting the chorionic villi and their underlying extraembryonic mesoderm as these structures provide high-yield fetal DNA while minimizing the risk of maternal cell contamination. This study aims evaluation of the method for various chromosomal abnormalities and statistical analysis of chromosomal pathologies in abortive tissues. While describing the use of ultra-low

coverage whole-genome sequencing (ulcWGS) for the molecular karyotyping of aborted material where standard cytogenetic methods were unfeasible, evaluating the frequency and structural landscape of chromosomal abnormalities in spontaneous miscarriages alongside several complex cases.

1. LITERATURE REVIEW

1.1 Pregnancy Losses

Abortions are defined as the expulsion or extraction of a fetus (embryo) weighing less than 500 g equivalent to approximately 22 weeks gestation, according to WHO [5]. Each year 23 million abortions occur worldwide or 44 abortions per minute, which is 15.3% from the recognized pregnancy losses [6]. Worldwide 10.8% of women have experienced at least one abortion and respectively 1.9% and 0.7% have experienced two or three abortions [6]. Abortions can be categorized into two as spontaneous or induced. Spontaneous abortions are also named as miscarriages which is the natural death of the fetus in the womb. While the induced abortions are the deliberate termination of a pregnancy through medical or surgical intervention. Miscarriages often historically dismissed as a "natural" mechanism for selecting against non-viable embryos, recent scientific advancements have revealed a complex web of genetic, immunological, and environmental factors that dictate pregnancy viability.

As spontaneous abortions is known as one of the most common which woman may encounter during the pregnancy which mostly predominates from hormonal problems maternal disorders, and chromosomal abnormalities [7]. Also other risk factors such as lifestyle, diet, smoking, and alcohol are also other important factors effect on spontaneous abortion [8]. Globally, the incidence of SA varies significantly, with the North Africa and Middle East (NAME) region recording rates as high as 2,409 per 100,000 women, well above the global average of approximately 2,172 [9].

The early pregnancy losses or the early miscarriages hold the most number miscarriages, approximately 80% of the miscarriages [10]. The early pregnancy losses are defined as a nonviable, intrauterine pregnancy with either an empty gestational sac or a gestational sac containing an embryo or fetus without fetal heart activity within the first 12 6/7 weeks of gestation [11]. The primary etiology of pregnancy loss, particularly in the first trimester, is chromosomal abnormality. Recent large-scale genomic studies have identified specific genetic variants that predispose individuals to miscarriage. Out of them 50% of all cases are due to the chromosomal abnormalities of the fetus [10]. These losses are frequently driven by errors in meiosis the specialized cell division process in germ cells. Variants associated with meiotic spindle assembly and chromosomal segregation are now recognized as critical markers for assessing miscarriage risk in both natural conceptions and *In Vitro* Fertilization (IVF).

These chromosomal abnormalities appear clinically as either isolated events or repeated reproductive failures. Clinically, early pregnancy terminations classified as sporadic pregnancy loss (SPL) and recurrent pregnancy loss (RPL) [12]. Sporadic pregnancy losses are defined as an isolated, non-consecutive miscarriage event, which occurs randomly in recognized pregnancies and predominantly caused by de novo parental meiotic errors [12]. Conversely, recurrent pregnancy losses are defined by international guidelines as the consecutive loss of two or more pregnancies [13]. While sporadic pregnancy losses are heavily dominated by numerical chromosomal errors such as aneuploidies and polyploidies. Recurrent pregnancy loss populations often present a distinct distribution of recurrent genetic, structural, or parental factors [12, 13].

Further it was identified that the most common risk factor among the women who have experienced the early pregnancy losses are advanced maternal age and a prior early pregnancy loss. 9-17% frequency of early pregnancy losses are observed among the women ages 20 to 30 and this rate increases dramatically from 20% at the age 35 ,40% at the age 40 and 80% at the age 45 [10].

Immunological mechanisms and maternal fetal tolerance also have a considerable effect on the pregnancy losses. As a successful pregnancy requires a sophisticated state of "immune privilege" where the maternal immune system recognizes but does not reject the semi-allogeneic fetus. Recurrent Pregnancy Loss (RPL) is understood as a failure of this tolerance. Contemporary research underscores the roles of regulatory T cells (T regs), natural killer cells, cytokine dysregulation, and disruptions in maternal–fetal immune tolerance [14]. As the reduction in Treg populations leads to an inflammatory environment hostile to the embryo, dysregulation in uterine natural killer cells can impair placental remodelling and vascularization and the cytokine dysregulation can increased Th1/Th2 ratio and a decreased Treg/Th17 ratio are hallmark immunological signatures of RPL [15]. Furthermore, recent evidence suggests that shared HLA alleles between partners and the presence of specific autoantibodies significantly increase the risk of SA [15].

The other factors which can affect the pregnancy losses are the impact of the environmental and lifestyle stressors and the socio-economic disparities. The impact of the environment on pregnancy maintenance is documented in the context of climate change and industrial pollution. Recent case-crossover studies have identified a "U-shaped association" between outdoor air temperature and the risk of spontaneous abortion, with extreme heat days (>95th percentile) increasing risk by 20% [16]. Additionally, lifestyle factors such as obesity, smoking, and high levels of psychoactive substance use create metabolic and endocrine disruptions that destabilize early pregnancy and occupational hazards as anesthetic gases, sterilization agents, and heavy metals (cadmium, lead, mercury) are proven causes of pregnancy losses.

Pregnancy loss is not distributed equally across populations, the socio-economic status (SES) remains a powerful predictor of pregnancy outcomes. Based on the research in UK and Republic of Ireland it was identified that the women from lower social classes or those facing unemployment have a 40% higher risk of experiencing stillbirth and neonatal mortality compared to those in the highest SES brackets [17]. These disparities are often attributed to "austerity," food insecurity, and a lack of access to specialized prenatal screening, reinforcing the need for systemic health intervention.

The clinical focus on the abortus often overshadows the profound psychological impact on the mother. A 2025 meta-analysis found that approximately 35.6% of women experience anxiety, 30.75% suffer from depression, and 33.66% report high levels of stress within six weeks following a miscarriage [18]. This "silent trauma" is often neglected in standard obstetric care, despite evidence that psychological distress can further exacerbate endocrine dysfunctions in subsequent pregnancies.

Overall, it emphasises the importance investigation of the underlying mechanisms of these losses, moving beyond the binary of "viable" versus "non-viable" to explore the nuanced pathophysiology of the spontaneous abortus. This study will focus on application of next generation sequencing for

the deeper understanding of the pregnancy losses and overcome bottlenecks of the prevailing diagnostic methods.

Overall, it emphasises the importance investigation of a population-level evaluation of genomic instability in spontaneous abortuses, specifically targeting aneuploidies, polyploidies, and subchromosomal CNVs. This study will focus on application of next generation sequencing for the deeper understanding of the pregnancy losses and overcome bottlenecks of the prevailing diagnostic methods.

1.2 Ontogeny of extraembryonic tissue and their diagnostic significance

1.2.1 Introduction

The transition from a single celled zygote to a complex multicellular organism involves the early segregation of cell lineages into embryonic and extraembryonic tissues. While the embryonic lineage focuses on the formation of the fetus, the extraembryonic tissues including the chorion, yolk sac, amnion, and allantois serve as the essential life-support system that mediates the maternal-fetal interface [19]. As the extraembryonic tissue emerges and differentiate simultaneously to the development of any embryonic precursors [19]. In clinical genetics and prenatal medicine, these tissues are not merely biological scaffolding, they are critical diagnostic windows. Because these structures share a common progenitor with the embryo, their cellular and molecular composition often mirrors the fetal genetic status and provide a unique opportunity for early intervention and screening [20]. This chapter focuses on the ontogeny of these tissues, elaborating on the chorionic villi, and evaluates their indispensable role in modern diagnostic methodologies.

1.2.2. Chorionic Villi

The chorionic villi represent the functional unit of the placenta, acting as the primary site for nutrient, gas, and waste exchange between the maternal and fetal circulations. Their development is a highly regulated process that begins shortly after implantation, transforming the simple trophoblastic shell into a complex, vascularized tree-like structure [21]. From a diagnostic perspective, the chorionic villi are the perfect fit for first-trimester prenatal diagnosis, as they allow for the extraction of fetal DNA at a significantly earlier stage than amniocentesis [22]. The maturation of the chorionic villi occurs in three distinct phases during the first weeks of gestation. This morphological progression is fundamental to the timing and success of diagnostic procedures, as the tissue harvested must contain sufficient cellular material for genetic analysis.

Primary villi begin to form at approximately the end of the second week of development (days 13–15). This stage is characterized by the rapid proliferation of cytotrophoblast cells which form solid, finger-like columns that project into the surrounding syncytiotrophoblast and the maternal blood-filled lacunae [21]. At this point, the villi consist purely of trophoblastic cells and lack an internal

mesenchymal core or blood vessels. While these structures are too immature for standard invasive diagnostic sampling, they represent the first stage of placental expansion and are the source of early trophoblastic signals detectable in maternal blood [20].

During the third week of gestation, the extraembryonic mesoderm splits in to somatic and splanchnic mesoderm and the splanchnic mesoderm a connective tissue layer begins to invade the center of the primary villi. Once this mesodermal core is established, the structures are classified as secondary chorionic villi [21]. This stage is a critical transition because the presence of mesodermal tissue provides the structural framework for the upcoming vascularization [23]. In diagnostic contexts, the mesenchymal cells within these secondary chorionic villi are often analyzed to identify mosaicism, as they can sometimes harbor different genetic lines than the overlying trophoblast [19].

By the end of the third week, the mesodermal cells within the secondary chorionic villi undergo vasculogenesis, differentiating into capillaries and blood cells. Once these secondary chorionic villi contain functional blood vessels, they are defined as tertiary chorionic villi [21]. These vessels eventually connect with the embryonic heart through the umbilical cord, establishing the fetoplacental circulation which provides the oxygen and nutrients for the fetus [21]. Tertiary villi are the primary target for Chorionic Villus Sampling (CVS), which is typically performed between the 10th and 13th weeks of gestation [22]. At this stage, the villi are highly branched, providing a large surface area from which clinicians can safely harvest tissue for comprehensive chromosomal and molecular testing [20].

The diagnostic ability of the chorionic villi stems from their genetic representativeness. Because both the embryo and the trophoblast originate from the same zygote, the genetic information contained within the villi is usually identical to that of the developing fetus [19]. This allows clinicians to use villous tissue to diagnose a wide range of clinical conditions, including chromosomal aneuploidies like Down Syndrome (Trisomy 21), Edwards Syndrome (Trisomy 18), and Patau Syndrome (Trisomy 13) [24]. Furthermore, molecular techniques such as chromosomal microarray (CMA) and whole-exome sequencing (WES) are increasingly applied to villous samples to detect microdeletions and single-gene disorders like Cystic Fibrosis or Thalassemia. One of the primary advantages of utilizing chorionic villi for diagnosis is the timing as the CVS can be performed as early as 10 weeks of gestation, whereas amniocentesis is rarely performed before 15 weeks [22]. This early window provides parents and healthcare providers with critical information during the first trimester, allowing for management options and reducing the psychological stress associated with late-gestation diagnosis [22].

However, the diagnostic use of this tissue requires a nuanced understanding of Confined Placental Mosaicism (CPM) which occurs when a chromosome anomaly exists in the placenta but not in fetus. In approximately 1–2% of cases, the chromosomal makeup of the placenta may differ from that of the fetus [25]. This occurs when a mutation or chromosomal error happens early in development but is restricted to the lineages forming the chorionic villi. While CVS is highly accurate but cases of mosaicism often require a follow-up amniocentesis to confirm whether the abnormality is also present in the fetus [19].

1.2.3. Yolk Sac

The human yolk sac (HYS) is an essential, albeit transient, extraembryonic organ that serves as the primary metabolic, hematopoietic, and endoderm-derived signaling center for the embryo during the first trimester. Historically regarded as a vestigial structure in placental mammals, recent breakthroughs in single-cell transcriptomics and high-resolution imaging have redefined the yolk sac as a multifunctional mediator that orchestrates embryonic survival before the definitive fetal liver and placenta assume these roles [26]. In clinical practice, the yolk sac is the first structure of the conceptus to be visualized via transvaginal ultrasonography, typically appearing between 4.5 and 5 weeks of gestation [27]. Its developmental trajectory, characterized by a complex transition from a primary to a secondary structure, provides a critical diagnostic window into the early gestational environment and the genetic integrity of the embryo. The ontogeny of the yolk sac begins during the second week of development, emerging from the hypoblast (primitive endoderm). The primary yolk sac, or exocoelomic cavity, forms around day 9 as cells from the hypoblast migrate and line the inner surface of the cytotrophoblast, forming Heuser's membrane [26]. This structure is ephemeral by day 12–13, the expansion of the extraembryonic mesoderm begins to compress the primary yolk sac, fragment into small vesicles which remain is the secondary yolk sac, a spherical, fluid-filled structure that becomes the functional interface between the embryo and the extraembryonic coelom [19].

Recent single-cell RNA sequencing (scRNA-seq) data has revealed that the secondary yolk sac is a complex multi-organ progenitor. During the fourth to seventh weeks of development, the yolk sac wall expresses transcripts traditionally associated with the liver, kidneys, and the bone marrow [26]. This functional interaction ensures that the embryo receives essential proteins and lipid transporters before the definitive fetal organs have differentiated. Furthermore, the yolk sac is the definitive site of primitive hematopoiesis and the mesodermal layer of the yolk sac wall develops "blood islands". Where hemogenic endothelial cells differentiate into the embryo's first erythroid and myeloid cells [26]. These cells migrate via the vitelline circulation to the fetal liver, which subsequently takes over hematopoietic functions. Additionally, the yolk sac serves as the essential reservoir for primordial germ cells (PGCs) which will eventually form oocytes or spermatozoa, are first identifiable in the yolk sac wall before they undergo a strictly orchestrated migration along the dorsal mesentery of the hindgut to the gonadal ridges [19].

In modern obstetric diagnostics, the yolk sac is recognized as a "sentinel" structure and its early appearance confirms an intrauterine pregnancy (IUP), distinguishing it from a pseudo-gestational sac sometimes seen in ectopic pregnancies [27]. Peer-reviewed studies have established rigorous morphological and biometric criteria for assessing yolk sac health to predict the first trimester pregnancy outcome. A normal secondary yolk sac typically has its diameter ranging from 3 mm to 5 mm between the 5th and 10th weeks. A yolk sac diameter exceeding 6 mm, less than 3 mm, irregular shape or having degenerative changes often associate with abnormal pregnancy outcomes [28].

1.2.4. Extraembryonic Mesoderm

The extraembryonic mesoderm (ExM) is a critical tissue layer that emerges during the second week of embryonic development, shortly after the formation of the primary yolk sac. It occupies the anatomical space between the primary yolk sac's exocoelomic membrane and the overlying cytotrophoblast, eventually expanding to form the lining of the chorionic cavity [23]. While historically described as a mere supportive or filling tissue, modern molecular biological findings has redefined the ExM as a dynamic regulator of the early gestational environment. Which is the definitive progenitor of the placental stroma and the fetal vasculature, making it of profound diagnostic significance for researchers investigating the origins of chromosomal mosaicism and the physiological mechanisms of cell-free DNA release [19].

The origin of the extraembryonic mesoderm in humans has been a subject of intense scientific debate, until recently clarified by scRNA-seq and stem cell-derived embryo models. Unlike the embryonic mesoderm, which forms during gastrulation via the primitive streak in the third week of development, the ExM specified significantly earlier likely derived from the primitive endoderm (hypoblast) or a unique early epiblast population [23, 29]. This distinct developmental origin is a critical diagnostic fact that implies ExM and its derivatives, such as the villous stroma, possess a different epigenetic signature and potentially a different chromosomal makeup than the embryo proper [21].

As the ExM matures, it undergoes a complex morphogenetic transformation, differentiating into two distinct functional layers as somatic and splanchnic mesoderm. The extraembryonic somatic mesoderm lines the inside of the trophoblast and forms the internal mesenchymal core of the chorionic villi. This layer provides the structural framework for the placenta and houses the developing fetal capillary networks [21]. Conversely, the extraembryonic splanchnic mesoderm surrounds the yolk sac, playing a vital role in establishing the vitelline circulation [23]. The expansion of these layers creates the extraembryonic coelom, or chorionic cavity. In early diagnostic imaging, the proper expansion and fluid pressure of this cavity are key indicators of normal fluid dynamics and metabolic exchange between the maternal environment and the developing conceptus [19].

In the field of prenatal genetics, the ExM is central to the interpretation of Confined Placental Mosaicism (CPM). This phenomenon occurs when a chromosomal abnormality is detected in the placenta but is absent in the fetus, or inversely. When clinicians perform Chorionic Villus Sampling, they typically analyze two distinct cell populations from the harvested tissue as the rapidly dividing trophoblasts and the mesenchymal cells of the ExM core. While the diagnostic significance arises when these two populations exhibit discordant genetic results [21]. Peer-reviewed research published between 2023 and 2025 emphasizes that ExM derived mesenchymal cells are often a more stable and high-fidelity indicator of the true fetal genetic status than the outer trophoblast layer. This is attributed to the fact that the ExM and the embryo share a closer developmental timeline than the early-segregating trophoblast [25].

If an abnormality is found in the trophoblast but not in the ExM derived mesenchyme (Type 1 CPM), the prognosis for the fetus is generally favorable. However, abnormalities detected within the ExM core (Type 2 or Type 3 CPM) are much more highly predictive of fetal involvement, intrauterine growth restriction (IUGR), and potential pregnancy complications, necessitating

follow-up amniocentesis for definitive confirmation [21, 25]. One of the most significant advancements in modern prenatal diagnostics is the refinement of Non-Invasive Prenatal Screening (NIPS), which analyzes fetal DNA fragments in maternal blood. While it is widely understood that NIPS measures fetal DNA, recent molecular studies have localized the source of this DNA primarily to the apoptosis of cells within the extraembryonic tissues where ExM contributes significantly to this pool of genetic material [25].

The transition from viewing extraembryonic tissues as peripheral developmental adjuncts to recognizing them as primary diagnostic reservoirs represents a fundamental shift in the landscape of prenatal medicine. As established throughout this chapter, the chorionic villi, yolk sac and extraembryonic mesoderm each provide a specific, temporal, and complementary window into the genetic and physiological health of the developing embryo. The clinical shift toward precision prenatal medicine relies on the synergy between the physical assessment of these membranes and the molecular analysis of the cells and fluid they contain.

1.3 Chromosomal Abnormalities

Chromosomal abnormalities serve as a hallmark of genomic instability and are fundamentally categorized into two major groups as numerical and structural alterations [30]. While numerical abnormalities involve changes in the total number of chromosomes leading to conditions such as aneuploidy or polyploidy. Structural abnormalities arise from DNA damage and illegitimate repair mechanisms that alter the physical arrangement or composition of the chromosome itself [31]. These structural rearrangements include deletions, duplications, translocations, inversions, and more complex forms like ring chromosomes or isochromosomes [30]. Although subsequent chapters will provide an exhaustive examination of aneuploidy, polyploidy, and CNV, this section establishes the broader landscape of chromosomal aberrations with a specific emphasis on structural rearrangements that not defined solely by simple gain or loss of genomic dosage.

Structural chromosomal abnormalities are primarily the consequence of DNA double-strand breaks (DSBs) followed by erroneous repair processes, such as non-homologous end joining (NHEJ) or non-allelic homologous recombination (NAHR) [32]. Among these, chromosomal translocations represent one of the most clinically significant categories. Translocations occur when broken ends from non-homologous chromosomes are illegitimately joined, resulting in the exchange or transfer of genetic material which are classified into reciprocal translocations, where segments are exchanged between two chromosomes, and non-reciprocal translocations, where a segment is transferred unidirectionally [32]. Furthermore, these chromosomal translocations can categorize as “balanced”, which means there is no net loss or gain of genetic material, or “unbalanced,” where the rearrangement leads to genomic imbalance [32]. While balanced translocation carriers are often phenotypically normal, they face a heightened risk of producing offspring with derivative chromosomes that lead to severe developmental disorders or infertility [33]. In somatic cells, translocations are frequently linked to tumorigenesis, often by creating fusion genes or placing proto-oncogenes under the control of highly active promoters [32].

Chromosomal inversions represent another critical class of structural genomic rearrangement where a segment of a chromosome is detached, flipped 180 degrees, and reinserted [34]. Inversions

are typically "balanced" and do not result in the immediate loss of genetic material. However, they can have profound biological impacts by disrupting genes at the breakpoints or altering the expression of nearby genes through position effects [33]. Furthermore, inversions act as suppressors of recombination in heterozygotes, which can lead to the accumulation of divergent allelic sets and play a significant role in evolutionary adaptation and speciation [34]. Recent studies have also highlighted that nearly 38% of cytogenetically visible inversions are actually formed by replication-based mechanisms, which can lead to cryptic copy-number variations at the breakpoints, potentially contributing to unexpected clinical phenotypes [33].

Beyond simple rearrangements, more complex structural abnormalities such as ring chromosomes and isochromosomes illustrate the extreme plasticity of the genome. Ring chromosomes generally form when breaks occur at both ends of a chromosome, and the sticky ends fuse to form a circular structure [35]. These are often unstable during cell division, leading to further aberrations like dicentric chromosomes or interlocking rings, which can result in significant growth deficiencies and neurological conditions such as epilepsy [35]. Isochromosomes, occur when a chromosome divides horizontally rather than vertically during meiosis or mitosis, resulting in a chromosome with two identical arms and the loss of the other arm [36]. Such aberrations are frequently observed in specific malignancies and are critical biomarkers for disease progression and risk stratification [36]. Collectively, these structural variations, alongside the numerical changes discussed later, constitute the complex genomic signature that influences both human development and the pathology of diseases like cancer.

1.4 Aneuploidy

1.4.1 Introduction to Aneuploidy and Chromosomal Dynamics

Aneuploidy is defined as an abnormal number of chromosomes that deviates from the standard haploid set complement of 23 which involves the gain or loss of specific individual chromosomes, leading to a state of genomic imbalance [37]. Aneuploidy represents the most pervasive genetic contributor to human pregnancy loss, characterized within the complex landscape of molecular embryology, aneuploidy is identified as the primary etiology in approximately 50% to 60% of first-trimester spontaneous abortions [38]. The high prevalence of these aberrations in humans compared to other mammalian species suggests a unique vulnerability in human gametogenesis, particularly during the protracted period of oocyte maturation. Recent advancements in molecular cytogenetics have refined our understanding of how these numerical errors manifest and the specific thresholds of genomic tolerance that dictate whether an embryo proceeds to implantation or undergoes spontaneous termination [38, 39].

The Molecular etiology of aneuploidy is predominantly rooted in meiotic nondisjunction, failure of homologous chromosomes or sister chromatids to segregate properly during meiosis I or II. While nondisjunction can occur during paternal spermatogenesis, the vast majority of aneuploidies identified in spontaneous abortions are maternal in origin [40]. Current research identifies the

deterioration of the cohesin complex as a critical mechanical driver of this phenomenon. Cohesin proteins, which are established during fetal development and must remain functional for decades in the arrested oocyte, gradually degrade over time. This degradation leads to the premature separation of sister chromatids or the weakening of bivalent attachments, resulting in a random segregation during the first meiotic division [40]. Furthermore, the inefficiency of the Spindle Assembly Checkpoint (SAC) in human oocytes allows cells with misaligned chromosomes to bypass regulatory gates and complete meiosis, which produce aneuploid gametes that, upon fertilization, form zygotes with lethal genomic imbalances [41].

Classification of aneuploidy is determined by the specific nature of the numerical deviation, represented by the formula $2n \pm x$. These classifications are essential for clinical prognosis, as the survival potential of an embryo is strictly tied to the specific chromosome involved and the degree of gene dosage imbalance.

Monosomy ($2n - 1$): This condition involves the loss of a single chromosome. Autosomal monosomies (loss of any chromosome from 1 to 22) are almost universally lethal in the pre-implantation or very early post-implantation stages [42]. High-resolution studies indicate that autosomal monosomies represent a severe form of genomic instability that usually results in developmental arrest even before a pregnancy is clinically recognized [43]. The sole exception to this rule of lethality is Monosomy X ($45,X$), also known as Turner Syndrome while $45,X$ is a frequent finding in products of conception from spontaneous abortions. It remains the only monosomy capable of surviving to term, albeit in only about 1% of cases [39, 43] (Figure 1).

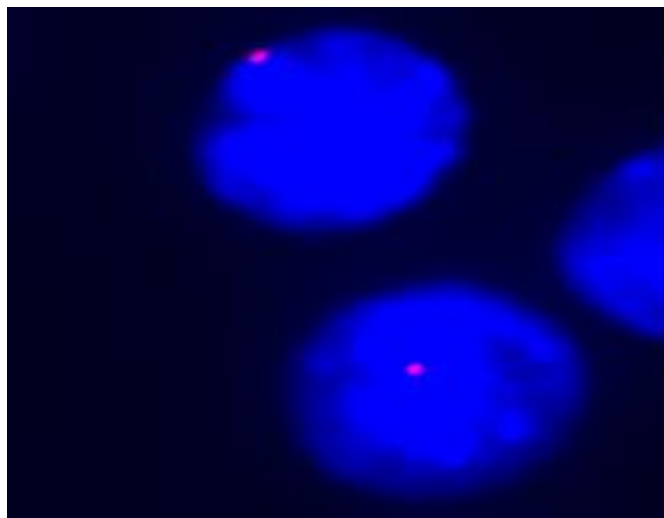


Figure 1: Interphase FISH analysis for chromosome X. Nuclei are counterstained blue with DAPI. The single red signal (probe specific to the X chromosome centromere with Tamra-dUTP) indicates a monosomy X constitution.

Trisomy ($2n + 1$): The gain of a single chromosome meaning there are three homologous copies of one chromosome, where in a human zygote that has 47 chromosomes [42]. Trisomies can involve any chromosome, but their frequency in spontaneous abortions is non-random. Trisomy 16 is the most prevalent, often referred to in literature as the "miscarriage chromosome" because it is never observed in live births and accounts for nearly 10% of all first-trimester losses [38]. Other frequent trisomies identified in abortuses include chromosomes 22, 21, 15, and 13. The

severity of the phenotypic impact is generally proportional to the size and gene density of the extra chromosome. However, specific regulatory regions on smaller chromosomes like 21 and 18 allow for rare instances of survival to birth, though often with significant developmental delays and multi-organ pathology [38, 40](Figure 2).

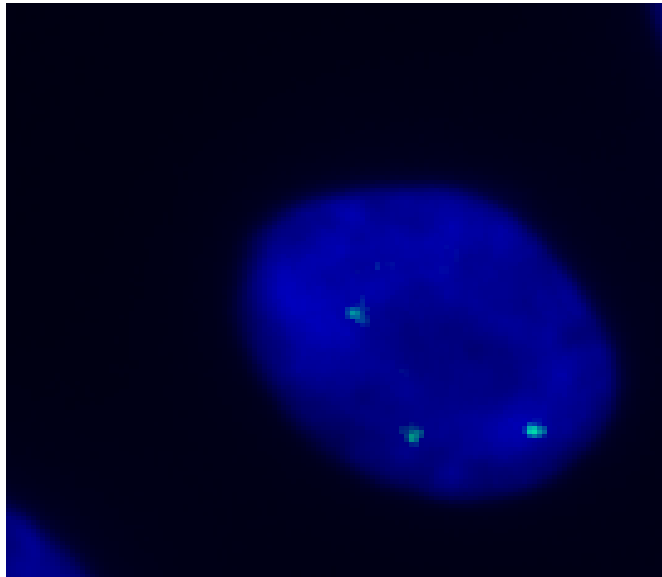


Figure 2: Interphase FISH analysis for chromosome 22. Nuclei are counterstained blue with DAPI. The three distinct green signals correspond to the DGCR (DiGeorge Critical Region) with Fluorescein-dUTP probe on chromosome 22q11.2, confirming a trisomy constitution.

Nullisomy ($2n - 2$) and Tetrasomy ($2n + 2$): Respectively, these represent the loss or gain of a homologous pair of chromosomes. Where Nullisomy is considered the most extreme form of aneuploidy and is incompatible with even the earliest stages of human embryonic cleavage and Tetrasomies are occasionally identified in specialized cytogenetic studies of spontaneous abortions, particularly involving the sex chromosomes (e.g., 48,XXXX) or small supernumerary marker chromosomes, but they are exceedingly rare compared to trisomies and monosomies [38, 39].

The biological consequence of these numerical shifts is a massive disruption of the cellular proteome as each extra or missing chromosome alters the expression levels of hundreds of genes simultaneously, leading to "proteotoxic stress." In the context of spontaneous abortion, this stress typically manifests as a failure in the development of the cytotrophoblast or severe malformation of the embryonic disc, leading to a missed abortion or a blighted ovum [43]. Understanding these fundamentals and their origins provides the necessary framework for the diagnostic and discussions in subsequent sections of this thesis.

1.4.2. Clinical Syndromes and Pregnancy Outcomes

The clinical expression of aneuploidy in human gestation is characterized by a high degree of selection against chromosomally imbalanced embryos. While the human genome possesses significant adaptability, the dosage imbalance resulting from an extra or missing chromosome typically disrupts the highly coordinated spatio-temporal expression of genes, leading to phenotypic failure [44]. In the context of spontaneous abortion, the "clinical syndrome" often refers to the specific physiological and morphological failures of the embryo or its supporting structures like placenta and yolk sac that leading to in pregnancy failure. Peer-reviewed literature from the last five years has increasingly utilized high-resolution genomic tools to categorize these syndromes not just by their karyotype, but by their specific lethality phases the gestational age at which development typically arrests [44].

As previously mentioned among the autosomal trisomies, Trisomy 16 holds a unique position. Unlike common viable aneuploidies, Trisomy 16 is considered universally lethal in its full form [38]. The clinical presentation of a Trisomy 16 pregnancy usually involves minimal development or an empty gestational sac, indicating that the developmental arrest occurs very early in the embryonic period, often before the formation of a discernible fetus [45]. Recent molecular studies suggest that the lethality of Trisomy 16 driven by the overexpression of genes located on chromosome 16. Which regulate early placental vascularization and trophoblast invasion. When these processes are disrupted, the embryo fails to establish a functional interface with the maternal blood supply, leading to rapid starvation and rejection by the uterine environment [45].

Other significant lethal trisomies include chromosomes 22 and 15. Trisomy 22 and 15 are frequently associated with severe growth restriction and are frequently observed in first-trimester miscarriages, though they are less common than Trisomy 16 [44]. Research indicates that the specific chromosomes duplicated correlates with the timing of pregnancy loss, for example, trisomies in chromosomes 15 and 22 significantly increased in earlier pregnancies and decline in later weeks, whereas trisomies that can survive until birth are more prevalent in advanced gestational weeks [44]. These findings highlight that the clinical outcome is not merely a result of having an extra chromosome, but specifically which chromosome is duplicated and how its unique genetic payload interacts with early developmental pathways.

Chromosomes 21, 18, and 13 are unique because they represent the subset of autosomal aneuploidies that can occasionally bypass the selection process of the first trimester and survive to birth [44]. However, even these trisomies are associated with high rates of spontaneous loss and the survivors leads for chromosomal disorders.

Down Syndrome occurred by the Trisomy 21, is the most common autosomal aneuploidy in live births, yet approximately 80% of its conceptions result in SA or stillbirth [46]. The clinical syndrome characterized by a specific set of morphological markers detectable through ultrasound, such as increased nuchal translucency and nasal bone hypoplasia [46]. Recent meta-analyses suggest that survival in Trisomy 21 potentially linked to early-life epigenetic compensation. Where certain embryos are able to regulate the over-expression of the "Down Syndrome Critical Region" (DSCR) on chromosome 21. Although trisomy 21 would be expected to cause 1.5-fold increase in expression of chromosome 21 genes, this is often not occurred. Due to epigenetic mechanisms,

such as compensatory hypermethylation, appear to regulate gene expression in the context of aneuploidy [47].

Trisomy 18 and Trisomy 13 leads rare autosomal aneuploidies, respectively to Edwards Syndrome and Patau Syndrome. These conditions characterized by multiple congenital anomalies and high neonatal and infant mortality. While historically perceived as mostly 'incompatible with life', a small but notable subgroup of patients survive for months to years [48]. Trisomy 18 normally presents with severe intrauterine growth restriction (IUGR), clenched fists with overlapping fingers, and complex cardiac defects. Trisomy 13 involves midline defects, including holoprosencephaly and cleft lip or palate [48].

Sex chromosome aneuploidies (SCAs) present a vast clinical profile compared to autosomal aneuploidies. Because of the biological mechanism of X-inactivation and the low gene density of the Y chromosome, the dosage imbalance in SCAs is generally less severe. However, Monosomy X which leads to Turner Syndrome remains a significant exception in the context of spontaneous abortion. Although 45, X is a common finding in live-born females, it represents for roughly 11% of all spontaneous abortions. This creates a "survival paradox" where 99% of 45,X embryos are lost during the first or second trimester [49]. Clinical syndrome in these cases is typically marked by severe cystic hygroma and cardiovascular malformations. Current research indicates that the 1% who survive likely to have cryptic mosaicism or specific epigenetic profile [49]. Klinefelter Syndrome (47, XXY) and Triple X Syndrome (47, XXX) are rarely identified aneuploidies in the tissues of spontaneous abortions. Their clinical syndromes are relatively mild, often going undiagnosed until puberty or adulthood. This suggests that the extra X chromosome is effectively silenced through X-inactivation, preventing the proteotoxic stress that causes the developmental arrest seen in autosomal trisomies [50].

1.5. Polyploidy

1.5.1 Introduction to Polyploidy and Chromosomal Set Imbalance.

The genomic integrity of the human zygote is a fundamental requirement for successful embryogenesis and subsequent clinical pregnancy. In reproductive genetics, genomic multiplicity or specifically polyploidy represents a distinct class of chromosomal aberration characterized by the presence of three or more complete sets of haploid chromosomes [42]. Polyploidy entails a whole-genome duplication event that fundamentally alters cellular stoichiometry and developmental potential. Despite advances in assisted reproductive technologies (ART), genomic instability remains a common feature of early human development, with de novo mutations and genomic imbalances estimated to occur in at least 70% of all conceptions [51]. Polyploidy, primarily manifesting as triploidy ($3n = 69$) or tetraploidy ($4n = 92$), is one of the most significant contributors that frequently leading to preimplantation embryo arrest or first-trimester spontaneous abortions [42]. In the germline and early embryo, polyploidy is almost universally lethal, leading to severe developmental disruption and eventual termination. The biological context of polyploidy presents a complex "Polyploidy Paradox" that is central to understanding its role in gestational

failure. In the germline and early embryo, polyploidy is almost universally lethal, leading to severe developmental disruption and eventual termination.

Conversely, polyploidy serves a functional role in specific human somatic tissues, a phenomenon known as developmentally programmed polyploidization. In tissues such as the liver, heart, and placenta, specifically within the trophoblast giant cells. Polyploidy is a strategy to increase cell size and biomass production without the structural disruption linked with cytokinesis [10]. This somatic polyploidy provides a higher tolerance to genomic stress or apoptotic signals and fitness benefits, including improved cell longevity [10]. Recent research shows that this transition to a polyploid state in somatic cells supports organ growth and tissue homeostasis, in regenerative environments specifically where tissues are less dependent on resident stem cells [10]. For example, polyploid cells are more efficient at forming protective barriers because larger cells require a lower density of cell-cell junctions [10].

In oncology, a similar paradox is observed in reversible polyploidy that may actually favour tumour progression and chemoresistance by providing an adaptive mechanism against the loss of heterozygosity [52]. Some researchers have suggested that polyploidy in cancer cells restores a tumour germ line in reproductive cycle, as a reciprocal relationship with the proliferative mitotic cycle to ensure survival under extreme genomic stress [52]. Understanding these epidemiological significances are critical for clinical diagnostics.

1.5.2 Cytogenetic Origins and Mechanisms of Formation

Polyploidy in human gestations is rooted in failures during gametogenesis or the earliest stages of zygotic division. The most prevalent form, triploidy (Figure 3), occurs in approximately 1% to 3% of all recognized conceptions which is a frequent finding in cytogenetic analyses of spontaneous abortions [53, 54]. The mechanical pathways leading to these states are classified into two broad categories. As paternal origin (diandry) and maternal origin (digyny), each influenced by specific cellular malfunction [55].

Paternal origin is the most frequent cause of human triploidy, which leads for dispermy approximately 66.4% of cases [55]. Which is the simultaneous fertilization of a single haploid oocyte by two separate haploid spermatozoa. In a normal fertilization event, the oocyte undergoes a zona reaction upon the entry of the first sperm. Which enzymatically alters the zona pellucida to prevent further sperm entry [54]. A failure in this allows a second sperm to penetrate, resulting in a zygote with one maternal and two paternal haploid sets. Paternal origin can also arise through the fertilization of a normal oocyte by a diploid sperm, resulting from a failure in paternal meiosis I or II [54].

Maternal origin or digyny, representing the remaining cases of triploid, occurs when a diploid oocyte is fertilized by a single haploid sperm [54]. The mechanical origin of a diploid oocyte, which can occur at two distinct stages. A failure of Meiosis I (MI) results in the retention of the first polar body and a failure of Meiosis II (MII) results in the retention of the second polar body [56]. The diandry triploids often abort between 10 and 20 weeks of gestation while the digyny

triploids generally abort before 8.5 weeks or less frequently survives late into the pregnancy as well developed fetuses [54].

Tetraploidy is a rarer and even more lethal genomic state found in spontaneous abortions which can be found approximately 2% in first-trimester [57]. Unlike triploidy, which is usually a fertilization error in tetraploidy most commonly occurs due to a failure in the first mitotic division [57]. This process, known as cytokinesis failure which occurs due to the DNA replicates during the S-phase, but the cell fails to divide into two daughter blastomeres. Which results zygote contains four copies of every chromosome (92,XXXX or 92,XXYY)[57] (Figure 4). Understanding these origins is not only important for classifying the etiology of pregnancy loss also for assessing the recurrence risk. Even though its being generally low, remains a point for investigation in the context of maternal genetic factors.

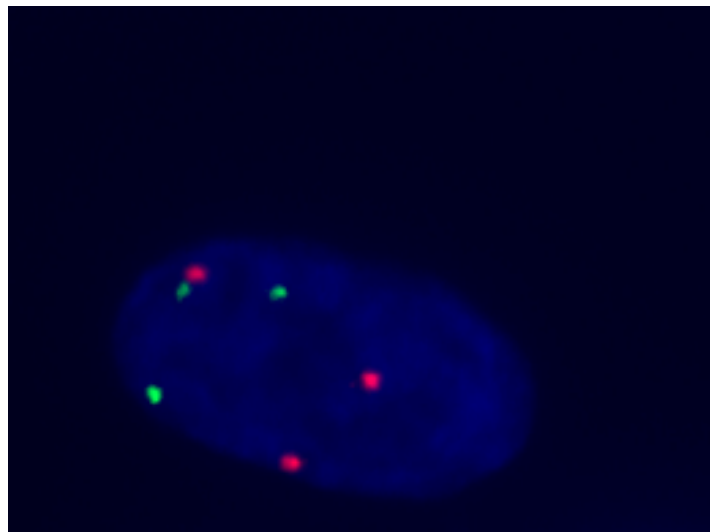


Figure 3: Interphase FISH analysis showing a triploid constitution. Nuclei are counterstained blue with DAPI. The three distinct red signals correspond to the chromosome X with Tamra-dUTP, and the three distinct green signals correspond to the chromosome 2 probe with Fluorescein-dUTP, confirming three complete sets of chromosomes (triploidy).

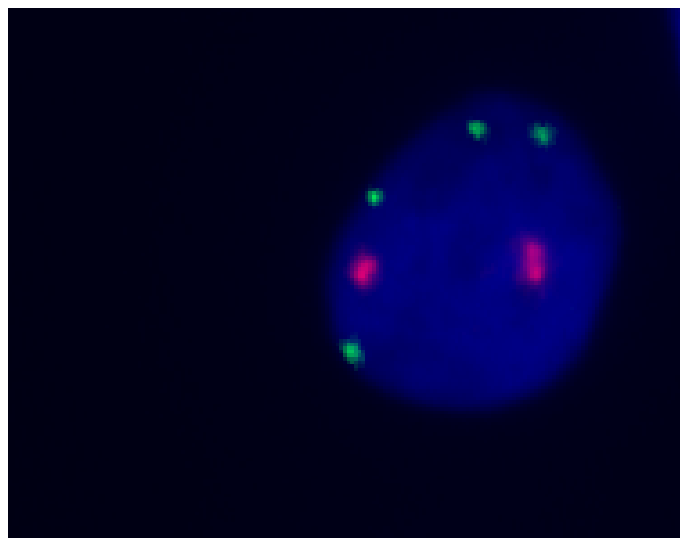


Figure 4: Interphase FISH analysis showing a tetraploid constitution. Nuclei are counterstained blue with DAPI. The four distinct red signals correspond to the chromosome X probe with Tamra-dUTP, and the four distinct green signals correspond to the chromosome 2 probe with Fluorescein-dUTP, confirming four complete sets of chromosomes (tetraploidy).

1.6. Copy Number Variations

Copy Number Variations (CNVs) have emerged as a critical genetic frontier in SA[58]. CNVs are changes in genomic segments, ranging from several dozens (> 50bp) to megabases, that deviate from the standard diploid state through deletions or duplications [59]. Even though some embryos previously screened as euploid from Preimplantation Genetic Testing for Aneuploidy (PGT-A), subsequent genetic analysis of miscarriage tissue often reveals hidden CNVs or mosaicism that explain the failure of the pregnancy [60].

The evolution of detection sensitivity has transformed the understanding of the genomic architecture of pregnancy loss. For last decades, traditional karyotyping remained the gold standard, yet its resolution is restricted only for large-scale changes. This low resolution meant that many minor structural abnormalities remained undetected. The field shifted toward Chromosomal Microarray (CMA), which applied array comparative genomic hybridization (CGH) or Single Nucleotide Polymorphism (SNP) arrays to detect sub-microscopic gains and losses [61]. CMA significantly improved diagnostic capabilities by identifying CNVs as small as 200kb while Low-pass Whole Genome Sequencing (LP-WGS) has emerged as a reliable and cost-effective alternative, recently. LP-WGS offers a resolution comparable to CMA while requiring less input DNA and providing better sensitivity for detecting low-level mosaicism and cytoband-level variations [61]. This detection capabilities has allowed clinicians to identify pathogenic CNVs (pCNVs) in approximately 2.7% to 5.26% of miscarriage cases, particularly in fetuses or before six weeks of gestation [59].

Molecular genesis of CNVs in the human germline fundamentally driven by errors in DNA repair and replication, through Non-Allelic Homologous Recombination (NAHR) and Microhomology-Mediated Break-Induced Replication (MMBIR) [62]. NAHR is the leading mechanism for repetitive CNVs as it occurs when the cellular machinery mistakenly aligns highly similar, non-allelic sequences such as segmental duplications or Low-Copy Repeats (LCRs) instead of the correct alleles during meiosis [62]. While these unequal crossing-over results in reciprocal deletion and duplication events. Conversely, non-recurrent and complex structural rearrangements often occur through MMBIR. This replication-based model suggests that when a replication fork stalls or collapses, the DNA polymerase jumps to a different template by utilizing tiny patches of microhomology, typically 2-33 base pairs. This template switching can result in non-reciprocal complex genomic imbalances. Which are particularly disruptive to early embryonic development [62]. These processes highlight how the human genome's complex structure, specifically its many repetitive elements makes specific regions prone to the structural instability seen in spontaneous abortions [62].

A fundamental challenge in clinical genetics is the "Dichotomy of Impact," which requires differentiation between benign polymorphic CNVs and those that are pathogenic. Not all genomic imbalances lead to harmful outcomes as many CNVs are common in healthy populations and represent normal genetic diversity. However, pathogenic CNVs are those that disrupt dosage-sensitive genes or regulatory elements essential for morphogenesis. According to 2023 ACMG guidelines, variants are classified into five categories as pathogenic, likely pathogenic, variant of uncertain significance (VUS), likely benign, and benign [63]. When it comes to spontaneous abortion, the pathogenicity of a CNV is often determined by its size, gene density, and whether it involves haploinsufficient genes.

1.7 Mosaic Embryos

The transition from a fertilized oocyte to a blastocyst is characterized by a high rate of chromosomal instability, which is mostly separate from the meiotic errors inherited from gametes. In previous chapters we have discussed constitutive aneuploidy, embryonic mosaicism can be defined as the presence of two or more genetically distinct cell populations within a single embryo, which occurs through post-zygotic mitotic errors [64]. The primary mechanical drivers of mosaicism are mitotic errors such as non-disjunction, anaphase lag, and endoreduplication occurring during the early cleavage stages [65, 66]. In mitotic non-disjunction, sister chromatids fail to separate properly during anaphase which results in one daughter cell with a trisomy and another with a monosomy. And anaphase lag occurs when a single chromosome fails to attach to the spindle apparatus or moves too slowly toward the cell pole. Which leads to its exclusion from reforming the nuclei and the subsequent formation of a micronucleus [66].

The molecular background of these mitotic failures is highly linked to weakened cell-cycle checkpoints and spindle instability in the embryo. Unlike somatic cells, cleavage-stage embryos often exhibit a relaxed spindle assembly checkpoint (SAC), which allows cells to proceed through mitosis even when chromosomes are not correctly aligned or attached [66]. This vulnerability is intensified by centrosome dysfunction and kinetic misalignment, which frequently result in the segregation of chromosomes into micronuclei spatially isolated from the main nucleus [66]. Recent transcriptomic analyses have found that mosaic embryos possess a genuine molecular identity, followed by significant deregulations in pathways involving apoptosis, protein degradation, and mitochondrial energy production [65]. These metabolic stresses lead the embryo to continued chromosomal instability during the development.

In the clinical context of spontaneous abortions and prenatal diagnosis, mosaicism is classified based on its tissue distribution. Therefore a critical distinction must be made between true fetal mosaicism and confined placental mosaicism (CPM) [67]. True fetal mosaicism (TFM) occurs when the abnormal cell population is present in both the fetus and the placenta and confined placental mosaicism is restricted only to the extraembryonic tissues, such as the trophoblast or chorionic villi [67]. Recent studies suggest that approximately 87% of mosaicism detected in chorionic villus sampling is limited to the placenta (CPM) and only about 13% involves the actual fetus (TFM) [64]. These classifications make a significant influence in clinical practice as CPM involving specific chromosomes such as 2, 7, 16, or 22. They are significantly associated with

intrauterine growth restriction and an increase risk of fetal demise, even if the fetus itself is chromosomally euploid [64].

Furthermore, the level of mosaicism categorized as low-level (20–50% abnormal cells) and high-level (>50% abnormal cells) is a decisive factor in pregnancy progression. Research indicates that high-level mosaic embryos are identified in a significantly higher rate of miscarriages compared to low-level mosaics [64]. The prevalence of mosaicism tends to target specific chromosomal regions, frequently affecting longer chromosomes, which observed between group A and group C and group B and group C [68]. However additional studies are necessary to determine the characteristics and mechanisms of mosaicism.

1.8 Standard Karyotyping Method and Its Limitations

The genetic investigation of spontaneous abortions (SA) has historically relied upon the visualization of the human genome through conventional cytogenetics methods. Standard karyotyping, specifically G-banded analysis, has remained the clinical "gold standard" since the late 1970s for identifying chromosomal abnormalities [69]. This technique involves the systematic pairing and arrangement of an individual's complete set of chromosomes, allowing for a comprehensive but a low-resolution view. However, as the field moves toward precision medicine, the biological and technical limitations of this analogue method have become increasingly observed, demanding the shift toward molecular platforms.

Standard karyotyping is a detailed process that depend on active participation of living biological material. Unlike modern molecular assays that use extracted DNA, karyotyping requires the successful cultivation of active fetal cells. For samples derived from spontaneous abortions, this typically involves the collection of chorionic villi, fetal skin, or internal fetal tissues. These tissues are processed and placed into specialized culture media, where they are incubated to encourage mitotic division. This process reaches a critical phase during the metaphase stage of mitosis. When chromosomes are at their most condensed and visible state. To capture cells at this stage a spindle-disrupting agent, such as Colcemid, is introduced to arrest the cells and it is followed by a hypotonic treatment, with a KCl solution usually. Which causes the cells to swell, thereby facilitating the effective spreading of chromosomes upon a microscope slide [70].

The most common staining technique is Giemsa banding (G-banding). In this procedure, the slides are briefly treated with trypsin a proteolytic enzyme that partially digests chromosomal proteins which followed by staining with Giemsa dye. This results in a unique pattern of light and dark bands. Dark G-bands generally represent AT-rich, gene-poor regions that replicate late in the S-phase. While light G-bands represent GC-rich, gene-dense regions. A standard clinical karyotype usually produce a resolution of 400 to 550 bands per haploid set, which translates to a detection threshold for structural changes of approximately 5 to 10 Mega bases (Mb) [70]. The primary application of standard karyotyping in analysing spontaneous abortions is its ability to provide a complete view of the genome. It is exceptionally effective at identifying whole-chromosome aneuploidies and the common findings include autosomal trisomies, monosomy X and polyploidies [71].

One uniqueness of conventional karyotyping and a reason it has preserved despite molecular advancements is its ability to detect balanced structural rearrangements, such as reciprocal or Robertsonian translocations and inversions. While these rearrangements do not involve a gain or loss of genetic material in the balanced state, their detection is vital for parental counselling. If a parent is a carrier of a balanced translocation, they have a significantly risk for future recurrent pregnancy losses or having offspring with unbalanced chromosomal profiles. So, karyotyping remains a primary tool for mapping these physical arrangements across the genome. Despite its wide usage, standard karyotyping is limited by significant technical vulnerabilities, the most prominent of which is the requirement for cell culture. In cases of spontaneous abortion, the tissue provided is often products of conception (POC). Which may have remained in utero for days or weeks following fetal demise. Consequently, these cells are often macerated or no longer viable. Studies indicates that culture failure occurs in approximately 25% of SA samples [72].

Furthermore, the culture process poses the risk of maternal cell contamination when POC samples cultured. There is a possibility that maternal decidual cells will proliferate instead of fetal cells. Because maternal cells are often more resilient and viable than the fetal tissue, they can overgrow the culture. This leads to a "false normal" 46, XX result which is shown by the studies using molecular confirmation that a significant portion of 46, XX results obtained by standard karyotyping are actually maternal in origin. Resulting masking the true fetal genotype, which may have been aneuploid [73]. Another major limitation of standard karyotyping is its low resolution. Because the method relies on the physical condensation of chromatin and the optical limits of light microscopy. It is insensitive to sub-microscopic aberrations such as copy number variants (CNVs), including microdeletions and microduplications that are smaller than 5–10 Mb [73].

This thesis centers on NGS to address the limitations of standard karyotyping as a tool which has reached its maximum potential. While it remains superior in its ability to detect truly balanced translocations, its failure rate and resolution limits in the context of pregnancy loss are considerably high in diagnostic environment. The limitations discussed in this chapter are key diagnostic barriers that challenge understanding the etiology of pregnancy loss and clinicians from providing accurate recurrence risks.

1.9. Karyotyping using Next Generation Sequencing methods

1.9.1 Comparative Analysis of Evolving Karyotyping Methodologies

Over the past decade the technological landscape for analysing genomic aberrations in POC from spontaneous abortions has evolved significantly. While classical cytogenetics relied on cell culture and the chromosomes analysis. Modern reproductive medicine increasingly prioritizes unbiased, high-throughput molecular platforms to overcome the biological limitations of necrotic or uncultured tissues. Consequently, the transition toward molecular karyotyping upgrade in diagnostic precision, enabling a comprehensive view of the fetal genomic landscape that was previously unavailable through optical macroscopic techniques.

The specific technical boundaries, diagnostic yields, and operational restrictions of traditional G-banding versus automated molecular karyotyping platforms are defined and contrasted in Table 9.1.

Table 9.1: Technical and Diagnostic Performance Comparison of Karyotyping Methodologies

Technical Parameter / Feature	Standard G-Banding Karyotyping	Molecular Karyotyping (NGS / CMA)	Supported Peer-Reviewed Literature (Post-2018)
Diagnostic Detection Resolution	Low resolution bound strictly to macroscopic bands (typically $\geq 5\text{--}10$ Mb).	High resolution enabling subchromosomal micro-aberration tracing down to < 1 Mb.	Bai et al., 2023 [58]; Dai et al., 2024 [74]
Tissue Culture and Viability Reliance	In some occasions requires viable, living cells arrested in metaphase; vulnerable to complete sample loss.	Completely independent of cell culture; utilizes genomic DNA extracted from non-viable tissue.	Chen et al., 2024 [75]; Dai et al., 2024 [74]
Diagnostic Failure and Culture Omission Rate	High operational failure rates (10%–40%) driven by tissue necrosis and maceration in spontaneous abortions.	Negligible failure rate ($< 1\%$); successful in degraded or long-retained products of conception (POC).	Chen et al., 2024 [75]
Maternal Cell Contamination (MCC) Bias	Risk of false-normal results due to the presence of normal-karyotype cell clones rather than maternal decidual contamination.	Mitigated through quantitative genotyping, digital read counting, or parent-offspring trio SNP comparison.	Chen et al., 2024 [75]; Lei et al., 2023 [3]
Structural Polyploidy Recognition (Triploidy/Tetraploidy)	Readily identifies complete triploidy (69, XXX) and balanced polyploid matrices through physical visual identification.	Challenging for standard NGS; requires precise B-allele frequency (BAF) parsing or specialized bioinformatic ploidy algorithms.	Chen et al., 2024 [75]; Geng et al., 2019 [76]
Truly Balanced Structural Rearrangements	Detectable; clearly visualizes balanced reciprocal translocations, robertsonian translocations, and inversions.	Invisible; copy-number-neutral structural variations show no net balance change in total read-depth or hybrid signal data.	Chen et al., 2024 [75]
Low-Level Mosaicism Tracking Sensitivity	Constrained by cell counts (standard 20-metaphase analysis); prone to cultivation-induced selective pressure artifacts.	Highly sensitive; statistical depth variations reliably capture somatic or placental mosaicism down to $< 10\%$ levels.	Dai et al., 2024 [74]

To address the high failure rates and low-resolution limits associated with conventional G-banding, alternative molecular karyotyping techniques have been integrated into clinical laboratories such as Chromosomal Microarray Analysis (CMA) and Next-Generation Sequencing (NGS)-based copy number variation sequencing (CNV-seq) [76]. Chromosomal Microarray Analysis includes array Comparative Genomic Hybridization (aCGH) and Single Nucleotide

Polymorphism (SNP) microarrays. These platforms depends on the competitive hybridization of fluorescently labelled patient and reference DNA to microchips embedded with locus-specific oligonucleotide probes. While CMA platforms advanced cytogenomics through kilobase-scale profiling of subchromosomal aberrations. However, reliance on pre-engineered physical probes imposes diagnostic limitations, particularly variable hybridization within regions of high sequence homology or repetitive elements [58, 76].

In contrast, Next-Generation Sequencing (NGS) based karyotyping such as low-pass whole-genome sequencing or CNV-seq, operates independently of target-specific physical hybridization probes [58]. NGS architectures implement target-free massively parallel sequencing to fragment, capture, and sequence the entire extracted genomic template uniformly across all autosomes and sex chromosomes, instead of scanning predefined loci. The resulting short reads are digitally mapped to a linear reference sequence, such as the GRCh38/hg38 assembly. By applying statistical binning models across the linear genome, CNV are reconstructed directly from the read depth or relative density of sequenced fragments within each designated genomic coordinate window [58]. This digital quantification approach eliminates the analog signal noise typical of fluorescent microarrays. Which offers an un-targeted, highly linear, and scalable method to map whole-chromosome aneuploidies, mosaicism, and micro-deletions or micro-duplications simultaneously from a single tissue sample [58, 74].

1.9.2 Overview of Whole-Genome Sequencing (WGS) Architectures

Short-read whole-genome sequencing architectures, predominantly dominated by Illumina and MGI sequencing platforms, serve as the technological foundation for standard copy number variation sequencing (CNV-seq). These instruments use sequencing-by-synthesis (SBS) or rolling circle replication technologies. Combining with fluorescently labeled reversible terminators to generate millions of highly accurate, uniform short reads ranging from 35 bp to 150 bp or 250 bp in length[77]. The primary strength of short-read architectures lies in their exceptionally high base-calling accuracy and massive data throughput. Which allow dozens of labeled patient samples to be pooled within a single sequencing flow cell lane.

For molecular karyotyping, short-read WGS does not require deep coverage to resolve single nucleotide variants (SNVs). Instead, it relies on the generation of millions of independent sequence tags that can be uniformly distributed across the reference genome [58]. Because these short fragments are easy to amplify and sequence even from highly compromised, partially degraded DNA templates. Short-read platforms have established themselves as the most cost effective and industrially scalable methodology for screening of early pregnancy losses [74].

While short-read sequencing currently dominates clinical diagnostics. However, long-read whole-genome sequencing architectures primarily Oxford Nanopore Technologies (ONT) and Pacific Bioscience (PacBio) Single-Molecule Real-Time (SMRT) sequencing have introduced a separate models for structural cytogenomics [77]. Nanopore platforms operate by passing single-stranded DNA molecule through synthetic protein nanopore embedded in an electrically resistant membrane. When the nucleotide strand translocates through the pore, it causes characteristic disruptions in the ionic current and that electrical fluctuations are translated into sequence data via

neural network base-calling algorithms. PacBio systems uses the zero-mode waveguides (ZMWs) to record the fluorescent signals of lone polymerases incorporating labelled nucleotides in real time, producing highly accurate long reads known as High-Fidelity (HiFi) reads.

The analytical advantage of long-read architectures is their capacity to span complex structural variants (SVs), large insertions or deletions, and repetitive heterochromatic or centromeric genomic landscapes that confound short-read alignment algorithms [77]. Because a single read can span tens to hundreds of kilobases, these platforms can map structural breakpoints directly and resolve copy-number-neutral chromosomal rearrangements [77].

However, several technical and financial barriers limit the ability to implementation of long-read WGS for screening POC in spontaneous abortions. The raw error rate of single-molecule sequencing, though significantly mitigated by PacBio circular consensus sequencing or Nanopore duplex configurations. Which remains higher than that of short-read platforms. Furthermore, long-read sequencing requires high molecular weight and un-degraded genomic DNA templates. This requirement is difficult to satisfy when processing tissues from spontaneous abortions. Coupled with a lower total throughput per flow cell, long-read platforms incur a substantially higher cost per sample compared to highly multiplexed short-read sequencing.

The implementation of an untargeted, whole-genome sequencing strategy rather than targeted enrichment protocols such as whole-exome sequencing (WES) or multi-gene panels is methodologically essential when configuring digital karyotyping for SA tissues [58]. Consequently, untargeted whole-genome strategies provide a reliable reference for genome-wide read-depth calculation. Which enables the simultaneous detection of large-scale aneuploidies, subchromosomal micro-aberrations, and structural imbalances with optimal sensitivity.

1.9.3 Ultra-Low Coverage Whole-Genome Sequencing (ulc-WGS)

Ultra-low coverage whole-genome sequencing (ulc-WGS), frequently referred as low-pass WGS or low-depth copy number variation sequencing (CNV-seq). Which represents a specialized application of high-throughput sequencing, optimized for macroscopic structural analysis [74]. In standard genomic pipelines, WGS requires a deep base-level coverage typically ranging from 30× to over 60× depth. Which means that every individual nucleotide position within the 3 billion base human genomes sequenced an average of 30 to 60 times. This high depth is mathematically necessary to identify heterozygous single nucleotide mutations, small indels, and low-frequency somatic variants against background sequencing noise.

In contrast, ulc-WGS operates within a significantly restricted sequencing depth paradigm, typically bounded precisely between 0.01× and 0.1× coverage per base position where the sequence reads cover only a small fraction of the genome [78]. This approach prioritizes genome-wide breadth and statistical counting efficiency over individual base-calling depth. Consequently, the analysis shifts from identifying localized point mutations to statistically tracking millions of independent, randomly distributed sequence tags across the entire genome [78].

The mathematical validity of ulc-WGS relies on the fact that macroscopic chromosomal aberrations including trisomies, monosomies, and large deletions that trigger spontaneous

abortions, measurably skew DNA volume across whole chromosomes or subchromosomal domains[58]. By fragmenting the genome and generating between 1 and 5 million short sequence reads per sample, the pipeline has no necessity to read every nucleotide to determine chromosomal status. Instead, the total volume of generated sequence data processed as an aggregate digital count. When these millions of reads are aligned to a reference genome, an increase or decrease in the local density of mapped reads within a specific chromosomal region provides a direct, statistically verifiable measure of copy number state [58]. This approach allows clinical laboratories to overcome the high financial costs associated with deep-coverage sequencing while achieving better diagnostic resolutions than traditional cytogenetic G-banding [58]. The clinical execution of ulc-WGS on products of conception tissues follows a highly optimized, multi-step workflow that transitions from wet-lab molecular extraction to dry-lab statistical modelling.

2. Methodology

2.1. Research Approach

A total of 9 medical abortions samples were used as negative control, 10 samples of spontaneous abortions with known karyotype with aneuploidy were used for positive control, and 186 samples of spontaneous abortions with an unknown karyotype were included in the analysis. Karyotyping the control group was performed using a standard cytogenetic method (G-o staining) and interphase FISH analysis. For interphase FISH, centromeric DNA probes for chromosomes 13/21, 18, and X were used. Extraembryonic mesoderm samples were separated under an Axiovert 200 inverted microscope (Carl Zeiss, Germany). Genomic DNA was isolated using the standard phenol-chloroform extraction method.

DNA quality and quantity were analyzed using a Nanodrop (Thermo, USA) and Qubit 2.0 (Thermo, USA) spectrophotometer. Libraries were prepared using the MBU Tn5 transposase (MBS-Technology, Russia) according to the manufacturer's recommendations. For adapter sequences preparation the Nextera XT DNA Library Preparation Kit (Illumina, USA) were used. Library quality was determined using a 2100 Bioanalyzer (Agilent Technologies, USA). Libraries were sequenced in FCM V2.0 flow cell on a Genolab M sequencer using the GenoLab M Sequencing Reagent (300 cycles) V2.0 kit (GeneMind, China), according to the manufacturer's recommendations.

Quality control of the raw sequence data was performed using FastQC v0.11.4 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The average genome coverage was 0.3x and read alignment was performed using bwa 0.7.17 against the human reference genome GRCh38 (hg38). Uniquely aligned reads were obtained and filtered for mapping quality using SAMtools v1.6. The resulting .bam files were processed using QDNAseq 1.40.0. For sex chromosomes, GC content correction was calculated using ComputeGCBias (a tool in the deepTools package) using the Benjaminy method. The genome was divided into 50 kbp regions, which coverage was calculated using bedtools v2.30.0.

For polyploidy analysis, chorionic tissue suspensions and centromeric DNA probes for chromosomes 2 and X were used. Between 100 and 400 interphase nuclei were counted in each sample. For all samples which preparations were possible Interphase chorionic nuclei were tested for polyploidy using DNA probes targeting the centromeres of chromosomes X and 2. Aneuploidy was verified using DNA probes targeting the centromeres of chromosomes 4, 7, 8, 9, 10, 15, 16, 18, 20, and 13/21, as well as unique probes targeting chromosomes 13 (the *RB1* gene region) and 22 (the *DGCR2* gene region). The analysis was performed using an Axio Imager Z2 microscope (Carl Zeiss, Germany) with Metafer and ISIS software (MetaSystems, Germany). When only DNA was available for the sample aneuploidy and CNV verification analysis was performed using real-time PCR. Oligonucleotide primers were selected using the NCBI Gene, UCSC Genome Browser, Ensembl databases, Primer-BLAST and Oligoanalyzer software for regions of various chromosome arms (4, 9, 16) or the long arm region of chromosomes 14 and 22. The analysis was

performed on an AriaMx instrument (Agilent Technologies, USA). Figure 5 illustrates the detailed schematic workflow of the research methodology

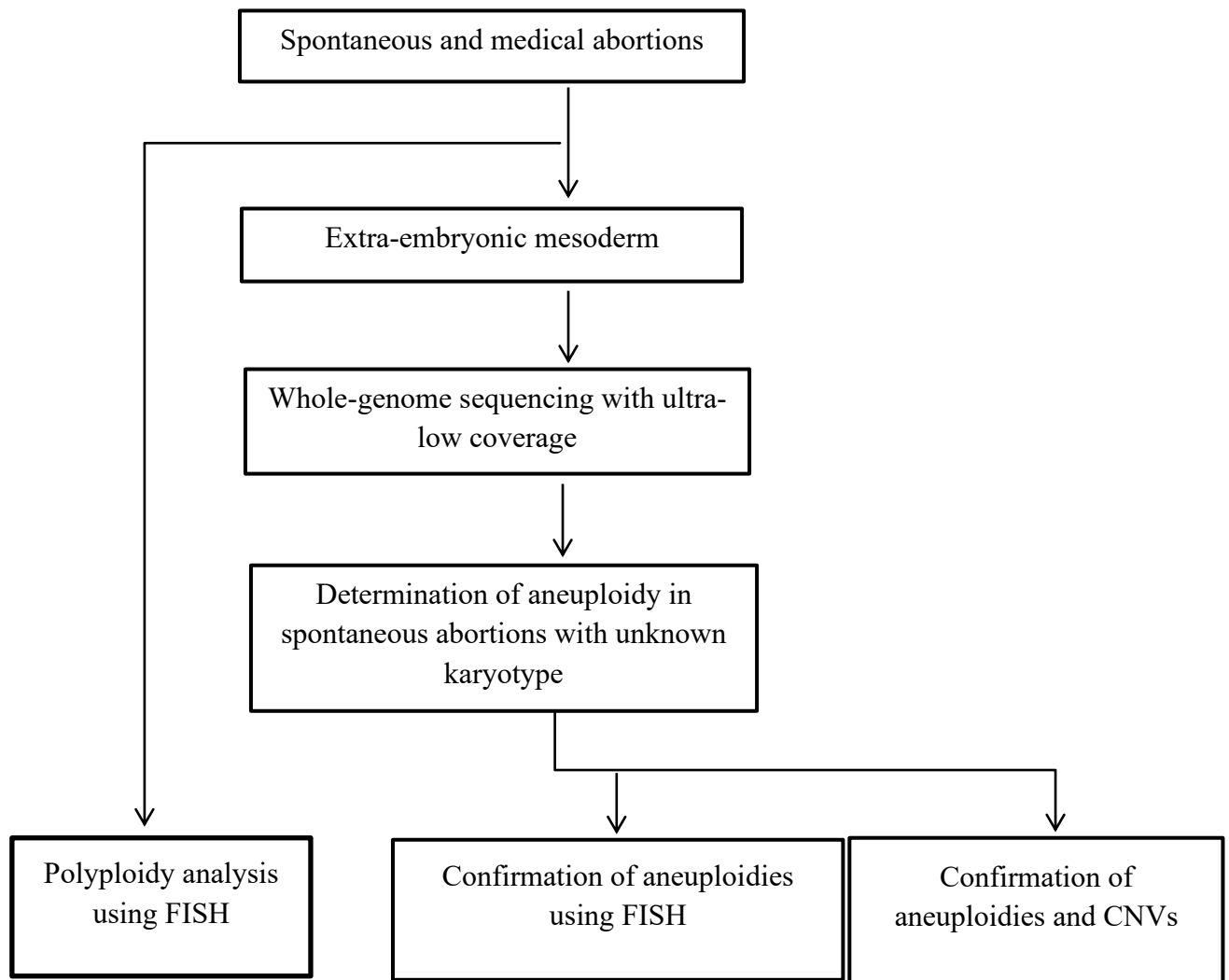


Figure 5: Sample verification workflow

2.2 Sample Collection

The biological material obtained for this research project was sourced from the biological collection "Biobank of the Population of Northern Eurasia," maintained at the Research Institute of Medical Genetics, Tomsk National Research Medical Center, Russia. Tissue samples were specifically isolated from uncultured extraembryonic mesoderm derived from both spontaneous and induced (medical) abortions. For analysis 9 samples from induced abortions and 196 samples of spontaneous abortions were used. The gestational age of the pregnancies, calculated from the date of the mother's last menstruation, ranged from 4 to 17.8 weeks, with a calculated mean gestational age of 8.3 ± 2.9 weeks. The maternal age distribution at the time of pregnancy loss spanned from 18 to 42 years, yielding an average maternal age of 28.5 ± 6.8 years. All sample collection protocols, clinical data aggregation, and experimental procedures were executed in compliance with the ethical principles outlined in the Declaration of Helsinki by the World

Medical Association(WMA). In accordance with these international ethical benchmarks, the study protocol was formally reviewed and approved by the local Biomedical Ethics Committee of the Tomsk National Research Medical Center prior to implementation (Protocol No.19, October 13,2023) To safeguard patient autonomy and rights, written informed consent was obtained from all participants before any clinical material or personal medical history was collected. Vulnerable populations were protected, and a risk-benefit assessment was performed to ensure that the molecular evaluation of the abortive tissue posed zero risk to the patients.

2.3 Procedure

2.3.1. DNA Extraction

DNA extraction was performed using the phenol-chloroform extraction method.

- Add 10–20 mg of clean chorionic villi tissue and place it into a 1.5 ml Eppendorf tube.
- Finely mince the tissue using a sterile scalpel, or homogenize it using a micro-pestle.
- Add 0.270 ml of 0.2 M sodium acetate directly to the homogenized tissue pellet and vortex for 10–15 seconds.
- Add 0.03 ml 10% SDS and 0.005 ml proteinase K.
- Incubate at 55°C for 1 hour, mixing by vortexing or inverting every 15 minutes.
- Add 0.15 ml phenol and 0.15 ml chloroform to the Eppendorf. Mix by vortexing or inverting for 10 minutes.
- Transfer the aqueous phase (upper layer) to a new tube, avoiding the interphase.
- Add 0.3 ml chloroform and mix for 10 minutes.
- Transfer the aqueous phase to a new tube, avoiding the interphase.
- Add 1 ml of 96% ethanol at room temperature and gently invert the tube until DNA precipitates. Spool the DNA using a pipette tip or glass rod.
- Centrifuge at 13,000 rpm for 2 minutes. Carefully remove the ethanol by pipetting or decanting.
- Add 0.5 ml of 70% ethanol and mix gently.
- Centrifuge at 13,000 rpm for 2 minutes and remove the ethanol.
- Air-dry the DNA pellet for 5–10 minutes or use a vacuum dryer.
- Resuspend the DNA in 0.05 ml TAE buffer or distilled water and store in -20°C.

2.3.2. Tissue Extraction and Suspension Preparation

Required materials and solutions

Using sterile fine-point forceps and scalpels, approximately 20–50 mg of clean, chorionic villi are dissected away from the spontaneous abortions samples.

0.6% (w/v) sodium citrate was prepared by dissolving 0.6 g of trisodium citrate dihydrate ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7 \cdot 2\text{H}_2\text{O}$) in 80 ml of deionized water. The mixture was stirred until complete dissolution was achieved and the final volume was adjusted to 100 ml using a volumetric flask. The solution was filtered and stored at 4°C to maintain stability and prevent microbial contamination prior to use in cell swelling procedures.

60% (v/v) aqueous acetic acid solution was prepared by adding 60 ml of glacial acetic acid ($\geq 99.8\%$ purity) to 40 ml of distilled water. The solution was allowed to equilibrate to room temperature before use.

Preparation of Chorionic Villus Cell Suspensions

- Transfer chorionic villi into 15 ml centrifuge tubes; They can be frozen or used immediately.
- Add 10 ml of 0.6% Sodium Citrate ($\text{Na}_3\text{C}_6\text{H}_5\text{O}_7 \cdot 2\text{H}_2\text{O}$) to the 15 ml tube; incubate for 20 minutes at 37°C.
- Centrifuge for 10 minutes at 1000 rpm and then remove supernatant, leaving up to 1 ml of sediment.
- Add 10 ml of fixative (Methanol: Acetic Acid = 3:1). Incubate for 15 minutes at -20°C.
- Centrifuge for 10 minutes at 1000 rpm and then remove supernatant, leaving up to 1 ml of sediment.
- Repeating the above step, add 10 ml of fixative again (Methanol: Acetic Acid = 3:1). Incubate for 15 minutes at -20°C.
- Centrifuge for 10 minutes at 1000 rpm. Remove supernatant, leaving up to 1 ml of sediment.
- Transfer the 1ml suspension to a 1.5 ml Eppendorf tube and centrifuge for 5 minutes at 6000 rpm. Remove supernatant, leaving up to 100 μl .
- Add 1 ml of 60% Acetic Acid.
- Incubate for 5 minutes at room temperature, periodically vortexing.
- Centrifuge for 5 minutes at 6000 rpm. Remove supernatant, leaving up to 100 μl .
- Add 1 ml of fixative. Store the suspension at -20°C.

2.3.3. Real-Time PCR

Design of Primers

The selection of optimal primers for the reaction was carried out using the Primer3 program (<http://bioinfo.ut.ee/primer3-0.4.0/>) and the Ensemble database (<http://www.ensembl.org/index.html>). The quality of the selected primers is verified using the Primer-BLAST program (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

When selecting primers, the following conditions are taken into account:

1. Primer length – 20–24 bp;
2. Product size – 100–150 bp;
3. Annealing temperature – 58–62 °C;

4. C/G content – 30–80%;
5. Four or more 3'-terminal nucleotides must not be complementary to the primer itself, the primer in the pair, or other synthetic oligonucleotides added to the reaction;
6. The primer should not contain four or more consecutively located identical nitrogenous bases.

The selection of primers for PCR was carried out according to the following protocol:

- To determine the sequence, the corresponding rearrangement coordinates were entered into the UCSC Genome Browser Home in the format.
- Direct selection of primers was performed in NCBI Primer BLAST.
- The joining or FASTA sequence was entered in the PCR Template field.
- In the Range section, in the Forward primer/From field, enter the coordinates of the beginning of one of the exons (taken from the description of the mRNA sequence). In the Range section, in the Reverse primer/To field, enter the coordinates of the end of one of the exons (taken from the description of the mRNA sequence).
- In the PCR product size field, we indicate min – 100, and max – 500.
- Primer melting temperatures (T_m): 57, 60, 63. Max T_m difference – 3.
- IN section Primer Pair Specificity Checking Parameters, in In the Database line, select Genomes for selected organisms (primary reference assembly only).
- IN section Advanced parameters, in In the Primer Size line we indicate: min – 17, opt – 22, max – 27.
- Primer GC content (%): 20.0, 80.0
- Max Poly – X: 5
- Max 3' Stability: 9
- Max GC in primer 3' end: 5
- We do not reset the site's automatic settings.
- At the end of the page, additionally check Show results in a new window and click the Get Primers button.
- After these steps, a window opens with the primers most suitable for the settings.

Found primers were checked for duplexes in the Vector NTI program.

- In the Vector NTI program, click on Analyses/Oligo Analyses, and then select Oligo Duplexes.
- In the new window, select Add New... On the General tab, enter the sequence of the currently selected primers. First for Forward, then for Reverse.
- Having selected both primer sequences, click on them with the mouse. Click Analyze.
- Evaluate the results obtained at the bottom of the window.
- Save results for the direct primer in a folder.
- If the primer pair satisfies the conditions, proceed to the next step. If not, try a different pair of primers.

In accordance with the described protocol, primers were designed to confirm aneuploidies (Table 1) and CNVs (Table 2).

Table 1. Primers used for checking aneuploidy samples.

Primer	Region	Subsequence
<i>OCIAD1ex7 F</i>	4p11	TGACTCAAGTGTGAGTGGTCAA
<i>OCIAD1ex7 R</i>	4p11	TCAGTAATACCAGTGGGAGCAG
<i>SGCBex2 F</i>	4q12	CAAAGTTCCAATGGTCCTGTAA
<i>SGCBex2 R</i>	4q12	TAAATTGCCCTTTCTTCCTCTC
<i>KANK1intr7 F</i>	9p24.3	TTCAGGGCAAGAACATTGTTAG
<i>KANK1intr7 R</i>	9p24.3	ATAGAGATGCGGGTCTCTTGTT
<i>TRPM6ex33 F</i>	9q21.13	GTCTACAGGTTGGAGGAGAGTT
<i>TRPM6ex33 R</i>	9q21.13	ACGACTCTCATAGCTTTACGGA
<i>FAM155Aex3 F</i>	13q33.3	AGTGTGACTGAGAGATGAAGGA
<i>FAM155Aex3 R</i>	13q33.3	GTCCGAGGGTTAAGTATGATGC
<i>DIAPH3ex21 F</i>	13q21.2	GAGCAATGTGAAGAGACTACGG
<i>DIAPH3ex21 R</i>	13q21.2	GAGCAATGTGAAGAGACTACGG
<i>ATXN3ex11</i>	14q32.12	CATGGAATGTAGGTGTCTGCTT
<i>ATXN3ex11</i>	14q32.12	CATGATGCAAGTGTCAAATCAC
<i>TGM1intr13 F</i>	14q12	CTGTGACAGCCTTTATTGAGGA
<i>TGM1intr13 R</i>	14q12	AAGCTGCTTAGAAGCAAAACCT
<i>PDPRex18 F</i>	16q22.1	TGTTGGCCAGAAATACGGAATC
<i>PDPRex18 R</i>	16q22.1	TCTCTAATTTACCCGAGACTCT
<i>PDZD9ex5 F</i>	16p12.2	CTGGATAATGGTGAAGCAAGACA
<i>PDZD9ex5 R</i>	16p12.2	GGCTTTACCTCTTCAACTTGG
<i>HIRAex16 F</i>	22q11.21	AGAGCAGAACCTTGTGAAAGAG
<i>HIRAex16 R</i>	22q11.21	TGTCTCTACCTCAAGCTCAAGT
<i>MAPK8IP2ex10 F</i>	22q13.33	TAGATACACTCGATTGGCCACA
<i>MAPK8IP2ex10 R</i>	22q13.33	TGTTCTTCAAGGAGTTCATGCC

Table 2. Primers used for checking CNV samples.

Primer	Region	Subsequence
<i>DEFB131A ex1 F</i>	4p16.1	TGATTCTCTCTAACCTGCTTTACCT
<i>DEFB131A ex1 R</i>	4p16.1	AGTTTCTGTCTTACCTGGAGGA
<i>FAM90A26 F</i>	4p16.1	CCTGGTTCCACCGACCTTG
<i>FAM90A26 R</i>	4p16.1	GTGGTGAAAACCCCTCTCACT
<i>KPNA5 ex15 F</i>	6q22.1	AGCAGGAAGCACCAATGGAT
<i>KPNA5 ex15 R</i>	6q22.1	TGGCAACAAAGAGGAGTTACC
<i>LOC101927664 ex3 F</i>	6q21.1	GCAGGAATGTGGGTTAGGGT
<i>LOC101927664 ex3 R</i>	6q21.1	ACTCACCGCAGCATGTTGTT
<i>TBP ex3 F</i>	6q27	CAAACACTTAGCAGCAGCCAG
<i>TBP ex3 R</i>	6q27	GGGTCAGTCCAGTGCCATAA
<i>RPS6KA2 int1 F</i>	6q27	GCTTCGGAATCTGTACTGCG
<i>RPS6KA2 int1 R</i>	6q27	TCACCCTCGTTTAGAGCCAAT
<i>PXDNL ex3 F</i>	8q11.22	AGTCAAACAACACCATAAGGCAT
<i>PXDNL ex3 R</i>	8q11.22	CAGTCTGCTGAACAACAACCA
<i>SNTG1 ex20 F</i>	8q11.21	TTGAAACCCACCAAAAGCAGC
<i>SNTG1 ex20 R</i>	8q11.21	AATGAACCACTTGCCACAGC
<i>DIAPH3 int19 F</i>	13q21.2	GAATCTGGCAAAGCATCTGGC
<i>DIAPH3 int19 R</i>	13q21.2	TGGTGAAGATCCCTAGCATGG
<i>DIAPH3 int25 F</i>	13q21.2	CTTTGGAGCACAAGCAGTCAC
<i>DIAPH3 int25 R</i>	13q21.2	ATCACTGGAAGCCTCTGGTC
<i>PDXDC1 ex11 F</i>	16p13.11	GAATCTGGCAACATTGGCTCTG
<i>PDXDC1 ex11 R</i>	16p13.11	CCATGAGGGACTGAGTGACTT
<i>XYLT1 int1 F</i>	16p12.3	TGGGAAATCAGAACTCGGGG
<i>XYLT1 int1 R</i>	16p12.3	TTGTGCTATGCCATTGGGGA
<i>18q21.2 F</i>	18q21.2	AAGCCAGGTTCTCCTCATGC
<i>18q21.2 R</i>	18q21.2	CAGAGTTGCCTCAGCCTAGA
<i>18q23 F</i>	18q23	TCTACACCTACTGGGCTCCT
<i>18q23 R</i>	18q23	CAGTGAACACTACAACACCATGTTT
<i>ZNF516 int2 F</i>	18q23	ACAAAAGGGGGTGAGTCAGG
<i>ZNF516 int2 R</i>	18q23	ACTAGGGACAGCGTCGTTAC
<i>ZNF763 ex3 F</i>	19p13.2	TGCTTCAGGACTACTTTTCTGTGT
<i>ZNF763 ex3 R</i>	19p13.2	CACTGCTTTCTCTTTAAGTGCCA
<i>MISP ex2 F</i>	19p13.3	ACTCGGAGAACAGGGAGGAT
<i>MISP ex2 R</i>	19p13.3	TCCAGGCGGTAGGTCTTCAT
<i>Y межген F</i>	Yq11.221	CCTTCAGAGCGTGGCTTCAT
<i>Y межген R</i>	Yq11.221	TCCCAGAATGTGTGAGCAGT
<i>NLGN4Y last exon F</i>	Yq11.221	CCATGAGACTCACAGGCACC
<i>NLGN4Y last exon R</i>	Yq11.221	CAGCGACTCACACTCGTGAT
<i>NLGN4Y first exon F</i>	Yq11.221	TTGGCTTGACCTGCGGAAG
<i>NLGN4Y first exon R</i>	Yq11.221	TGAACCTCTTGTCACACAGC
<i>Y межген F</i>	Yq11.221	CCTTCAGAGCGTGGCTTCAT
<i>Y межген R</i>	Yq11.221	TCCCAGAATGTGTGAGCAGT

DNA Sample Preparation

DNA purity and concentration for both test and control samples are verified using a NanoDrop ND-1000 spectrophotometer (USA). The optimal DNA concentration for real-time PCR is 10 ng/μl (9.8–10.2 ng/μl), with an A260/280 ratio of $\approx$ 1.8–2.0 and an A260/230 ratio of $>$ 1. DNA degradation levels were assessed via 1% agarose gel electrophoresis.

To prepare a 1% agarose gel

- mix 0.5 g of agarose, 1 ml of TAE, and 50 ml of water in a 100 ml flat-bottom flask.
- Heat the mixture in a microwave until the agarose completely dissolves.
- Cool the mixture to 50–60 °C, then add ethidium bromide (0.01 μg/μl) under a fume hood and mix thoroughly.
- Rinse the electrophoresis chamber with running water (10 times) and distilled water (5 times).
- Pour the solution onto the gel tray. Once the gel has fully solidified, carefully remove the combs and place the gel into the electrophoresis chamber.
- Parameters for the electrophoresis power supply is fixed as;
 - Time- 30 - 40 min
 - Voltage- 80 V
 - Amperage limit- 100 mA
 - Power- 80W
- DNA fragments are visualized under transmitted UV light using a GelDoc gel documentation system (Bio-Rad, USA).

Optimization of PCR Conditions for Primer Pairs

Once the primer sequences have been selected and synthesized, verification process was conducted to ensure that they yield the target DNA products of the correct quantity and size, without producing non-specific products. Initial test was conducted using primers in a stable and verified DNA preparation as a template. For each real-time PCR run, a protocol is drafted detailing the plate layout and the composition of the reaction mixture.

The PCR UV cabinet is prepared for handling DNA samples. To protect the samples from contamination, the tools and reagents were decontaminated under ultraviolet (UV) light. Dry primers were reconstituted with deionized water to a concentration of 100 μmol/L and mixed on an FV–2400 Microspin (Biosan, Latvia) to ensure complete dissolution. To obtain a working primer concentration of 1 μmol/L, 5 μL of the stock primers were mixed with 495 μL of deionized water.

A control sample with varying DNA concentrations (0.625 ng/μL, 1.25 ng/μL, 2.5 ng/μL, and 5 ng/μL) was added to a 96-well plate at 5 μL per well. In an Eppendorf tube on ice, the master mix is prepared. Calculated per single well, it contains 2.5 μL each of the forward and reverse primers, 2.5 μL of deionized water, and 12.5 μL of Maxima™ SYBR Green qPCR Master Mix (Thermo Scientific, USA). The mixture was gently mixed and spun down using an FV–2400 Microspin microcentrifuge (Biosan, Latvia).

According to the protocol, the contents of the tubes were dispensed into the 96-well plate (20 μL per well). All manipulations were performed on ice. The plate is then sealed with an optically clear film. The plate is centrifuged in a GRANT–BIO LMC3000 centrifuge for 1 minute at 1,000 rpm. Real-time PCR was performed using the AriaMX Real-Time PCR System thermal cycler (Agilent Technologies, USA).

Real-time PCR conditions:

1. Initial denaturation: 95°C – 10 minutes;
2. Amplification (40 cycles):
 - 95°C – 15 seconds;
 - 60°C – 30 seconds;
 - 72°C – 30 seconds.

Upon reaction completion, the DNA amplification plots were analyzed using computer software.

Real-Time PCR Testing of Spontaneous Abortus Samples.

The PCR UV cabinet was prepared for operation. Each PCR mixture contains 25 ng of DNA, 12.5 μ L of Maxima™ SYBR Green qPCR Master Mix (Thermo Scientific, USA), and 2.5 μ L of each primer (1 μ mol/L). The total reaction volume was adjusted to 25 μ L with RNase-free water. The mixture was dispensed into the wells according to the reaction protocol. The plate was then sealed with an optically clear film. The plate was centrifuged in a GRANT–BIO LMC3000 centrifuge for 1 minute at 1,000 rpm.

Real-time PCR was performed using the AriaMX Real-Time PCR System thermal cycler (Agilent Technologies, USA). Upon reaction completion, the DNA amplification plots were analyzed using computer software (Figure 6).

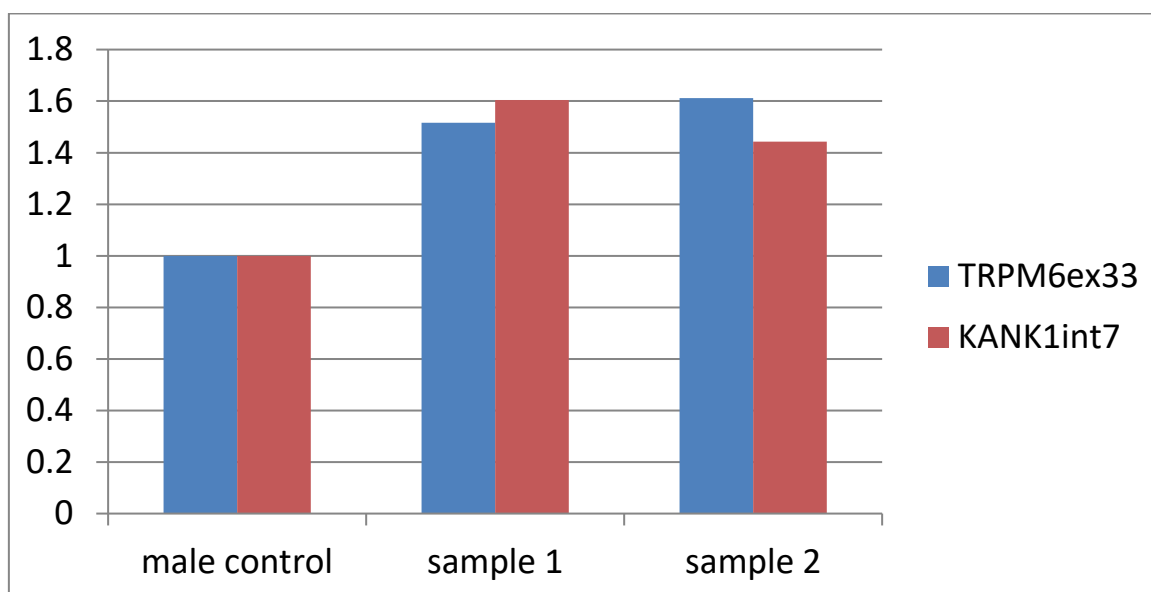


Figure 6: Real-time PCR analysis using TRPM6ex33 and KANK1int7 primers for chromosome 9, revealing a duplication of chromosomal material.

Data analysis was performed using MS Excel software in accordance with the algorithm described in the Guidelines for Medical Technologies of Molecular Diagnostics of Chromosomal and Inherited Diseases[79].

2.3.4. FISH Sample Preparation

Synthesis of DNA Probes

The total length of the amplified DNA fragments used to create DNA probes were exceeding 30 kbp, to ensure good visualization of the DNA probe on interphase chromosomes during FISH analysis. The resulting fragments were located as close as possible on the chromosome but without overlap. One pair of primers was selected for every 10 kbp. The UCSC genome browser (University of California, Santa Cruz), which contains information on genome sequences, was used to obtain the nucleotide sequence. The obtained nucleotide sequence was then used to select primers using standard bioinformatics programs (BLAST) provided by the National Center for Biotechnological Information (NCBI). Primers were selected taking into account the following conditions:

1. All primers had similar temperatures of melting (T_m) and temperatures of annealing ($\pm$ length of the primer can vary from 15 to 30 pairs of nucleotides)
2. The specificity of primers had been at least 70%;
3. Internal homology had not exceed 3 bp;
4. The content of GC dinucleotides had been vary in the range from 45% to 55%.

We excluded primers with extended polypyrimidine or polypurine regions, as well as primers and target sites whose secondary structure melting temperature was higher than or equivalent to the melting temperature of the primer itself. We also assessed the possibility of primer homodimerization and heterodimerization, as this can lead to a decrease in the yield of the final PCR product.

DNA products were generated using long fragment PCR using the BioMaster LR HS-PCR-Color (2 $\times$) kit. (#MHC040-100, Biolabmix, Russia). Fragments were generated on a SureCycler 8800 thermocycler (Agilent Technologies, USA) using the following mode:

94°C – 4 minutes;
10 cycles: 94 °C – 20 seconds;
 59 °C – 30 seconds;
 68 °C – 8 minutes;
25 cycles: 94 °C – 20 seconds;
 59 °C – 30 seconds;
 68 °C – 8 minutes (+10 seconds per cycle to the elongation stage).

Corresponding bands were detected for all amplified fragments by agarose gel electrophoresis. Using this optimized protocol, custom probes were generated for chromosomes 13, 21 and 22.

To evaluate the polyploidy state commercial centromeric probes were used, as detailed in Table 3.

Table 3. DNA probes used for testing polyploidy

DNA probe	Locus	Plasmid	Insert Size(bp)	Region
pZ5.1	D1Z7 / D5Z2 / D19Z3	pUC13	680	1p11-q11, 5p11-q11, 19p11-q11
pBS4D	D2Z	bluescribe	1300	2p11-q11
pAE0.68	D3Z1	bluescribe	680	3p11-q11
p4n1/4	D4Z1	bluescript	680	4p11-q11
pEDZ6	D6Z1	bluescript	340	6p11-q11
pZ7.6B	D7Z1	pUC13	680	7p11-q11
pZ8.4	D8Z2	bluescribe	1200	8p11-q11
pMR9A	D9Z	bluescribe	340	9p11-q11
pZ101.3	D10Z1	bluescript	2700	10p11-q11
pRB11	D11Z1	bluescribe	680	11p11-q11
pRB12	D12Z3	bluescribe	680	12p11-q11
p14.1	D14Z1 / D22Z1	pUC13	1360	14p11.1, 22p11.1
pMC15	D15Z1	bluescript	3500	15p11.1
pZ16A	D16Z2	pUC13	680	16p11-q11
pZ17-1.6A	D17Z1	pUC19	1020	17p11-q11
2Xba	D18Z1	pUC18	1700	18p11-q11
pZ20	D20Z1	bluescribe	1360	20p11-q11
pZ21A	D13Z1 / D21Z1	bluescribe	850	13p11.1, 21p11.1
pI90.22	D22Z	bluescribe	2000	22p11.1
pDMX1	DXZ1	bluescript	2000	Xp11.1-q11.1
pLAY5.5	DYZ3	bluescript	5500	Yp11.1-q11.1

Nick Translation

DNA libraries were labeled by nick translation using fluorescent dyes flu-12-dUTP and TAMRA-5-dUTP.

The reaction mixture contained 5 μ L of a deoxyrimonucleoside triphosphate mixture (0.5 mM dATP, dCTP, dGTP, 0.1 mM dTTP), 5 μ L of 10 $\times$ DNA polymerase I buffer, 1 μ L of biotin-16-deoxyuridine triphosphate or digoxigenin-11-deoxyuridine triphosphate (1 nM/ μ L), 5 μ L of bovine serum albumin (0.5 mg/mL), 5 μ L of DNase I diluted 1:49 with deionized water (from a 1 U/ μ L stock solution), 1 μ L of DNA polymerase I (10 U/ μ L), and 1 μ g of DNA obtained by LR-PCR. The mixture volume was adjusted to 50 μ L with deionized water. The reaction was carried out at 15 $^{\circ}$ C for 1.5 hours and stopped by adding 1 μ L of 0.5 M EDTA (pH 8.0). The probe was precipitated in the presence of a 100-fold excess of Cot1 DNA, 1/20 volume of sodium acetate, and 2.5 volumes of 100% ethanol (-20 $^{\circ}$ C). After incubating the mixture at -20 $^{\circ}$ C overnight, it was centrifuged twice for 30 minutes at 4 $^{\circ}$ C and 12,000 rpm, dried at 45 $^{\circ}$ C in a concentrator (V-

AL) for 3–5 minutes until the droplets completely disappeared, and dissolved in 20 μ L of hybridization mixture (2 parts of 100% formamide, 1 part of 50% dextran sulfate).

Label incorporation into the DNA was performed by the standard nick translation method (Maniatis T. et al., 1984) using *E. coli* DNA polymerase I enzymes (ThermoFisher Scientific, USA), DNase I (ThermoFisher Scientific, USA), and a mixture of 4 deoxyrimonucleotide triphosphates (Sigma, USA): dATP, dCTP, dGTP, and dTTP in a 5:5:5:1 ratio. Flu-dUTP and TAMRA-dUTP (BioSan, Russia) were used as dNTPs carrying reporter groups. For the nick translation reaction, 1 μ g of DNA was used.

Hybridization setup

The FISH reaction was performed according to the following protocol.

- Drop the suspension on to the slides
- Slides were steamed using the water bath at 90°C and dried in an incubator at 37°C.
- Next, the slide was washed again in three changes of 2×SSC for 5 minutes each at 37 °C.
- Incubated slides in a 0.01% pepsin solution for 5 minutes at 37 °C.
- Followed by post-fixation in a 0.5% paraformaldehyde solution in phosphate-buffered saline (PBS) for 10 minutes at room temperature
- Washing the slides in PBS for 5 minutes in room temperature.
- The slide was then dehydrated by passing it through three jars of ethanol with increasing concentrations (70% – 80% – 100%) for 5 minutes each, after which it was dried in an incubator at 37 °C.
- Ten microliters (10 μ L) of the hybridization mixture was applied to the slide.
- DNA probe denaturation for 5 minutes at 75 °C followed by DNA probe hybridization on the slide for 24 hours at 37 °C was performed in a ThermoBrite programmable temperature controlled humid chamber (Abbott Molecular, USA).
- The next day, the slides were washed to remove non-specifically bound DNA probe in a solution of 0.4×SSC, 0.3% NP-40 (Amresco, USA) for 2 minutes at 73 °C
- Then in a solution of 2×SSC, 0.1% NP-40 (Amresco, USA) for 5 minutes at 37 °C.
- The slides were then mounted in Vectashield medium (VectorLab, USA) with DAPI counterstain under coverslips.
- Fluorescent signals were evaluated using an Axio Imager Z2 microscope (Zeiss, Germany) at 630× magnification using optical filters specific for the emission spectra transmission of DAPI, fluorescein, and TAMRA dyes.

2.3.5. Ultra-Low Coverage Whole Genome Sequencing

DNA quality and quantity were analyzed on Nanodrop (Thermo, USA) and Qubit 2.0 (Thermo, USA) spectrophotometers.

Nanodrop (Thermo, USA)

During the initial stages of genomic DNA preparation, the degree of DNA purity from protein contamination and salts is determined by measuring the A260/A280 (1.8–2.0) and A260/A230 (> 1) ratios using a spectrophotometer. Only DNA samples meeting these requirements were used for further steps.

Qubit 2.0 (Thermo, USA) spectrophotometers

- Equilibrate all assay reagents, kit standards, and experimental samples to room temperature (22–28°C) before initiating the assay. Keep all fluorescent target-selective dyes protected from direct light exposure to avoid photobleaching.
- Prepared a fresh batch of working solution for the total number of samples and standards required. Dilute the specific Qubit Reagent (selective fluorophore dye) at a 1:200 ratio into the provided Qubit Buffer following 1 µL of Qubit Reagent into 199 µL of Qubit Buffer per reaction tube.
- Used only clean, thin-wall 0.5 mL polypropylene Real-Time PCR tubes, such as Qubit Assay Tubes (Cat. no. Q32856) to ensure optical compatibility.
- Transfer 190 µL of the freshly prepared Qubit™ Working Solution into two designated tubes, and added exactly 10 µL of Standard 1 and Standard 2 respectively.
- Transfer 5 µL into a clean tube, and add 195 µL of the prepared Qubit Working Solution to bring the final total volume to exactly 200 µL.
- Vortex all standard and sample assay tubes for 2–3 seconds to achieve full homogeneity. Spin the tubes briefly or tap gently against a surface to fully dissipate any minute air bubbles. Incubate the tubes at room temperature (protected from light) for 2 minutes.
- The Qubit® 2.0 Fluorometer initializes automatically when connected to its power line. Using the touchscreen interface, select DNA.
- When prompted by the instrument user interface to read new standards, press Yes. Insert Standard 1 firmly into the Sample Chamber, close the protective chamber lid completely, and tap Read. Repeat this step sequentially for Standard 2. The device generates the standard curve utilizing its built in curve fitting algorithm.
- Insert the prepared sample tube into the Sample Chamber, close the lid, and press Read. The device will evaluate fluorescence intensity and display the concentration within approximately 3 seconds. For all subsequent samples, remove the measured tube, insert the new one, and select Read Next Sample.
- Export the data.

Libraries were prepared using MBU Tn5 transposase (MBS-Technology, Russia) in accordance with the manufacturer's recommendations. Adapter sequences from the Nextera XT DNA Library Preparation Kit (Illumina, USA) were used during preparation.

Library quality was assessed using a 2100 Bioanalyzer (Agilent Technologies, USA). The libraries were sequenced in an FCM V2.0 flow cell on a Genolab M sequencer using the GenoLab M Sequencing Reagent (300 cycles) V2.0 kit (GeneMind, China) according to the manufacturer's instructions.

Preparation and Assembly of Tn5 Transposomes

- In a sterile microcentrifuge tube, combine 600 μ L of 100 μ M Tn5ME-rev/A oligo and 600 μ L of 100 μ M Tn5ME-rev/B oligo.
- Transfer 7.5 mL of 3xFlag-pATn5 transposase (formulated in 50% glycerol) to the oligonucleotide mixture. Gently pipette up and down to ensure complete homogeneity.
- Incubate the mixture at room temperature for approximately 50 minutes to allow the pre-annealed MEDS oligonucleotides to bind to the transposase enzyme.
- Following incubation, immediately transfer the assembled transposome complex to -20 $^{\circ}$ C for long-term storage.

***In Vitro* Tagmentation Activity Assay**

- Transfer the substrate DNA with the assembled Tn5 transposome in a reaction tube and incubate for 7 minutes at 55 $^{\circ}$ C.
- Add 2 μ L of 1% sodium dodecyl sulfate (SDS) directly to each reaction tube to denature the transposase and halt the tagmentation process.
- Incubate the samples for 7 minutes at either room temperature or 55 $^{\circ}$ C.
- Add 0.63 μ L of Proteinase K (20 mg/mL) to each reaction to degrade the remaining 3xFlag-pATn5 protein and fully release the tagmented DNA fragments.
- Incubate the mixture for 7 minutes at 55 $^{\circ}$ C.
- Immediately place the samples on ice or store them at 4 $^{\circ}$ C to completely arrest all enzymatic activity.
- Analyze the resulting DNA fragments by migrating the samples on a 0.7% agarose gel prepared in 1x TAE buffer to evaluate tagmentation efficiency and size distribution.

2.3.6. Bioinformatic Processing

Quality control of the raw sequencing data was performed using FastQC v0.11.4.

The average genome coverage was 0.3x. Read alignment was performed against the GRCh38 (hg38) human reference genome using bwa 0.7.17.

Uniquely aligned reads were obtained and filtered by mapping quality using SAMtools v1.6.

The resulting .bam files were processed using QDNAseq 1.40.0.

For sex chromosomes, GC content correction was calculated using ComputeGCBias (a tool from the deepTools package) employing the Benjamini method [80].

The genome was divided into 50 kb bins, and coverage for these regions was calculated using bedtools v2.30.0 [81].

4. DISCUSSION

In this study, the human reference genome GRCh38 (hg38) was partitioned into contiguous 50 kilobase (kb) regions to calculate localized read-depth profiles. The clinical utility of this cytogenomic approach relies completely on the mathematical thresholds established to distinguish true biological CNVs and numerical aneuploidies from background technical noise, sequencing artifacts, and GC-content biases.

During our optimization protocol, we established relative thresholds of 1.2 and above for the classification of trisomies and 0.8 and below for monosomies. The selection of these specific limits represents a strategic diagnostic compromise between sensitivity and specificity. By lowering the trisomy threshold down to 1.2, we substantially optimized the diagnostic sensitivity of the pipeline, enabling the successful identification of low-level tissue mosaicism (down to a boundary of approximately 20% mosaic clonal representation) and subtle, subchromosomal CNVs. However, this analytical sensitivity inherently carries the mathematical trade-off of inflating the false-positive rate, particularly within genomic regions predisposed to mapping uncertainties or low DNA quality.

The ability to detect genetic variants depends directly on achieving sufficient sample coverage. More importantly, this coverage dictates the minimum size of CNVs that can be reliably identified. While theoretical models suggested that low-coverage sequencing would limit detection to CNVs larger than 1 million base pairs (Mb), practical application demonstrated a significantly higher resolution. Specifically, the chosen methodology successfully identified a submicroscopic variant spanning only 350,000 base pairs, which was subsequently validated via real-time qPCR. Consequently, establishing these empirical detection limits is essential for defining the clinical sensitivity of the assay regarding smaller structural variations.

To evaluate the clinical performance, sensitivity, and specificity of ulcWGS, validation was conducted using interphase FISH and real-time qPCR sub-cohorts. This validation design allowed us to cross-verify structural genomic calls at different levels of resolution. The operational validation demonstrated a high level of technical concordance. Out of 36 samples analyzed through FISH to confirm the ulcWGS findings, 32 samples demonstrated completely concordant results. For the real-time qPCR validation, 16 samples were tested, with 14 samples confirming the exact chromosomal alterations identified by ulcWGS.

The overall sensitivity of the ulcWGS was calculated at 88.5%. Importantly, all discordant findings across the validation cohorts were identified as false positives. A critical analysis of these misclassifications revealed that 50% of the false-positive results were specifically linked to copy number assessments of the X chromosome. Furthermore, all instances of false-positive classification occurred exclusively in samples characterized by the lowest average coverage depth and the highest background noise metrics during ulcWGS. Beyond these technical limitations, a subset of these discordant cases may also be attributed to tissue-specific mosaicism.

These technical limitations underscore the impact of sequencing metrics on diagnostic specificity. The concentration of errors within the X chromosome highlights a known vulnerability in low-coverage sequencing, where sex chromosome dosage anomalies are frequently exacerbated by

mapping artifacts and baseline background noise[83]. The direct correlation between lower depth of coverage, elevated noise, and false-positive calls demonstrates that a strict technical threshold for sample quality control must be established. To mitigate these diagnostic errors, samples falling below the optimal coverage and noise baselines will require mandatory reflex testing using orthogonal methods like qPCR or FISH.

In the analyzed SA tissues, chromosomal aberrations were identified in a significant proportion of cases, with numerical aneuploidies making up the vast majority of the genetic pathology. A precise breakdown of the abnormal karyotypes revealed that autosomal aneuploidies comprised 48.7% of all detected abnormalities, while sex chromosome numerical variations accounted for the remaining 9.4%. Among the autosomal disruptions, sequencing data identified a broad spectrum of imbalances involving chromosomes 4, 7, 9, 10, 13, 14, 15, 16, 18, 20, and 22. Notably, trisomy 22 emerged as the most common individual autosomal aberration within this cohort. The slightly contrasts with typical literature findings identifying trisomy 16 as the most common aberration[38]. But closely followed by recurrent instances of trisomy 16 and trisomy 13. This shift, where trisomy 22 exceeded trisomy 16 by only a single case, likely stems from random sampling variation rather than a statistically significant biological trend. Numerical anomalies affecting the sex chromosomes were exclusively represented by monosomy X (Turner syndrome), which strongly aligns with the common research outcomes.

The combination of multiple aneuploidies in a single organism is a rare event, with the main mechanism being non-disjunction in two meiotic cycles, which can come from either the same parent or different parental origins. In this study comparatively rare chromosomal combinations were identified. Chromosomal combination identified in the first case including the chromosomes 7 and 8 was a significantly rare combination that hadn't been mentioned in previous studies. In case two chromosome 13 and 16 was identified which is also not a common combination in double aneuploidy.

A significant technical boundary encountered during the digital profiling of embryonic tissues is the inherent inability of standard ulcWGS to directly identify balanced polyploid states. To overcome this major limitation and prevent the misclassification of complete genome duplications as normal karyotypes, we integrated an interphase Fluorescence *In Situ* Hybridization (FISH) with a dual-centromeric probe setup. This cytogenetic validation tier successfully identified 15 polyploid cases (7 triploidies and 8 tetraploidies), representing 7.28% of the evaluated sub-cohort. Additionally, this multi-targeted workflow successfully identified highly intricate configurations, such as a sample with three independent structural CNVs located on the X chromosome occurring within an expanded, polyploid genome. It also revealed complex multilineage mosaicism, including a triple-clone mosaic matrix featuring a dominant tetraploid line alongside a minor monosomy X line (mos: 92, XXXX [78] / 45, X [14] / 46, XX [25]).

To determine the reproductive pathogenesis of these subchromosomal alterations, all identified variants were classified in accordance with the American College of Medical Genetics and Genomics (ACMG) guidelines. For variants localized to the X chromosome, sex-specific inheritance patterns, hemizygosity, and dosage compensation mechanisms might affect on the pathogenicity assessment. Out of the 27 unique samples carrying 35 subchromosomal rearrangements, 11 autosomal and three X-linked variants which makes 14 total rearrangements were classified as Pathogenic, establishing a pathogenic CNV (pCNV) rate of 40.0% within the

structural aberration sub-cohort. Calculated across the entire screening cohort, pathogenic CNVs accounted for 4.71% of all spontaneous miscarriages.

This diagnostic yield aligns tightly with contemporary clinical literature. Large-scale genomic studies utilize chromosomal microarray analysis and low-pass or copy-number sequencing to isolate pathogenic imbalances in approximately 2.7% to 5.26% of unexplained first-trimester pregnancy losses [55]. A vast number of subchromosomal alterations were related to chromosome X (40%) reflecting the fragile nature of it. By successfully mapping these pathogenic alterations at an average genome coverage of 0.3×, our results prove that low-depth untargeted whole-genome strategies can reliably match the diagnostic utility. The remaining 21 structural variants were classified as Variants of Uncertain Significance (VUS) or benign variants. Many of these VUS are low-frequency structural modifications or microaberrations lacking comprehensive documentation in embryological databases.

However, 48.06% of SA cohort presented in the study exhibited a completely normal, balanced diploid sequencing profile received by ulcWGS and verified after FISH analysis. This diagnostic discrepancy might affect several factors. Although one maybe, microdeletions or microduplications smaller than the 100–300 kb biological resolution threshold of low-pass whole-genome sequencing escape detection entirely. Additionally, alternative epigenetic mechanisms may actively contribute to the formation of the lethal phenotype. Investigating the precise impact of these non-genetic factors remains a subject for further research.

5. CONCLUSION

The comprehensive molecular screening of 206 SA conducted in this thesis reveals a highly non-random, multi-scale landscape of genomic instability underlying early gestational failure. The structural and numerical chromosomal configurations validated across the study cohort resolved into four distinct profiles. Nearly half of the abortus samples presented a normal chromosomal complement (48.06%). This prominent sub-cohort highlights that while genetic defects dominate first-trimester loss, a large proportion of spontaneous abortions are driven by alternative aetiologies.

Numerical aneuploidies (31.55%) constituted the primary genetic cause of embryonic mortality. Autosomal trisomies predominated, driven by chromosome specific maternal meiotic nondisjunction. Trisomy 22 (n=12) and trisomy 16 (n=11) emerged as the dominant autosomal hyperploidies, confirming that small, gene-dense chromosomes or those with highly permissive spindle assembly checkpoints (SAC) exhibit an increased propensity for meiotic segregation failures. Among sex chromosomes, monosomy X (45, X; n=11) stood as the single most prevalent numerical loss, reflecting a frequent paternal sex-chromosome shedding event during spermatogenesis or an early post-zygotic mitotic error.

These primary sub-cohorts were supplemented by the minor yet significant sub-cohorts of polyploidies (9.71%) and subchromosomal structural variations (6.31%). The identification of pathogenic subchromosomal CNVs, exemplified by a rare, high-consequence 16p13.11 triplication, demonstrates that submicroscopic chromosomal imbalance is a significant, underreported driver of spontaneous abortion.

The ulcWGS pipeline demonstrated exceptional specificity for detecting unbalanced genomic mutations. Following cross-platform calibration, the final integrated pipeline achieved an analytical sensitivity of 88.5%. While theoretical models restrict low-coverage sequencing to CNVs larger than 1 Mb, practical application achieves a significantly higher sub-microscopic resolution. By mapping short sequencing reads into contiguous bins across a reference genome (GRCh38), the system bypasses the core limitations of traditional G-banding. These results confirm that ultra-low coverage genome-wide sequencing is effective in diagnosing CNVs and aneuploidies in the absence of alternative methods.

This thesis demonstrates that while standalone sequencing exhibits distinct structural blind spots, a dual-platform framework coupling ultra-low-coverage whole-genome sequencing (ulcWGS) with fluorescence *in situ* hybridization (FISH) and Real-Time PCR (qPCR) delivers a reliable, scalable, and definitive paradigm for karyotyping spontaneous abortions effectively transitioning reproductive medicine toward precise, data-driven diagnostic clarity.

REFERENCES

1. ACOG Practice Bulletin No. 200: Early Pregnancy Loss. *Obstet Gynecol*, 2018. **132**(5): p. e197-e207.
2. Essers, R., et al., *Prevalence of chromosomal alterations in first-trimester spontaneous pregnancy loss*. *Nat Med*, 2023. **29**(12): p. 3233-3242.
3. Smolander, J., et al., *Evaluation of tools for identifying large copy number variations from ultra-low-coverage whole-genome sequencing data*. *BMC Genomics*, 2021. **22**(1): p. 357.
4. Lund, R.J., et al., *High-throughput karyotyping of human pluripotent stem cells*. *Stem Cell Research*, 2012. **9**(3): p. 192-195.
5. WHO, *INTERNATIONAL CLASSIFICATION OF DISEASES MANUAL OF THE INTERNATIONAL STATISTICAL CLASSIFICATION OF DISEASES, INJURIES, AND CAUSES OF DEATH Based on the Recommendations of the Ninth Revision Conference, 1975, and Adopted by the Twenty-ninth World Health Assembly 1975*.
6. Quenby, S., et al., *Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss*. *The Lancet*, 2021. **397**(10285): p. 1658-1667.
7. SIMPSON, J.L., *Causes of Fetal Wastage*. *Clinical Obstetrics and Gynecology*, 2007. **50**(1): p. 10-30.
8. Chatenoud, L., et al., *Paternal and Maternal Smoking Habits before Conception and During the First Trimester: Relation to Spontaneous Abortion*. *Annals of Epidemiology*, 1998. **8**(8): p. 520-526.
9. Moradinazar, M., et al., *Epidemiological Features of Spontaneous Abortion in the North Africa and the Middle East From 1990 to 2019: Results From the Global Burden of Disease Study 2019*. *J Family Reprod Health*, 2022. **16**(3): p. 183-191.
10. Øvrebø, J.I. and B.A. Edgar, *Polyploidy in tissue homeostasis and regeneration*. *Development*, 2018. **145**(14).
11. Excellence., N.I.f.H.a.C., *Ectopic pregnancy and miscarriage: diagnosis and initial management in early pregnancy of ectopic pregnancy and miscarriage*. NICE Clinical Guideline 2012.
12. Magnus, M.C., et al., *Role of maternal age and pregnancy history in risk of miscarriage: prospective register based study*. *Bmj*, 2019. **364**: p. l869.
13. Bender Atik, R., et al., *ESHRE guideline: recurrent pregnancy loss: an update in 2022*. *Hum Reprod Open*, 2023. **2023**(1): p. hoad002.
14. Moustakli, E., et al., *Immunological Factors in Recurrent Pregnancy Loss: Mechanisms, Controversies, and Emerging Therapies*. *Biology (Basel)*, 2025. **14**(7).
15. Garmendia, J.V., et al., *Exploring the Immunological Aspects and Treatments of Recurrent Pregnancy Loss and Recurrent Implantation Failure*. *International Journal of Molecular Sciences*, 2025. **26**(3): p. 1295.
16. Wesselink, A.K., et al., *Exposure to Ambient Heat and Risk of Spontaneous Abortion: A Case-Crossover Study*. *Epidemiology*, 2024. **35**(6): p. 864-873.
17. Thomson, K., et al., *Socioeconomic inequalities and adverse pregnancy outcomes in the UK and Republic of Ireland: a systematic review and meta-analysis*. *BMJ Open*, 2021. **11**(3): p. e042753.

18. Shetty, A., et al., *Global prevalence of post-miscarriage anxiety, depression, and stress: a systematic review and meta-analysis*. J Glob Health, 2025. **15**: p. 04245.
19. Thowfeequ, S. and S. Srinivas, *Embryonic and extraembryonic tissues during mammalian development: shifting boundaries in time and space*. Philos Trans R Soc Lond B Biol Sci, 2022. **377**(1865): p. 20210255.
20. Chen, K., et al., *The diagnostic accuracy of preimplantation genetic testing (PGT) in assessing the genetic status of embryos: a systematic review and meta-analysis*. Reprod Biol Endocrinol, 2025. **23**(1): p. 39.
21. Elizabeth J. Herrick, B.B., *Embryology, Placenta*, in *StatPearls* 2026, StatPearls Publishing.
22. Montero., J.M.F.J., *Chorionic Villus Sampling*, in *StatPearls*. 2026, StatPearls Publishing.
23. Niu, B., et al., *Deciphering signaling mechanisms and developmental dynamics in extraembryonic mesoderm specification from hESCs*. Nat Commun, 2025. **16**(1): p. 4688.
24. Srebniak, M.I., et al., *Chorionic Villus Sampling for Rapid Confirmation of High-Risk NIPT Results for Trisomy 21, 18, and 13*. Prenat Diagn, 2025. **45**(8): p. 988-993.
25. Flowers, N.J., et al., *Assessment of Placental Chromosomal Mosaicism during Prenatal Cell-Free DNA Screening Refines Positive Predictive Values for Fetal Trisomy*. Clinical Chemistry, 2025. **71**(10): p. 1036-1046.
26. Goh, I., et al., *Yolk sac cell atlas reveals multiorgan functions during human early development*. Science, 2023. **381**(6659): p. eadd7564.
27. Dewald O, H.J., *Gestational Sac Evaluation.*, in *StatPearls*. 2026, StatPearls Publishing.
28. Suguna, B. and K. Sukanya, *Yolk sac size & shape as predictors of first trimester pregnancy outcome: A prospective observational study*. Journal of Gynecology Obstetrics and Human Reproduction, 2019. **48**(3): p. 159-164.
29. Zheng, Y., et al., *Single-cell analysis of embryoids reveals lineage diversification roadmaps of early human development*. Cell Stem Cell, 2022. **29**(9): p. 1402-1419.e8.
30. Aygun, N., *Acquired Chromosomal Abnormalities and Their Potential Formation Mechanisms in Solid Tumours*, in *Chromosomal Abnormalities - A Hallmark Manifestation of Genomic Instability*, M.L. Larramendy and S. Soloneski, Editors. 2017, IntechOpen: London.
31. Kou, F., et al., *Chromosome Abnormalities: New Insights into Their Clinical Significance in Cancer*. Molecular Therapy - Oncolytics, 2020. **17**: p. 562-570.
32. Canoy, R.J., et al., *Factors That Affect the Formation of Chromosomal Translocations in Cells*. Cancers, 2022. **14**(20): p. 5110.
33. Vervoort, L. and J.R. Vermeesch, *Low copy repeats in the genome: from neglected to respected*. Exploration of Medicine, 2023. **4**(2): p. 166-175.
34. Sharakhov, I.V. and M.V. Sharakhova, *Chromosomal inversions and their impact on insect evolution*. Current Opinion in Insect Science, 2024. **66**: p. 101280.
35. Yip, M.-Y., *Autosomal ring chromosomes in human genetic disorders*. Translational Pediatrics, 2015. **4**(2): p. 164-174.
36. de Leval, L., et al., *Genomic profiling for clinical decision making in lymphoid neoplasms*. Blood, 2022. **140**(21): p. 2193-2227.
37. Orr, B., K.M. Godek, and D. Compton, *Aneuploidy*. Current Biology, 2015. **25**(13): p. R538-R542.

38. Jia, C.W., et al., *Aneuploidy in Early Miscarriage and its Related Factors*. Chin Med J (Engl), 2015. **128**(20): p. 2772-6.
39. Gu, C., et al., *Chromosomal Aneuploidy Associated With Clinical Characteristics of Pregnancy Loss*. Frontiers in Genetics, 2021. **Volume 12 - 2021**.
40. Gruhn, J.R., et al., *Chromosome errors in human eggs shape natural fertility over reproductive life span*. Science, 2019. **365**(6460): p. 1466-1469.
41. Thomas, C., T. Cavazza, and M. Schuh, *Aneuploidy in human eggs: contributions of the meiotic spindle*. Biochem Soc Trans, 2021. **49**(1): p. 107-118.
42. Pierce, B.A., *Genetics essentials: concepts and connections*. 2018.
43. Vázquez-Diez, C. and G. FitzHarris, *Causes and consequences of chromosome segregation error in preimplantation embryos*. Reproduction, 2018. **155**(1): p. R63-r76.
44. Soler, A., et al., *Overview of Chromosome Abnormalities in First Trimester Miscarriages: A Series of 1,011 Consecutive Chorionic Villi Sample Karyotypes*. Cytogenet Genome Res, 2017. **152**(2): p. 81-89.
45. Shahbazi, M.N., et al., *Developmental potential of aneuploid human embryos cultured beyond implantation*. Nature Communications, 2020. **11**(1): p. 3987.
46. Antonarakis, S.E., et al., *Down syndrome*. Nat Rev Dis Primers, 2020. **6**(1): p. 9.
47. Muskens, I.S., et al., *The genome-wide impact of trisomy 21 on DNA methylation and its implications for hematopoiesis*. Nature Communications, 2021. **12**(1): p. 821.
48. Uygur Külcü, N., et al., *Trisomy 18 and Trisomy 13: A Retrospective Cohort Study at a Tertiary Hospital*. Children (Basel), 2026. **13**(2).
49. Hook, E.B. and D. Warburton, *Turner syndrome revisited: review of new data supports the hypothesis that all viable 45,X cases are cryptic mosaics with a rescue cell line, implying an origin by mitotic loss*. Human Genetics, 2014. **133**(4): p. 417-424.
50. Militaru, M.S., et al., *The Impact of Chromosomal Mosaicisms on Prenatal Diagnosis and Genetic Counseling-A Narrative Review*. J Pers Med, 2024. **14**(7).
51. Shukla, V., M.K. Høffding, and E.R. Hoffmann, *Genome diversity and instability in human germ cells and preimplantation embryos*. Seminars in Cell & Developmental Biology, 2021. **113**: p. 132-147.
52. Salmina, K., et al., *The Cancer Aneuploidy Paradox: In the Light of Evolution*. Genes, 2019. **10**(2): p. 83.
53. Toufaily, M.H., et al., *Triploidy: Variation of Phenotype*. American Journal of Clinical Pathology, 2016. **145**(1): p. 86-95.
54. Zaragoza, M.V., et al., *Parental Origin and Phenotype of Triploidy in Spontaneous Abortions: Predominance of Diandry and Association with the Partial Hydatidiform Mole*. The American Journal of Human Genetics, 2000. **66**(6): p. 1807-1820.
55. Rosenbusch, B.E., *Mechanisms giving rise to triploid zygotes during assisted reproduction*. Fertility and Sterility, 2008. **90**(1): p. 49-55.
56. Heslop-Harrison, J.S., *Polyploidy*, in *Brenner's Encyclopedia of Genetics (Second Edition)*, S. Maloy and K. Hughes, Editors. 2013, Academic Press: San Diego. p. 402-403.
57. Yatsenko, S.A., et al., *Cytogenetic signatures of recurrent pregnancy losses*. Prenatal Diagnosis, 2021. **41**(1): p. 70-78.
58. Bai, W., et al., *Analysis of copy number variations and possible candidate genes in spontaneous abortion by copy number variation sequencing*. Frontiers in Endocrinology, 2023. **Volume 14 - 2023**.

59. Wu, H., et al., *Analysis of Genomic Copy Number Variation in Miscarriages During Early and Middle Pregnancy*. Frontiers in Genetics, 2021. **Volume 12 - 2021**.
60. Bai, H., C. Zhang, and J. Shi, *Case Report: Discordant genetic findings in miscarriage tissue following transfer of euploid embryos*. Frontiers in Reproductive Health, 2026. **Volume 8 - 2026**.
61. Li, F.X., et al., *Detection of chromosomal abnormalities in spontaneous miscarriage by low-coverage next-generation sequencing*. Mol Med Rep, 2020. **22**(2): p. 1269-1276.
62. Carvalho, C.M.B. and J.R. Lupski, *Mechanisms underlying structural variant formation in genomic disorders*. Nature Reviews Genetics, 2016. **17**(4): p. 224-238.
63. Miranda Durkie, E.-J.C., Ian Berry, Martina Owens, Clare Turnbull, Richard H Scott, Robert W Taylor, Zandra C Dean, Sian Ellard, Emma L Baple and Dominic J McMullan, *ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2023*. ACMG guidelines, 2023: p. 56.
64. Abhari, S. and J.F. Kawwass, *Pregnancy and Neonatal Outcomes after Transfer of Mosaic Embryos: A Review*. Journal of Clinical Medicine, 2021. **10**(7): p. 1369.
65. Martin, A., et al., *Mosaic results after preimplantation genetic testing for aneuploidy may be accompanied by changes in global gene expression*. Frontiers in Molecular Biosciences, 2023. **Volume 10 - 2023**.
66. Budrewicz, J. and S.L. Chavez, *Insights into embryonic chromosomal instability: mechanisms of DNA elimination during mammalian preimplantation development*. Frontiers in Cell and Developmental Biology, 2024. **Volume 12 - 2024**.
67. Grati, F.R., *Chromosomal Mosaicism in Human Feto-Placental Development: Implications for Prenatal Diagnosis*. Journal of Clinical Medicine, 2014. **3**(3): p. 809-837.
68. Chuang, T.H., et al., *The Incidence of Mosaicism for Individual Chromosome in Human Blastocysts Is Correlated With Chromosome Length*. Front Genet, 2020. **11**: p. 565348.
69. Zuev A.S., K.P.M.B., Fonova E.A., Shevtsov D.G., Babay T.S., Nikitina T.V., Fedotov D.A., Sazhenova E.A., Tolmacheva E.N., Vasiliev S.A., *Evaluation of the efficiency of full-genome sequencing for karyotyping of spontaneous abortus cells with no proliferative activity*. . Medical Genetics, 2024. **2024;23**(12): p. 30-36.
70. Popescu-Hobeanu, G., et al., *Cytogenetic Analysis of Sporadic First-Trimester Miscarriage Specimens Using Karyotyping and QF-PCR: A Retrospective Romanian Cohort Study*. Genes (Basel), 2022. **13**(12).
71. Sahoo, T., et al., *Comprehensive genetic analysis of pregnancy loss by chromosomal microarrays: outcomes, benefits, and challenges*. Genet Med, 2017. **19**(1): p. 83-89.
72. Wang, B.T., et al., *Abnormalities in spontaneous abortions detected by G-banding and chromosomal microarray analysis (CMA) at a national reference laboratory*. Mol Cytogenet, 2014. **7**: p. 33.
73. Levy, B., et al., *Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis*. Obstet Gynecol, 2014. **124**(2 Pt 1): p. 202-209.
74. Dai, Y.-F., et al., *Experience of copy number variation sequencing applied in spontaneous abortion*. BMC Medical Genomics, 2024. **17**(1): p. 15.
75. Chen, Y., et al., *Copy number variation sequencing for the products of conception: What is the optimal testing strategy*. Clinica Chimica Acta, 2024. **557**: p. 117884.

76. Geng, Q., et al., *Screening of triploid with low-coverage whole-genome sequencing by a single-nucleotide polymorphism-based test in miscarriage tissue*. Journal of Assisted Reproduction and Genetics, 2019. **36**(12): p. 2525-2531.
77. McCombie, W.R., J.D. McPherson, and E.R. Mardis, *Next-Generation Sequencing Technologies*. Cold Spring Harb Perspect Med, 2019. **9**(11).
78. Li, S., et al., *Ultra-low coverage genome-wide association study – insights into gestational age using 17,844 embryo samples with preimplantation genetic testing*. medRxiv, 2022: p. 2022.06.15.22276464.
79. Stepanov., A.o.t.R.A.o.S.V.A., *Guidelines for medical technologies for molecular diagnostics of chromosomal and hereditary diseases* 2022: p. 234.
80. Benjamini, Y. and T.P. Speed, *Summarizing and correcting the GC content bias in high-throughput sequencing*. Nucleic Acids Res, 2012. **40**(10): p. e72.
81. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features*. Bioinformatics, 2010. **26**(6): p. 841-2.
82. Murry, J.B. and Y.S. Zou, *Live-Born Double Aneuploidy at the Johns Hopkins Cytogenomics Laboratory: Case Report and Review of the Literature*. OBM Genetics, 2022. **06**(04): p. 168.
83. Webster, T.H., et al., *Identifying, understanding, and correcting technical artifacts on the sex chromosomes in next-generation sequencing data*. GigaScience, 2019. **8**(7).



АНТИПЛАГИАТ
ОБНАРУЖЕНИЕ ЗАИМСТВОВАНИЙ

СПРАВКА

о результатах проверки текстового документа
на наличие заимствований

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
"НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ"

ПРОВЕРКА ВЫПОЛНЕНА В СИСТЕМЕ АНТИПЛАГИАТ.ВУЗ

Автор работы: Канканам Патиранаге Манит Банула
Самоцитирование
рассчитано для: Канканам Патиранаге Манит Банула
Название работы: ВКР Канканам Патиранаге Манит Банула
Тип работы: Выпускная квалификационная работа
Подразделение: САЕ ИУМТ

РЕЗУЛЬТАТЫ

■ ОТЧЕТ О ПРОВЕРКЕ КОРРЕКТИРОВАЛСЯ: НИЖЕ ПРЕДСТАВЛЕНЫ РЕЗУЛЬТАТЫ ПРОВЕРКИ ДО КОРРЕКТИРОВКИ

СОВПАДЕНИЯ 12.94%
ОРИГИНАЛЬНОСТЬ 87.06%
ЦИТИРОВАНИЯ 0%
САМОЦИТИРОВАНИЯ 0%
ИИ-КОНТЕНТ 0%

СОВПАДЕНИЯ 12.94%
ОРИГИНАЛЬНОСТЬ 87.06%
ЦИТИРОВАНИЯ 0%
САМОЦИТИРОВАНИЯ 0%

ДАТА И ВРЕМЯ КОРРЕКТИРОВКИ: 05.06.2026 10:08

ДАТА ПОСЛЕДНЕЙ ПРОВЕРКИ: 05.06.2026

Структура документа: Проверенные разделы: основная часть с.4, 6-7, 26-28, 43-63, титульный лист с.1-2, содержание с.3, 5, введение с.8-26, методы с.32-43, выводы с.64-67
Модули поиска: Профессиональная лексика. Юриспруденция; Переводные заимствования; Профессиональная лексика. АПК и биотех; Цитирование; Шаблонные фразы; ИПС Адилет; Профессиональная лексика. Медицина; Сводная коллекция научных работ Беларуси; СПС ГАРАНТ: аналитика; Публикации eLIBRARY; СМИ России и СНГ; Патенты СССР, РФ, СНГ; Коллекция НБУ; IEEE; Переводные заимствования по базе публикаций открытого доступа PubMed; PubMed; Перефразирования по СПС ГАРАНТ: аналитика; СПС ГАРАНТ: нормативно-правовая документация; Кольцо вузов; Публикации РГБ; Переводные заимствования IEEE; Переводные заимствования по Коллекции открытых публикаций международных издательств; Сводная коллекция ЭБС; Публикации РГБ (переводы и перефразирования); Кольцо вузов (переводы и перефразирования); Перефр...

Работу проверил: Воронова Гульнара Альфридовна

ФИО проверяющего

Дата подписи:

05.06.2026

Подпись проверяющего



Чтобы убедиться
в подлинности справки, используйте QR-код,
который содержит ссылку на отчет.

Ответ на вопрос, является ли обнаруженное заимствование
корректным, система оставляет на усмотрение проверяющего.
Предоставленная информация не подлежит использованию
в коммерческих целях.