DISSERTATIONES
MEDICINAE
UNIVERSITATIS
TARTUENSIS
326

HINDREK TEDER

Developing computational methods and workflows for targeted and whole-genome sequencing based non-invasive prenatal testing

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
326

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS
326

HINDREK TEDER

Developing computational methods and workflows for targeted and whole-genome sequencing based non-invasive prenatal testing

UNIVERSITY OF TARTU
Press

Institute of Clinical Medicine, Faculty of Medicine, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy (Medicine) on March 16th, 2022 by the Council of the Faculty of Medicine, University of Tartu, Estonia.

Supervisors Associate Professor Priit Palta (PhD),
Estonian Genome Center, Institute of Genomics,
University of Tartu, Tartu, Estonia

Associate Professor Kaarel Krjutškov (PhD),
Institute of Clinical Medicine,
University of Tartu, Tartu, Estonia

Professor Andres Salumets (PhD),
Institute of Clinical Medicine,
University of Tartu, Tartu, Estonia

Reviewed Lecturer Kai Muru (MD, PhD),
Department of Clinical Genetics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia

Associate Professor Sander Pajusalu (MD, PhD),
Department of Clinical Genetics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia

Opponent Associate Professor Erik Sistermans (PhD),
Amsterdam Reproduction & Development,
Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Commencement: June 10, 2022

ISSN 1024-395X
ISBN 978-9949-03-884-8 (print)
ISBN 978-9949-03-885-5 (pdf)

Copyright: Hindrek Teder, 2022

University of Tartu Press
www.tyk.ee

CONTENTS

1. LIST OF ORIGINAL PUBLICATIONS ... 7

2. ABBREVIATIONS ... 8

3. INTRODUCTION ... 9

4. REVIEW OF THE LITERATURE ... 11
  4.1. Origin of fetal aneuploidy ... 11
    4.1.1. Parental origin ... 11
    4.1.2. Maternal age effect ... 13
  4.2. Common fetal aneuploidies ... 15
    4.2.1. Trisomy 21 – Down syndrome ... 15
    4.2.2. Trisomy 18 – Edwards syndrome ... 16
    4.2.3. Trisomy 13 – Patau syndrome ... 17
    4.2.4. Monosomy X – Turner syndrome ... 18
  4.3. Prenatal screening and diagnosis ... 18
    4.3.1. First trimester combined test ... 20
    4.3.2. Prenatal screening and diagnosis in Estonia ... 21
  4.4. Cell-free DNA ... 22
  4.5. Non-invasive prenatal testing ... 24
    4.5.1. Whole-genome sequencing approach ... 25
    4.5.2. Targeted sequencing approach ... 26
    4.5.3. SNP-based targeted sequencing approach ... 26
  4.6. Hidden Markov model ... 27
  4.7. Summary of the literature ... 28

5. AIMS OF THE STUDY ... 29

6. MATERIAL AND METHODS ... 30
  6.1. Study subjects ... 30
    6.1.1. *In vitro* (Papers I–II) ... 30
    6.1.2. *In silico* (Paper II) ... 30
    6.1.3. Clinical patients (Paper III) ... 31
  6.2. Methods ... 32
    6.2.1. Allelic ratio (Paper II) ... 32
    6.2.2. Hidden Markov model (Paper II) ... 33
    6.2.3. Supplemental machine learning methods (Paper II) ... 35
    6.2.4. Fetal fraction estimation for targeted sequencing approach (Paper II) ... 35
    6.2.5. Fetal fraction estimation for whole-genome sequencing approach (Paper III) ... 36
    6.2.6. Detection of fetal trisomies (Paper III) ... 36
    6.2.7. Statistical analysis ... 36
  6.3. Ethics ... 37

7. RESULTS 38
7.1. Paper I 38
7.1.1. TAC-seq assay 38
7.1.1.1. Target selection and probe design 39
7.1.1.2. Library preparation and sequencing 39
7.1.1.3. Data preprocessing and analysis 40
7.1.2. Experimental evaluation of TAC-seq 40
7.1.3. Detection of trisomy using TAC-seq assay 41
7.2. Paper II 43
7.2.1. Computational NIPT analysis framework for targeted sequencing assays 43
7.2.1.1. Locus classification 44
7.2.1.2. Chromosome classification 45
7.2.2. *In vitro* analysis 48
7.3. Paper III 50
7.3.1. Analysis pipeline 50
7.3.2. Fetal fraction analysis 52
7.3.3. Fetal aneuploidy analysis 52
7.3.4. Fetal sex determination 53
8. DISCUSSION 54
9. CONCLUSION 57
10. SUMMARY IN ESTONIAN 58
11. REFERENCES 64
12. ACKNOWLEDGMENTS 74
13. PUBLICATIONS 75
14. CURRICULUM VITAE 118
15. ELULOOKIRJELDUS 120

1. LIST OF ORIGINAL PUBLICATIONS

Paper I **Teder, H.**, Koel, M., Paluoja, P., Jatsenko, T., Rekker, K., Laisk-Podar, T., Kukuškina, V., Velthut-Meikas, A., Fjodorova, O., Peters, M., Kere, J., Salumets, A., Palta, P., & Krjutškov, K. (2018). TAC-seq: targeted DNA and RNA sequencing for precise biomarker molecule counting. *Npj Genomic Medicine*, *3*(1), 34. https://doi.org/10.1038/s41525-018-0072-5

Paper II **Teder, H.**, Paluoja, P., Rekker, K., Salumets, A., Krjutškov, K., & Palta, P. (2019). Computational framework for targeted high-coverage sequencing based NIPT. *PLOS ONE*, *14*(7), e0209139. https://doi.org/10.1371/journal.pone.0209139

Paper III Žilina, O., Rekker, K., Kaplinski, L., Sauk, M., Paluoja, P., **Teder, H.**, Ustav, E., Tõnisson, N., Reimand, T., Ridnõi, K., Palta, P., Vermeesch, J. R., Krjutškov, K., Kurg, A., & Salumets, A. (2019). Creating basis for introducing non-invasive prenatal testing in the Estonian public health setting. *Prenatal Diagnosis*, *39*(13), 1262–1268. https://doi.org/10.1002/pd.5578

Contribution of the author to the original publications:

Paper I Developed the TAC-seq oligonucleotide probe design and TAC-seq sequencing data analysis software tools, *in silico* designed the oligonucleotide targeting probes, co-performed the data analyses, co-prepared the figures, co-wrote the first draft, participated in manuscript preparation.

Paper II Participated in the study design, developed the computational data analysis methods and framework, performed the *in silico* simulated and *in vitro* generated data analyses, prepared the figures and tables, wrote the first draft, participated in manuscript preparation.

Paper III Co-developed the data analysis workflow for production, co-performed the validation study data processing and analyses. Participated in manuscript preparation.

2. ABBREVIATIONS

45,X	monosomy X
AC	amniocentesis
AR	allelic ratio
BMI	body mass index
cfDNA	cell-free DNA
cffDNA	cell-free fetal DNA
chr	chromosome
CI	confidence interval
CVS	chorionic villus sampling
DANSR	Digital Analysis of Selected Regions
DR	detection rate
DT	decision tree
FA	fetal allele
FCT	first trimester combined test
FORTE	Fetal Fraction Optimized Risk of Trisomy Evaluation
FPR	false positive rate
hCG	human chorionic gonadotropin
HMM	hidden Markov model
MA	maternal allele
MAX_i	major allele count of locus i
MI	meiosis I
MII	meiosis II
MIN_i	minor allele count of locus i
ML	machine learning
NATUS	Next-generation Aneuploidy Test Using SNPs
NIPT	non-invasive prenatal testing
NT	nuchal translucency
PAPP-A	pregnancy-associated plasma protein A
PCR	polymerase chain reaction
PPV	positive predictive value
PZM	post-zygotic mitosis
RC	read count
RCAR	read count and allelic ratio
SD	standard deviation
SNP	single-nucleotide polymorphisms
SVM	support vector machine
T13	trisomy 13
T18	trisomy 18
T21	trisomy 21
TAC-seq	Targeted Allele Counting by sequencing
UMI	unique molecular identifiers
WGS	whole-genome sequencing
β-hCG	human β subunit of chorionic gonadotropin

3. INTRODUCTION

Over the past half-century, the field of medicine has witnessed a series of remarkable advances in prenatal screening and diagnostics. It is a result of scientific discoveries and technological development, which has led to new and improved methods for fetal aneuploidy detection. Even though amniocentesis (AC) and fetal ultrasound were first used in the 1950s and late 1960s, respectively, prenatal screening and diagnostic methods were not accessible for the majority of pregnant women at first.

During the 1970s and early 1980s, advanced maternal age – commonly defined as over 35 years of age – was the only indication why the fetal chromosomal abnormalities were assessed in the general population of pregnant women. Less than 1/3 of Down syndrome cases were prenatally diagnosed in these early days. Only about 2% of those who followed invasive prenatal procedures had fetal karyotype abnormalities, which is comparable to the chance of procedure-related miscarriage with AC or chorionic villus sampling (CVS).

The introduction of second trimester maternal serum screening in the form of "double", "triple", and "quadruple" marker testing in the late 1980s and early 1990s significantly improved the accuracy of fetal aneuploidy detection. The proportion of prenatally diagnosed Down syndrome cases more than doubled, with chromosomal abnormalities discovered in as many as 4% of cases with positive screening results.

In the late 1990s and early 2000s, the introduction of the first trimester combined test (FCT), which uses ultrasound measurement of nuchal translucency (NT) thickness along with concentrations of human β subunit of chorionic gonadotropin (β-hCG) and pregnancy-associated plasma protein-A (PAPP-A) in maternal serum as risk markers, shifted fetal aneuploidy screening from the second trimester to the first trimester. As a result, screening performance has improved such that more than 9/10 of Down syndrome cases are prenatally diagnosed, and the yield from invasive testing has risen nearly to 6%, still indicating that only a small number of pregnant women screened as high risk of fetal chromosomal aberration are confirmed by invasive prenatal diagnosis. FCT has been widely adopted as the primary prenatal screening test in many countries, including Estonia.

The latest addition to prenatal screening is non-invasive prenatal testing (NIPT), which uses cell-free DNA (cfDNA) from maternal blood sample to detect fetal aneuploidy. In 1997, Lo et al. first reported that pregnant women's blood sample includes fetal DNA, a founding discovery for cfDNA-based prenatal screening. Cell-free fetal DNA (cffDNA) constitutes about 10% of total cfDNA and is thought to be placental in origin, resulting from apoptotic trophoblasts. cffDNA fragments are typically small, approximately 150 bp in length, and can be detected as early as four weeks of gestation. Importantly, cffDNA can be used to detect common fetal trisomies, such as trisomy 21

(T21), 18 (T18), and 13 (T13), also rare aneuploidies of other chromosomes, and even sub-chromosomal aberrations as the entire fetal genome is represented by the fragments and can be quantified by sequencing. NIPT has shown exceptional performance in detecting fetal aneuploidy compared to the conventional prenatal screening methods previously mentioned. Even though NIPT has demonstrated high sensitivity and specificity for common fetal trisomies, it is still a screening test and, therefore, positive test result needs to be confirmed by a follow-up invasive diagnostic procedure.

The aim of the doctoral thesis was to develop a NIPT data analysis workflow and computational framework for targeted sequencing data. In addition, we implemented and validated a low-coverage whole genome based NIPT analysis pipeline for Estonian prenatal clinical care to detect common fetal aneuploidies.

4. REVIEW OF THE LITERATURE

4.1. Origin of fetal aneuploidy

In the early 1970s, cytogenetic studies of human oocytes made the search for the source of chromosomal aneuploidy possible (Pellestor et al., 2005). The primary causes of spontaneous miscarriages and congenital abnormalities in humans are known to be chromosomal aneuploidies (Jia et al., 2015). Chromosomal abnormalities affect about 30% of eggs (Wartosch et al., 2021). In addition, trisomic or monosomic pregnancies account for at least 10% of all clinically diagnosed pregnancies (Nagaoka et al., 2012). Many aneuploid conceptions are thought to be eliminated and thus miscarried during the first stages of pregnancy (Nagaoka et al., 2012). The majority of aneuploidies are caused by errors originating during oocyte maturation (Nagaoka et al., 2012). The main differences between female and male gametogenesis, resulting in many more aneuploidies in oocytes rather than in sperm cells, are related to the timing of cell division and the number of cells produced from one parent cell (Hassold et al., 2007).

The fundamental mechanisms of chromosomal missegregation errors that result in aneuploidy or mosaicism are non-disjunction and anaphase lag. Chromosomal non-disjunction is a condition in which chromosomes in meiosis I (MI) or sister chromatids in meiosis II (MII) fail to separate equally into the daughter cells. As a consequence, one cell has two chromosomes or chromatids, whereas the other cell has none (Lamb et al., 2005). Chromosome segregation errors are more prevalent during MI division compared to MII in human oocytes. Moreover, the incidence of errors is chromosome-dependent (Wartosch et al., 2021). Some non-disjunction mechanisms are responsible for errors that are associated with failure to crossover, others with crossovers that occur too close or too far from the centromere. Also, some mechanisms are caused by anomalies in other meiotic processes, such as loss of sister chromatid cohesion or defects in spindle assembly/disassembly (Hassold et al., 2007; Sherman et al., 2006).

4.1.1. Parental origin

More than 90% of aneuploidies are thought to be caused by maternal meiotic errors, as illustrated in Table 1 (Hassold et al., 2007). The significantly increased proportion of maternal mistakes has been attributed to several factors. Most of the information and cases are based on T21, the most common and studied autosomal trisomy (Loane et al., 2013). These include an unusually long MI division during oogenesis, which spans from fetal development to menopause (Hassold et al., 2007), and a stronger stringency in the removal of T21 cells during fetal testicular development than ovarian development (Iwarsson et al., 2015). Another idea suggests that female embryos already contain a trisomic cell population in their ovaries with T21 (Hultén et al., 2010).

Because paternal aneuploidy is thought to account for less than 10% of aneuploidy, there have been significantly fewer investigations on the subject to understand its paternal origin (Hassold & Hunt, 2001). Male gametes are continually created from male germ cells that are actively mitotically dividing before entering meiosis in mature adults. As a result, male gametes are subjected to far less temporal stress, and sperm production is maintained throughout lifetime. The majority of aneuploidies in male gametes are connected to sex chromosomes since during male meiosis, XY chromosomes with only a small region of homology must pair and segregate, whereas in female meiosis, the XX chromosomal pairing does not pose any additional obstacles (Hassold & Hunt, 2001). Similar to the maternal origin, non-disjunction is the leading cause of meiotic segregation errors in human sperm, followed by anaphase lag (Uroz & Templado, 2012).

Despite the unbalanced contribution of maternal origin, the parental origin is discovered to be chromosome-dependent. In cases of pregnancy loss or in embryos created using in vitro fertilization and undergoing preimplantation genetic testing for aneuploidies, distinguishing parental origin can be crucial. Men with female partners who experience recurrent pregnancy loss have been found to have higher sperm aneuploidy than controls (Ramasamy et al., 2015). Furthermore, there seem to be more chromosomal errors originating from the MI division of oocytes for T21 and T13, while more errors are happening in the MII division of oocytes for T18. The incidence of post-zygotic mitotic errors, usually manifested as a mosaic form of trisomies, tends to be relatively small (less than 10%) compared to trisomies of meiotic origin (Table 1).

Table 1. Origin of the most common trisomies. T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy 13; MI – meiosis I; MII – meiosis II; PZM – post-zygotic mitosis. Adapted from (Hassold et al., 2007).

Trisomy	Cases	Maternal		Paternal		PZM (%)
		MI (%)	MII (%)	MI (%)	MII (%)	
T21	782	69.6	23.6	1.7	2.3	2.7
T18	150	33.3	58.7	0.0	0.0	8.0
T13	74	56.6	33.9	2.7	5.4	1.4

4.1.2. Maternal age effect

According to studies of trisomies, the incidence and risk of aneuploidy relate to increasing maternal age. Because of the considerable time period between meiotic arrest in the fetus and ovulation in the adult woman, maternal age can influence the occurrence of aneuploidy (Hassold & Hunt, 2001). Recombination mistakes in early meiosis, a deficient spindle assembly checkpoint, deterioration of the connection between the sister chromatids with age, and maternal "genetic age" are all theories for age-related aneuploidies (Chiang et al., 2012). The impact of maternal age is significant – trisomic pregnancies account for 2% of all pregnancies among women under the age of 25, whereas they account for 35% of all conceptions among women over 40 years of age. Many fetuses with chromosomal abnormalities are eliminated during pregnancy (Figure 1). There is no known evidence that race, geographical region, or socioeconomic status may have an impact on maternal age-specific trisomy rates (Hassold & Hunt, 2001).

The vast majority of aneuploidies resulting from maternal errors are meiotic and thus age-dependent (Chiang et al., 2012). This is not valid for all trisomy cases, as non-disjunction is a complicated mechanism influenced by both age-dependent and age-independent factors and events, and thus fetal trisomies may also occur in younger mothers (Ghosh et al., 2010). The two-step mechanism explains how these two are related. The first step is the formation of a susceptible chiasmatic structure in the fetal oocyte, which is age-independent. The second step is aberrant processing of the vulnerable bivalent (i.e., the connected pair of homologous chromosomes), occurring at MI, dependent on the mother's age (Muller et al., 2000). Therefore, the non-disjunctional process occurs in both young and old women. It simply happens more frequently as people become older, potentially because of age-related degradation of the cell cycle or meiotic proteins involved in sister chromatid cohesion or homologue segregation (Muller et al., 2000).

After establishing a clear association between maternal aneuploidy and advanced maternal age, no conclusive connection between paternal age and the frequency of disomic sperm has been established (Templado et al., 2013).

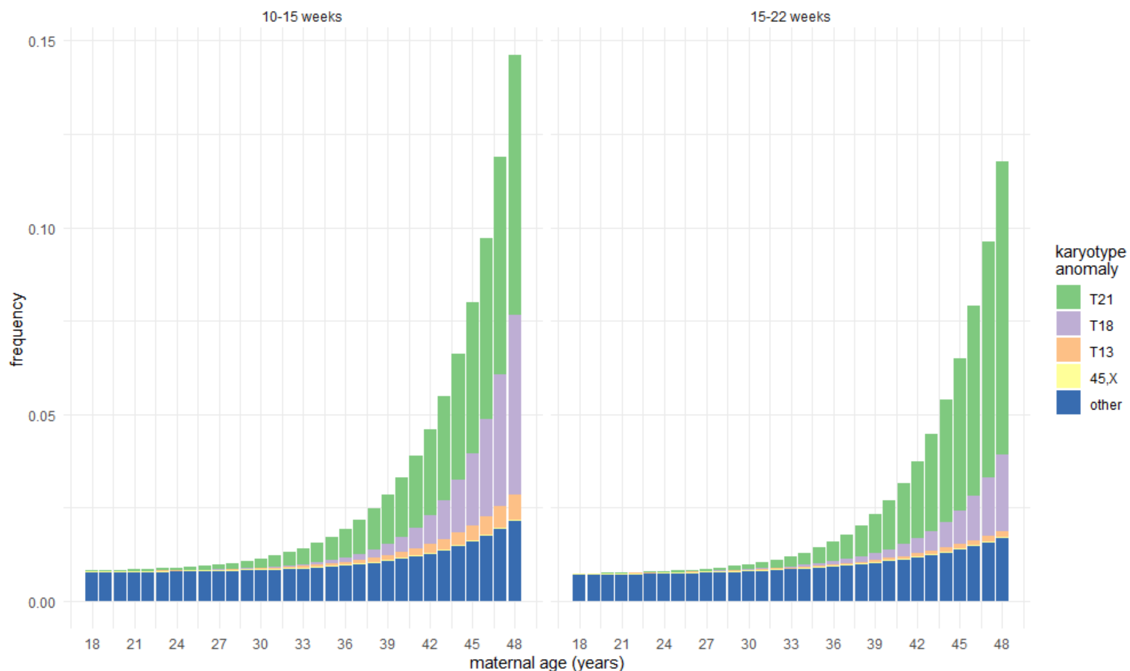

Figure 1. Relative frequency of karyotype anomalies per maternal age at 10 to 15 weeks and 15 to 20 weeks. Stacked bar plot of the frequency of each chromosomal defect for each maternal age and gestational age group. The x-axis corresponds to maternal age; the y-axis corresponds to the estimated frequency. T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy 13; 45,X – monosomy X; other includes the following chromosome anomalies: apparently balanced familial reciprocal translocation; inherited and satellite de novo small marker chromosome without euchromatic material; inherited inversion; de novo Robertsonian translocation; familial Robertsonian translocation; normal balanced de novo translocations and inversions; true fetal mosaicism 46/45,X (45,X<30%); true fetal mosaicism (autosomes and mos46/45,X >30%); dicentric chromosomes; abnormal balanced de novo translocations and inversions; unbalanced structural rearrangements; non-satellite and satellite de novo small marker chromosome containing euchromatic material; complex unbalanced karyotypes; autosomal aneuploidies other than T21, T18, T13; triploidy; sex chromosome anomalies other than 45,X (mosaic and homogeneous); 47, XXY. Adopted and modified from (Ferreira et al., 2016).

4.2. Common fetal aneuploidies

Trisomies are the most prevalent aneuploidies in humans, accounting for about 0.3% of all live births. Trisomies are defined by the presence of an extra chromosome and are induced mainly by segregation errors, which most commonly occur during meiotic divisions. Autosomal trisomies account for around 35% of spontaneous abortions and do not appear compatible with life, except for T21, and to some extent T18 and T13 (Hassold & Hunt, 2001). Analyses demonstrate that most trisomies, regardless of the chromosome, arise during oogenesis and are hence maternal origin. The most prevalent trisomy is chromosome 16 trisomy, affecting more than 1% of all pregnancies (Hassold et al., 1995). Only fetuses with the mosaic form may survive trisomy 16, and usually this condition causes spontaneous miscarriage in the first trimester (Langlois et al., 2006). T21, T18, and T13 are the most prevalent autosomal trisomies that survive to birth and, therefore, are collectively referred to as common fetal aneuploidies (Table 2). Furthermore, monosomy X (45,X) is also a relatively common aneuploidy, affecting about one in 2,000 liveborn females (Stochholm et al., 2006). Among common fetal autosomal aneuploidies only people with T21 typically live to adulthood and their lifespan is not significantly shortened (Weijerman & De Winter, 2010). On the other hand, T13 and T18 are the only additional autosomal trisomies found relatively often prenatally, yet newborns rarely survive past the first few months (Hassold & Hunt, 2001). Compared to the common autosomal trisomies, the disease phenotype of 45,X is relatively mild.

Table 2. Prevalence of common autosomal trisomies per 10,000 births. T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy; CI – confidence interval (Loane et al., 2013).

Trisomy	Total prevalence (95% CI)	Live birth prevalence (95% CI)
T21	22.0 (21.7-22.4)	11.2 (10.9-11.5)
T18	5.0 (4.8-5.1)	1.04 (0.96-1.12)
T13	2.0 (1.9-2.2)	0.48 (0.43-0.54)

4.2.1. Trisomy 21 – Down syndrome

T21, also known as Down syndrome, is a genetic disorder caused by an extra copy of chromosome 21. Down syndrome is the most prevalent chromosomal abnormality in humans (Holmes, 2014). In addition to the three copies of chromosome 21, around 3% of Down syndrome cases are caused by a Robertsonian translocation, the most prevalent of which is t(14;21); while other forms of translocations are a relatively rare cause of this condition (Kusre et al., 2015). In addition, 2% to 4% of people with Down syndrome have mosaicism (Papavassiliou et al., 2015). Approximately 5% of additional chromosomes in T21 are of paternal origin, whereas more than 90% are of maternal origin. Because human oocytes can be arrested in prophase I of the MI for several

decades, this phenomenon seems to be easy to understand. Thus, maternal MI errors are more prevalent than MII errors in the etiology of T21 (Hassold et al., 2007; Wartosch et al., 2021).

Down syndrome occurrence is dependent on maternal age (Asim et al., 2015). At the time of conception, the frequency of T21 is high, but roughly half to seventy percent of these pregnancies are lost (Roper & Reeves, 2006). Other autosomal trisomies may be even more frequent than T21, yet they have a much lower postnatal survival rate than Down syndrome. A small number of genes on chromosome 21, which is the smallest and least gene-dense of the autosomes, are assumed to be responsible for the high survival rate of patients with T21 (Roper & Reeves, 2006). There are several hypotheses on the disease origin of Down syndrome and the association between different genotypes and phenotypes. Gene dosage imbalance is one of them, in which the dosage (copy number) or quantity of genes of chromosome 21 is increased, resulting in enhanced gene expansion (Antonarakis et al., 2004).

People with Down syndrome have a higher risk of acquiring a variety of medical conditions and diseases. They usually have mild to severe intellectual impairment, growth restriction, and distinctive dysmorphic features (Chitty, 1998). Additional characteristics also linked to Down syndrome are congenital cardiac disease, skeletal dysplasia, gastrointestinal problems, immunosuppression, and visual and hearing impairment (Chitty, 1998). Congenital heart abnormalities affect up to half of all newborns with Down syndrome (Asim et al., 2015).

4.2.2. Trisomy 18 – Edwards syndrome

T18, commonly known as Edwards syndrome, is a chromosomal disorder caused by an extra copy of chromosome 18. Edwards syndrome is the second most frequent syndrome next to T21 (Bugge et al., 1998). Edwards syndrome is divided into three types: complete, partial, and mosaic T18. The most prevalent type of T18 is complete T18 making 94% of the cases. Non-disjunction, which usually occurs during maternal MII, is the cause of the extra chromosome 18. The additional chromosome is usually inherited from the mother (Bugge et al., 1998). The second most prevalent type of T18, following meiotic T18, is mosaic T18, with a frequency of less than 5%. Mosaic trisomy consists of a full T18 cell line along with the normal euploid cell line (Cereda & Carey, 2012). Edwards syndrome may also be caused by a partial (segmental) T18 that affects 2% of people with this particular syndrome, in which only a portion of chromosome 18q (long chromosome arm) is present in triplicate. A balanced translocation or inversion carried by one of the parents is frequently the cause of the partial triplicate of chromosome 18 (Cereda & Carey, 2012).

T18 pregnancies are known to have a high risk of fetal loss and stillbirth (Goel et al., 2019). In one study, the median survival time was 14.5 days for people with T18 and more than 90% died within the first year (Rasmussen et al., 2003). The frequency of Edwards syndrome rises in correlation with

maternal age. Due to a rise in the average maternal age, the prevalence of T18 has increased recently (Goel et al., 2019). Edwards syndrome incidence varies by country, mainly being related to the pregnancy termination policy. Interestingly, the sex of the affected fetus is biased, as more female than male fetuses have been found (ratio 3:2). This means that male fetuses have a greater risk for fetal loss than pregnancies with female fetuses, while the latter pregnancies have a better survival rate (Goel et al., 2019).

Edwards syndrome is defined by a wide range of clinical signs and symptoms, as well as the involvement of multi-organ dysfunction (Cereda & Carey, 2012). It is characterized by malformations of different organs that develop before birth, specific craniofacial features, prenatal growth deficiency, distinctive hand posture (overriding fingers), nail hypoplasia, short hallux, and short sternum, as the main clinical features that serve as diagnostic clues in the perinatal period. 90% of Edwards syndrome individuals have severe heart abnormalities, therefore only 5% to 10% of children with this condition live past their first year (Cereda & Carey, 2012).

4.2.3. Trisomy 13 – Patau syndrome

T13, commonly known as Patau syndrome, is an autosomal chromosomal condition caused by an extra copy of chromosome 13. The syndrome can occur as a full, partial (segmental), or mosaic form of T13 (Cammarata-Scalisi et al., 2019). The complete trisomy is the most frequent type of T13, accounting approximately for 80% of all cases (Cammarata-Scalisi et al., 2019). A Robertsonian translocation of t(13;14) in one of the parents leading to unbalanced three copies of chromosome 13 instead of the usual two in the fetus is significantly less prevalent, mosaicism for T13 is only found in 5% of all cases (Miryounesi et al., 2016). T13 is the third most prevalent trisomy. Patients carrying the genomic mutation have a postnatal survival rate of 6% to 12% beyond the first year of life (Satgé et al., 2017).

T13 has become more common in recent decades as the number of pregnancies among women of advanced maternal age has increased (Loane et al., 2013). According to the study, the median survival time of T13 was 7 days and approximately 90% of infants died within the first year (Rasmussen et al., 2003). Recent advancements in clinical management, such as surgical and intensive medicinal treatment, have shown progress in survival rate. As a result, more T13 patients are now living past their first year of life (Goel et al., 2019).

The most prevalent defects in T13 are cardiac and nervous system abnormalities (Springett et al., 2015). Fetal growth restriction and holoprosencephaly, in which the embryonic forebrain fails to develop into two cerebral hemispheres, are common symptoms of T13. Cyclopia, cleft lip, and cleft palate are examples of midline facial abnormalities. Patients with T13 can develop congenital heart disease, polycystic kidney disease, urogenital malformations, and gynecological dysgenesis (Watson et al., 2007).

4.2.4. Monosomy X – Turner syndrome

45,X, also known as Turner syndrome, is caused by a partial or total loss of one of the X chromosomes (Saenger, 2001). Congenital ovarian hypoplasia syndrome is another name for 45,X. In females, it is the most frequent chromosomal defect (Saenger, 2001). Non-mosaic 45,X affects around half of the people with Turner syndrome, while in the other half, a mosaic 45,X form is found (Wolff et al., 2010). Turner syndrome usually is *de novo* in origin and occurs due to a random event during meiosis. Non-disjunction is a type of cell division defect that can result in gametes with an incorrect number of chromosomes (Wolff et al., 2010). Turner syndrome, unlike trisomies, is unrelated to maternal age and is more likely due to instability of chromosome Y leading to its loss during male meiosis. Therefore, in 75 to 80% of the Turner syndrome cases, chromosome X of maternal origin is still present (Renna et al., 2013).

Turner syndrome may remain undiagnosed due to the absence or mild phenotypic effects at birth. As many girls or women with a modest phenotypic outcome remain undiagnosed or are discovered later in life, its real frequency is somewhat unclear (Gunther et al., 2004). Turner syndrome affects people of all ethnicities and nations to a similar extent. Because some pregnant women choose to terminate their pregnancies when their fetuses have 45,X, the prevalence of Turner syndrome at birth is decreasing as prenatal screening becomes more widely used (Kikkeri & Nagalli, 2021).

The prominent feature of Turner syndrome patients is their small height, which is present in all of them. A wide chest, short neck, *genu valgum*, and nail dysplasia are some of the other traits (Sävendahl & Davenport, 2000). Turner syndrome patients have normal intellect but may develop neurocognitive abnormalities, such as memory, attention, and learning issues. Turner syndrome patients have a higher incidence of cardiovascular abnormalities, which increases their mortality risk (Kikkeri & Nagalli, 2021). Women with Turner syndrome are at extremely high risk for primary ovarian insufficiency and infertility due to the atresia of follicles, which may already start before puberty. Therefore, the patients might appear with either a delayed onset of puberty or amenorrhea in adolescent age. As Turner syndrome is associated with follicular death, the elevated level of follicle-stimulating hormone and decreased level of the anti-Mullerian hormone are considered sensitive markers for predicting ovarian failure (Lunding et al., 2015).

4.3. Prenatal screening and diagnosis

For more than 40 years, prenatal genetic diagnosis for fetal aneuploidies, such as T21, T18, and T13, has been an important element of prenatal care. Only fetal genetic material can give a definitive prenatal diagnosis, which requires invasive testing and is accompanied by a risk of miscarriage (Chitayat et al., 2011). Due to that, more convenient and safe screening approaches are used to

limit the practice of invasive procedures and thus the risk for pregnancy loss. To assess the risk of chromosomal abnormalities, a variety of screening technologies are used, each with its own set of advantages and disadvantages in terms of gestational time, cost, and accuracy. Table 3 shows a list of different prenatal screening and diagnostics approaches. The majority of prenatal follow-up is done for the purpose of screening in order to limit the population of pregnant women truly requiring invasive prenatal diagnostics (Kotsopoulou et al., 2015).

Starting around the 1980s the second trimester screening was used for aneuploidy detection (Russo & Blakemore, 2014). Nevertheless, in the late 1990s the introduction of the FCT (Spencer et al., 1999), which uses ultrasound measurement of NT thickness along with maternal serum risk markers shifted fetal aneuploidy screening from the second trimester to the first trimester (Carlson & Vora, 2017). The number of fetuses in the current pregnancy (singletons vs multiple pregnancies), maternal age, previous reproductive history, weight/body mass index of the mother, race, serum trisomy biomarkers, and NT measurement are all used to establish an individual risk assessment. The presence or lack of visualized nasal bone is also incorporated into some risk calculations. Thereafter, the risk is expressed as a ratio, for example, 1 in 10, indicating the risk for fetal trisomies for individual chromosomes (Carlson & Vora, 2017).

The karyotype analysis is the golden standard for the genetic diagnosis of common fetal aneuploidies by AC or CVS (Renna et al., 2013). An invasive test, such as AC, CVS, or cordocentesis, is used to obtain fetal genetic material for a conclusive diagnosis once the high risk has been identified by screening (Alldred et al., 2017). AC is an invasive procedure that includes obtaining a small sample of the amniotic fluid around the fetus with a needle that enters through the abdominal wall into the uterine cavity, under ultrasound guidance. CVS provides a sample of placental or trophoblastic tissue using a similar method as AC. CVS is normally done between 10 and 13 weeks of pregnancy (Alldred et al., 2017). Even when performed by a skilled specialist, AC and CVS may have a small risk of miscarriage. According to a systematic analysis, the procedure-related risk of pregnancy loss was 0.11% for AC and 0.22% for CVS (Akolekar et al., 2015).

Nevertheless, if more precise primary screening approaches are correctly utilized, the number of unnecessary invasive tests can be significantly decreased. As a result, a more reliable, convenient, and safe prenatal diagnosis program with a lower risk of fetal loss and a lower rate of false positive results is possible (Wright & Chitty, 2009).

Table 3. Methods of prenatal screening and diagnosis based on their invasiveness and timing. Adopted and modified from (Kotsopoulou et al., 2015)

Invasiveness	Method	Timing
Non-invasive	Ultrasonographic examination of nuchal translucency	First or second trimester
	Maternal serum screening	
	Fetal heartbeat	
Invasive	Chorionic villus sampling	After 10 weeks
	Amniocentesis	After 15 weeks
	Cordocentesis	After 20 weeks

4.3.1. First trimester combined test

Bogart and colleagues discovered that high hCG levels in maternal serum were linked to Down syndrome and could be used as a possible screening marker for fetal chromosomal abnormalities in the second trimester (Bogart et al., 1987). Later research revealed that the free β subunit of hCG was an even more useful marker than total hCG (Macri et al., 1993). Another blood marker, PAPP-A, which is released by the placenta, has also been linked to Down syndrome. PAPP-A levels have been found to be lower in the first trimester of Down syndrome pregnancies (Sersinger et al., 1995). With T13, T18, and X,45 maternal serum free β-hCG and PAPP-A levels are lower (Renna et al., 2013; Shiefa et al., 2013).

NT is the anatomical feature of the fetus as measured in fetal ultrasound scanning and indicating the accumulation of the fluid behind the neck of the fetus. NT testing is reliable by ultrasound up to 15 weeks of gestation. Gestational age influences the size of the NT and is a normal part of the development (Pajkrt et al., 1995). Increased NT has been linked to a wide range of genetic disorders, morphological abnormalities, a higher risk of miscarriage, and intrauterine fetal death. A chromosomal defect affects about 20% of fetuses with increased NT (Bakker et al., 2014). The thickening of the neck region was the initial sign linked with Down syndrome. 40% to 50% of affected fetuses had a thickened nuchal fold measuring at least 6 mm in the second trimester, with a FPR of 0.1% (Renna et al., 2013). T18 and T13 are detected in 66.7% of cases when the nuchal fold and nasal bone are assessed (Geipel et al., 2010). Around 90% of 45,X cases have increased nuchal fold (Surerus et al., 2003).

The FCT model incorporating the analysis of maternal serum biomarkers, PAPP-A, and free β-hCG, as well as fetal NT by ultrasound, was introduced in 1999 and is widely used as the primary prenatal screening test in many countries (Spencer et al., 1999). FCT has considerably better accuracy than the NT and maternal age-based tests with or without either PAPP-A or free β-hCG (Alldred et al., 2017). The development of FCT showed that a combined test detects about 90% of Down syndrome pregnancies, with a FPR of 5% (Spencer et al., 1999). However, over 90% of pregnant women who test positive by FCT and have an invasive follow-up diagnostic procedure are found to carry a fetus

with no chromosomal abnormalities, causing unneeded stress and risk of miscarriage (Table 4).

Table 4. Performance characteristics of first trimester combined screening. PPV – positive predictive value; CI – confidence interval; T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy. Adapted and modified from (van Elslande et al., 2019).

Statistics	Combined (95% CI)	T21 (95% CI)	T18 (95% CI)	T13 (95% CI)
Sensitivity	86.1% (70.5–95.3)	90.0% (68.3–98.8)	90.0% (55.5–98.8)	33.3% (0.8–90.6)
Specificity	97.3% (97.2–97.7)	97.7% (97.5–97.9)	99.0% (98.9–99.2)	99.2% (99.0–99.3)
PPV	6.3% (5.4–7.3)	4.3% (3.6–5.0)	5.2% (4.1–6.7)	0.7% (0.1–3.4)

4.3.2. Prenatal screening and diagnosis in Estonia

The national prenatal screening program in Estonia was initiated in 1995 with maternal age being the primary risk factor. In 1999, second trimester serum screening was introduced with the support of the Estonian Health Insurance Fund. The first attempts to perform FCT in the Estonian public healthcare system was in 2005 (Ustav et al., 2016). Since 2016, FCT has been declared as the primary prenatal screening test for fetal chromosomal disorders by the national prenatal diagnostics guideline (Ustav et al., 2016). As a result, the invasive testing rate decreased from 12% in 2007 to about 5% in 2016 (Sitska, 2018). Further on, the invasive testing rate has declined even more to only 2.7%, which is attributed to universal FCT (Ridnõi et al., 2021). The study also shows that FCT is an effective method for prenatal screening of fetal aneuploidies with a false positive rate of 3.12% (Ridnõi et al., 2021). Currently, the overall coverage of prenatal screening in Estonia is more than 90% (Sitska, 2018).

Table 5. Screening performance of first trimester combined test for common autosomal trisomies in Estonia. T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy 13; DR – detection rate; CI – confidence interval; PPV – positive predictive value; NPV – negative predictive value. Adopted and modified from (Ridnõi et al., 2021).

Trisomy	Cases	Total	DR (95% CI)	Sensitivity (95% CI)	95.38% (86.24–98.80)
T21	51	14,566	94% (84.09–97.98)	Specificity (95% CI)	96.86% (86.24–98.80)
T18	11		100% (74.21–100)	PPV (95% CI)	11.99% (9.38–15.18)
T13	3		100% (43.85–100)	NPV (95% CI)	99.97% (99.93–99.99)

In 1990, the first amniocentesis was performed in Tartu University Hospital which marks the beginning of prenatal diagnosis in Estonia. Initially, chromosomal disorders were diagnosed by karyotyping. As the development of prenatal diagnosis advanced new approaches became available and were integrated into clinical practice in the Estonian public healthcare system – interphase fluorescence *in situ* hybridization in 2000, chromosomal microarray analysis in 2011, and next generation sequencing in 2015 (Sitska, 2018).

Every pregnant woman with an Estonian residency visa, regardless of citizenship, is covered by Estonian Health Insurance Fund for prenatal care. During pregnancy, all routine screening and diagnostic procedures are free of charge (Ustav et al., 2016). This helps to provide high-quality prenatal screening regardless of socioeconomic background.

4.4. Cell-free DNA

cfDNA plays a central role in non-invasive screening for many diseases, especially in fetal genetic abnormality and solid tumor oncology analysis. The presence of cfDNA in human blood was first reported in 1948 by Mandel & Métais (Mandel & Métais, 1948). cfDNA is mostly a short double-stranded DNA fragment freely circulating in the bloodstream. Although the origin of the cfDNA is still in dispute, it is commonly believed that the main sources are cellular breakdown mechanisms and active DNA release mechanisms, including the extracellular vesicle-based transport of nucleic acids. The debate regarding the origins of cfDNA emerged as a result of the widespread perception that most, if not all, biological parameters and processes (e.g., aging, exercising, diseases like cancer, inflammatory and immune reactions, etc.) are influencing the release of DNA into the circulation (Aucamp et al., 2018).

cfDNA in a pregnant woman's circulation also includes a DNA of fetal origin, also known as cffDNA. In 1997, Lo et al. were the first to demonstrate the presence of male fetal DNA in pregnant woman's plasma and serum samples (Dennis Lo et al., 1997). This shows that the placenta forms a bidirectional permeable barrier between the pregnant woman and her fetus (Figure 2). It is also supported by the presence of intact fetal cells in a pregnant woman's circulation (Bianchi, 1999). The discovery of fetal-originated DNA from maternal blood has led to a new generation of prenatal screening approaches which target cffDNA as the source of fetal genetic information.

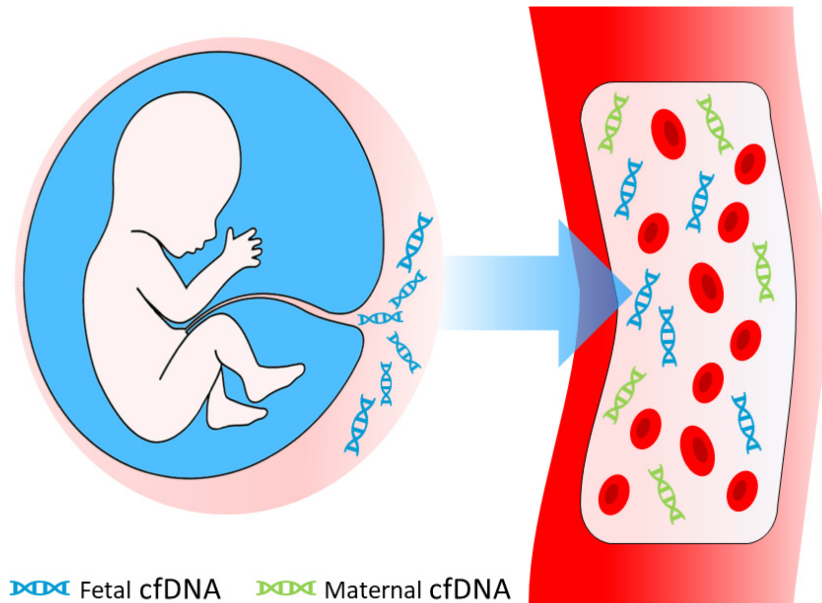

Figure 2. Origin of cell-free fetal DNA (cffDNA). During pregnancy, some cells of the placenta, mainly trophoblasts, undergo apoptosis – a programmed cell death. As a result, small DNA fragments, known as cffDNA, are released into maternal circulation for utilization. cffDNA disappears within a few hours after delivery.

Trophoblasts, which form the outer layer of a blastocyst and develop into a large part of the placenta, and fetal hematopoietic cells are considered to be the main source of the cffDNA (Alberry et al., 2007; Aucamp et al., 2018). In addition to maternal blood, cffDNA has also been detected in amniotic, cerebrospinal and peritoneal fluid, and urine of a pregnant woman (Bianchi, 2004). cffDNA accounts for approximately 10% of the total cfDNA in pregnant woman's circulation at 10 to 13 weeks of gestation (Ashoor, Syngelaki, Poon, et al., 2013; Wang et al., 2013). It is detectable from as early as the 7th week of gestation, and the concentration gradually increases as the pregnancy progresses, with a noticeable increase during the last eight weeks of pregnancy (Lo et al., 1998). The estimated half-life of the cffDNA fragment is up to 30 min, and within a few hours after delivery, the cffDNA has disappeared from maternal plasma (Dennis Lo et al., 1999).

cffDNA fragments are approximately 150 bp in length. It is shown that cfDNA fragments of fetal origin are shorter than cfDNA fragments of maternal origin (Fan et al., 2010). The short length of cffDNA fragments implies apoptosis – a programmed cell death – as the releasing mechanism rather than necrosis – accidental cell death (Gahan, 2013). The difference in size allows cffDNA to be distinguished from maternal DNA fragments. This can be utilized in prenatal screening as cffDNA enrichment by size-selection reduces back-

ground noise originating from maternal DNA fragments leading to more precise determination of more subtle fetal genetic traits (Li et al., 2004).

4.5. Non-invasive prenatal testing

NIPT is a common term for cfDNA-based prenatal screening approaches. As the cfDNA of a pregnant woman also includes DNA fragments of fetal origin, it can be used to assess the fetal genetic profile (Dennis Lo et al., 1999). For prenatal screening, cfDNA is extracted from a blood sample of a pregnant woman. The extracted cfDNA is sequenced and analyzed to detect fetal chromosomal aberrations. In addition to chromosomal aneuploidies, NIPT can also be used to detect fetal sex, sub-chromosomal and single-gene disorders (Ayse Kirbas, Korkut Daglar, 2016). In 2011, NIPT became clinically available for the detection of common fetal trisomies, such as T21, T18, and T13 (Palomaki et al., 2011). In recent years, NIPT has been widely integrated into public healthcare systems as either a first-line test or an addition to the existing prenatal screening programs. For example, most European countries, including Estonia, have implemented NIPT as an offer for higher risk women after FCT, while in Belgium and the Netherlands NIPT is provided for all pregnant women (Gadsbøll et al., 2020).

Compared to the conventional prenatal screening methods like ultrasound and maternal serum screening, NIPT offers higher sensitivity and specificity (Iwarsson et al., 2017; Taylor-Phillips et al., 2016). A meta-analysis reviewing the screening performance of NIPT for common fetal aneuploidies, concluded in Table 6, supports that cfDNA analysis of maternal blood can be used safely for prenatal screening in singleton pregnancies (Gil et al., 2017). In addition, aneuploidy screening with NIPT has been shown to be superior compared to the combination of conventional methods – mother's age, ultrasound results, and biochemical analysis of the maternal serum (Norton et al., 2015).

Table 6. Screening performance of non-invasive prenatal testing studies for common aneuploidies. T21 – trisomy 21; T18 – trisomy 18; T13 – trisomy; 45,X – monosomy X; DR – weighted pooled detection rate; CI – confidence interval; FPR – weighted pooled false positive rate. Adapted and modified from (Gil et al., 2017).

Condition	Cases	Non-cases	DR (95% CI)	FPR (95% CI)
T21	1,963	223,932	99.7% (99.1–99.9)	0.04% (0.02–0.07)
T18	563	222,013	97.9% (94.9–99.1)	0.04% (0.03–0.07)
T13	119	212,883	99.0% (65.8–100.0)	0.04% (0.02–0.07)
45,X	36	7,676	95.8% (70.3–99.5)	0.14% (0.05–0.38)

Fetal fraction concentration in a cfDNA sample is an important quality metric for cfDNA-based prenatal screening. The overall amount of cffDNA in total cfDNA extracted from maternal plasma is low, and there are no reliable laboratory methods to efficiently separate cffDNA molecules from maternal cfDNA molecules. Distinguishing the fetal genomic signal from maternal background noise poses a significant technical obstacle in the development of NIPT. Currently, the threshold of 4% for cffDNA fraction has been agreed, below which the determination of fetal trisomies becomes unreliable (Canick et al., 2013; Ehrich et al., 2011; Gregg et al., 2016; Norton et al., 2012).

Depending on the assay, 0.1% to 6.1% of tests fail due to the low fetal fraction (Gil et al., 2017). The most well-known biological factors affecting fetal fraction are maternal body mass index (BMI) and gestational age. It has been shown that fetal fraction decreases as maternal BMI increases (Kinnings et al., 2015). Therefore, cfDNA-based prenatal screening may not be the best method for obese women. As fetal fraction concentration gradually increases with gestational age, the optimal time for NIPT is usually from 10 weeks of gestation onwards to reduce failures due to the low fetal fraction (Wang et al., 2013).

A number of approaches have been suggested for fetal fraction enrichment by taking advantage of cfDNA fragment size difference and immunoprecipitation of hypo- or hypermethylated DNA (Li et al., 2004; Papageorgiou et al., 2009; Welker et al., 2020). Even without cffDNA enrichment, several approaches have proven to be effective and are used in clinical practice.

4.5.1. Whole-genome sequencing approach

The whole-genome sequencing (WGS) based approach allows to sequence millions of both fetal and maternal cfDNA fragments simultaneously. Each sequencing read is aligned to the human reference genome and linked to the chromosome from which it originates. In the case of fetal aneuploidy, there should be a relative excess or deficit of read counts for the chromosome of interest compared to the reference set of known euploid cases (Chiu et al., 2008; Fan et al., 2008). The difference in counts will be small as cffDNA constitutes only a fraction of the total cfDNA. To be able to detect minor differences in counts reliably, the fetal fraction and/or the sequencing coverage needs to be sufficient. As fetal fraction is a biological variable with limited control, WGS requires millions of reads per sample to reliably detect fetal aneuploidy when the fetal fraction is suboptimal (Paluoja et al., 2021).

The comparison between the reference set and the tested sample is often expressed as a z-score, which measures the deviation from the diploid situation of the reference set in standard deviations. An aneuploidy is called if a z-score of the chromosome of interest is over or under the predefined upper or lower z-score thresholds, respectively. Well-known commercial tests using the WGS-based approach are MaterniT21 PLUS™ by Sequenom (San Diego, USA) and Verifi™ by Illumina (San Diego, USA).

4.5.2. Targeted sequencing approach

The targeted sequencing based approach differs from the WGS by selectively amplifying and sequencing only the specific genomic regions of interest. Targeting involves the use of hybridization-based capture and allows to focus only on clinically important chromosomes (21, 18, 13, and X). The reads are counted by chromosome and compared against the expected proportion calculated over all observed proportions within the sequencing lane (Sparks, Struble, et al., 2012). As nearly all targeted sequences are useful in fetal aneuploidy detection, the total number of reads is significantly reduced. This enables to sequence more samples in a single run and thereby reduces the sequencing costs per sample. Targeted sequencing followed by read counting with respective statistical analyses has shown similar sensitivities and specificities for common fetal aneuploidies as the WGS-based approach (Ashoor, Syngelaki, Wang, et al., 2013; Liao et al., 2012; Norton et al., 2012).

In 2012, Ariosa Diagnostics (San Jose, USA) developed a targeted assay known as Digital Analysis of Selected Regions (DANSR™). The assay is based on the amplification of specific regions under investigation (Sparks, Wang, et al., 2012). DANSR™, in combination with the Fetal Fraction Optimized Risk of Trisomy Evaluation (FORTE™) algorithm, which uses non-polymorphic loci for the analysis of chromosome proportions, are the technologies behind the Harmony™ prenatal test by Roche (Basel, Switzerland). The method includes an initial step of targeted amplification of loci on the chromosome of interest prior to analysis, which results in an improvement in sequencing efficiency per target chromosome. Compared to the WGS approach, DANSR™ enables aneuploidy detection using approximately 1 million raw reads per sample. As with all quantitative methods, the approach is dependent on the amplification variability of the target region, which may limit its diagnostic accuracy for some aneuploidy regions (Norwitz & Levy, 2013). Although quantification of cfDNA for fetal aneuploidy screening has mostly relied upon sequencing, DANSR™ has been shown to accurately quantify the targeted loci for NIPT using DNA microarrays, thus, having an advantage in terms of sequencing cost and throughput (Juneau et al., 2014; Stokowski et al., 2015).

4.5.3. SNP-based targeted sequencing approach

The targeted sequencing based approach with the inclusion of single-nucleotide polymorphisms (SNPs) based on allelic genotype data utilizes significantly more information than sequencing-based methods, which consider only the number of reads to identify fetal aneuploidies. The advantage of using SNP allelic information allows distinguishing which sequencing reads originate from the fetus and which from the mother. In addition to detecting fetal chromosomal aneuploidies, the approach also enables haplotype reconstruction and identification of genetic aberrations that cannot be detected using counting methods, such as triploidy and uniparental disomy.

A leading SNP-based commercial test, Panorama™ by Natera (San Carlos, USA), targets 19,488 SNPs for aneuploidy screening. Their Next-generation Aneuploidy Test Using SNPs (NATUS™) algorithm considers parental SNP genotypes or, in the absence of a paternal sample, population allele frequencies and crossover frequency data to calculate the expected allele distribution and possible fetal genotypes based on the estimated recombination sites in the parental chromosomes. The model compares predicted allelic distributions to observed allelic distributions by employing a Bayesian-based Maximum Likelihood approach to determine the relative likelihood of chromosomal copy number hypothesis family, such as monosomy, disomy, or trisomy, given the data. The likelihoods of each copy number hypothesis family are summed, and the hypothesis with the maximum likelihood is selected as a result (Samango-Sprouse et al., 2013; Zimmermann et al., 2012).

4.6. Hidden Markov model

Hidden Markov model (HMM), named after the Russian mathematician Andrey Andreyevich Markov, who developed much of the relevant statistical theory, is a statistical model that was first proposed by Baum (Baum & Petrie, 1966). In a HMM, the system being modeled is assumed to be a Markov process where the observed parameters are used to identify the unknown parameters. These parameters are then used for further analysis.

The basic principle of HMM is that the observed events have no direct linkage with hidden states. Observed events are connected to states through the probability distribution. The model includes a Markov chain as the basic stochastic process and describes state transitions and the statistical correspondence between the hidden states and observed values.

HMMs have been successfully applied in pattern recognition such as speech, handwriting, and gesture recognition. Since the second half of the 1980s, HMMs have been adopted in computational biology to resolve various problems of biological sequence analysis including gene prediction, pairwise and multiple sequence alignment, base-calling, modeling DNA sequencing errors, protein secondary structure prediction, non-coding RNA identification, RNA structural alignment, acceleration of RNA folding and alignment, fast non-coding RNA annotation, etc (Yoon, 2009). HMMs have become a fundamental tool in bioinformatics for their robust statistical foundation, conceptual simplicity, and fit for diverse classification problems.

In this thesis, our idea was to model a studied chromosome as a HMM of sequential loci with fetal chromosomal copy number as the hidden state. Classifying sequential loci distributed on clinically important chromosomes or sub-chromosomal regions is a novel approach by the author of the thesis to detect fetal chromosomal trisomies in the context of NIPT.

4.7. Summary of the literature

Chromosomal aneuploidies are the leading cause of spontaneous miscarriages and congenital disorders in humans. Even though at least 10% of all clinically confirmed pregnancies are aneuploid, most of them are assumed to abort spontaneously during the earliest stages of pregnancy. The majority of aneuploidies are caused by the error occurred in meiosis.

One of the most well-known and frequent fetal aneuploidy is Down syndrome – a genetic disorder caused by the presence of a third copy of chromosome 21. Down syndrome occurrence increases with maternal age, while the probability of having a child with T21 is less than 0.1% for a 20-year-old mother compared to 3% for a 45-year-old mother.

Conventional prenatal screening for fetal aneuploidies relies mostly on FCT, which includes the analysis of maternal serum biomarkers and fetal ultrasound. Pregnant women with high-risk pregnancies are offered an invasive procedure such as AC or CVS followed by karyotyping for confirmation of the fetal genomic disease. As a result, most of the high-risk pregnancies are detected successfully. Unfortunately, the PPV of the FCT is low, and, thereby, most of the patients with high-risk results have actually healthy fetuses. This causes extra stress and procedure-related risk of miscarriage for a lot of patients who are unnecessarily undergoing an invasive confirmatory procedure.

A novel cfDNA-based approach, known as NIPT, has demonstrated superior results compared to conventional prenatal screening methods. NIPT has great potential as the number of FPR is significantly lower. On the other side, due to the sequencing cost, NIPT is still a more expensive alternative for routine use in clinical care for all pregnant women.

Therefore, novel laboratory techniques and technological advancements are essential to make NIPT more affordable and widely adopted for prenatal screening. In addition, new approaches also require appropriate computational solutions and analytical software to be developed, tested, and validated. Furthermore, after the corresponding scientific research and development phase, these software tools have to be optimized for the task and implemented into automated and scalable computational workflows, which can be utilized for routine clinical practice.

5. AIMS OF THE STUDY

1. To develop the appropriate sequencing data processing and analysis workflow for TAC-seq assay. Apply the developed tools on TAC-seq assay to analyze synthetic trisomic samples, and evaluate the accuracy of the corresponding results.
2. To develop a computational analysis framework and related statistical methods for targeted high-coverage NIPT sequencing assay and study the biological and technical limitations and accuracy of these methods using a variety of *in silico* simulated datasets. Apply these computational tools for NIPT analysis on *in vitro* generated synthetic trisomic samples.
3. To set up, optimize and implement a low-coverage WGS-based NIPT workflow into the Estonian healthcare system. Conduct a validation study using clinical samples of pregnant patients.

6. MATERIAL AND METHODS

6.1. Study subjects

6.1.1. *In vitro* (Papers I–II)

The genomic DNA of T21 cell line (GM04616) and non-T21 cell line (GM01359) were mixed to create experimentally controlled samples corresponding to fetal T21 pregnancy. Artificially constructed samples followed specific library preparation. Extracted genomic DNA was sheared by sonication to 150–200 bp fragments to imitate cfDNA and mixed in different ratios to represent a wide range of fetal fraction. Specifically designed TAC-seq detection probes for NIPT were used to target the reference chromosomes (2 and 3) and the chromosome of interest (21). Experiment 1 included 114 targeted loci for chromosomes 2 and 21. 16 samples split into six equal-sized groups where T21 DNA fraction was 0%, 5%, 10%, 15%, 30%, and 100%, respectively. Experiment 2 included the extended set of 224 targeted loci for chromosomes 2, 3, and 21. 12 samples were split into groups of three where T21 DNA fractions were 0%, 10%, 20%, and 100%, respectively. Both libraries were sequenced using Illumina NextSeq 500 platform. UMI threshold equal to one was applied to remove PCR duplicates from the dataset using TAC-seq data analysis tool (Teder & Paluoja, TAC-seq data analysis, 2018). Obtained molecule counts were further processed to eliminate outliers using the 1.5 × interquartile range rule. Filtered counts were used to evaluate TAC-seq trisomy detection accuracy (the read count based model) and trisomic DNA fraction estimation.

6.1.2. *In silico* (Paper II)

A computer simulation was used to model allele counts corresponding to the targeted sequencing dataset of pregnant women's cfDNA samples. A total of 1,800 different datasets were generated, which varied in the context of fetal chromosomal condition (euploidy, maternally or paternally originated trisomy), read depth (500 to 15,000 reads per locus), and fetal fraction (1 to 20% of total cfDNA). Each dataset included 100 training and 10,000 test samples. Each sample was a mixture of maternal and fetal components, which incorporated 1,000 targeted loci per chromosome of interest (used for estimating fetal chromosomal condition) and 1,000 targeted loci per reference chromosome (used for estimating fetal fraction and HMM emission parameters). Each locus was either homozygous or heterozygous, as both alleles had an equal likelihood of occurrence in parents (minor allele frequency = 0.5). Fetal alleles, which originated from parental alleles, underwent randomly simulated chromosomal crossover by exchanging homologous alleles. Generated allele counts followed negative binomial distribution (as real-life 2nd generation sequencing data) with a mean proportional to the chromosomal copy number and with a variance-to-mean ratio of 3 (Miller et al., 2011). The simulated datasets were used to train

the developed models (to estimate the corresponding parameters) and to evaluate and compare fetal trisomy detection accuracy and fetal fraction estimation. The computer simulation code is deposited online (Teder, 2018).

6.1.3. Clinical patients (Paper III)

A total of 424 patients with a singleton pregnancy were included in the validation study from Estonian and Belgium cohorts. The samples were collected and processed according to previously published guidelines from KU Leuven (Leuven, Belgium) with minor modifications (Bayindir et al., 2015). Briefly, peripheral blood samples were collected in cfDNA BCT tubes (Streck, USA) and plasma was separated within 72 h by standard dual centrifugation. cfDNA was extracted from 4 to 5 ml plasma using QIAamp Circulating Nucleic Acid Kit (Qiagen, Germany). Libraries were prepared using the TruSeq ChIP Library Preparation kit (Illumina Inc., USA) with 12 cycles for the final PCR enrichment step. Equal amounts of 24 libraries were pooled and the quality of the pool was assessed on Agilent 2200 TapeStation (Agilent Technologies, USA). Sequencing was performed on the Illumina NextSeq 500 (Illumina Inc., USA) with an average coverage of 0.32 and producing 85 bp single-end reads.

The Estonian cohort included 259 and 149 blood samples from high-risk and low-risk pregnancies (also referred to as general population), respectively. The blood samples were collected at Tartu University Hospital (Tartu, Estonia) and East-Tallinn Central Hospital (Tallinn, Estonia) from 2015 to 2018. The mean gestational age at the time of the blood sampling was 15 and 25 weeks for high-risk and low-risk pregnancies, respectively. The high-risk factors included increased risk (>1/300) of any common fetal trisomy (T21, T18, and T13) based on the FCT (analysis of maternal serum biomarkers and fetal NT ultrasound), family history of genetic diseases, ultrasound abnormalities, and advanced maternal age (≥36 years). Pregnant patients from the high-risk group underwent the invasive prenatal diagnostic procedure (CVS or AC followed by karyotyping, fluorescence in situ hybridization, or chromosomal microarray analysis). Patients with low risk from the FCT and no detectable fetal abnormalities on ultrasound examination were defined as low-risk pregnancies. The general health status of a baby was confirmed at birth by a neonatologist. In total, the Estonian cohort included 28 aneuploidy cases (13 T21, 7 T18, 4 45,X, 1 mosaic T18, 1 T13, 1 mosaic T8, and 1 47,XXX).

The Belgium cohort included 16 cfDNA samples from the Center for Human Genetics, UZ Leuven (Leuven, Belgium) enriched for common fetal trisomy cases. The cohort included 8 common trisomy cases (2 T21, 3 T18, and 3 T13) previously confirmed by the KU Leuven laboratory (Leuven, Belgium).

6.2. Methods

6.2.1. Allelic ratio (Paper II)

Allelic ratio is defined as the ratio of major allele specific sequencing read count to minor allele count of a locus. Allelic ratio depends on the maternal and fetal genotypes and proportions. This means that allelic ratio can be used to detect fetal trisomy if the fetal fraction is known (Table 7).

Table 7. Allelic patterns for biallelic loci. The allelic pattern of a biallelic locus depends on maternal and fetal genotypes (homozygous or heterozygous), and fetal condition (euploidy, maternally or paternally inherited trisomy). More frequently occurring allele represents the major allele, and less frequently occurring allele represents the minor allele. The ratio of major allele count over minor allele count of a locus is called allelic ratio. MA – maternal allele count; FA – fetal allele count. Adopted from (Teder et al., 2019).

Fetal condition	Maternal genotype	Fetal genotype	Major allele	Minor allele	Allelic ratio
Euploidy	Homo-zygous	Homozygous	$2MA + 2FA$	-	-
		Heterozygous	$2MA + FA$	FA	$\frac{2MA + FA}{FA}$
Maternally inherited trisomy		Homozygous	$2MA + 3FA$	-	-
		Heterozygous	$2MA + 2FA$	FA	$\frac{2MA + 2FA}{FA}$
Paternally inherited trisomy		Homozygous	$2MA + 3FA$	-	-
		Heterozygous	$2MA + 2FA$	FA	$\frac{2MA + 2FA}{FA}$
			$2MA + FA$	$2FA$	$\frac{2MA + FA}{2FA}$
Euploidy	Hetero-zygous	Homozygous	$MA + 2FA$	MA	$\frac{MA + 2FA}{MA}$
		Heterozygous	$MA + FA$	$MA + FA$	$\frac{MA + FA}{MA + FA}$
Maternally inherited trisomy		Heterozygous	$MA + 2FA$	$MA + FA$	$\frac{MA + 2FA}{MA + FA}$
Paternally inherited trisomy		Homozygous	$MA + 3FA$	MA	$\frac{MA + 3FA}{MA}$
		Heterozygous	$MA + 2FA$	$MA + FA$	$\frac{MA + 2FA}{MA + FA}$

Allelic ratio can be calculated for informative loci where either mother or fetus is heterozygous. On average, 75% of biallelic loci are informative, assuming that both alleles have an equal likelihood of occurrence (minor allele frequency = 0.5). The proportion of informative loci is even higher in the case of paternally originated trisomy as both paternally inherited alleles contribute to heterozygosity.

6.2.2. Hidden Markov model (Paper II)

A developed hidden Markov model (HMM) was used to classify sequential targeted loci into predefined states representing different fetal conditions like euploidy and trisomy. The HMM models were implemented in Python version 3.6.2 (https://www.python.org/) using the *hmmlearn* package version 0.2.0 (https://hmmlearn.readthedocs.io/). First, three distinct models were created based on the observed measurements of targeted loci. Second, the model parameters were estimated using the training samples. Finally, the Viterbi algorithm was used to find the most likely underlying fetal condition behind sequential targeted loci.

The read count (RC) model is a 2-state HMM which utilizes read counts of sequential targeted loci to estimate the underlying fetal condition (euploidy or trisomy) of each target (Figure 3A). The RC model is based on the assumption that the mean sequencing coverage of a given region is proportional to the DNA copy number of the region. In the case of fetal trisomy, the sequencing read counts would increase by half of the fetal fraction due to the extra fetal chromosome compared to fetal euploidy.

The allelic ratio (AR) model and the combined model of read count and allelic ratio (RCAR) are both 7-state HMMs, which enable the detection of underlying fetal conditions and the parental origin of extra chromosomes. The AR model utilizes allelic ratios, and the RCAR model utilizes read counts and allelic ratios of sequential informative loci to detect the underlying fetal condition (euploidy, maternally or paternally originated trisomy) of each target chromosome (Figure 3B). Both models classify informative loci into seven distinguishable states by the allelic patterns of maternal and fetal genotypes and fetal condition (Table 7).

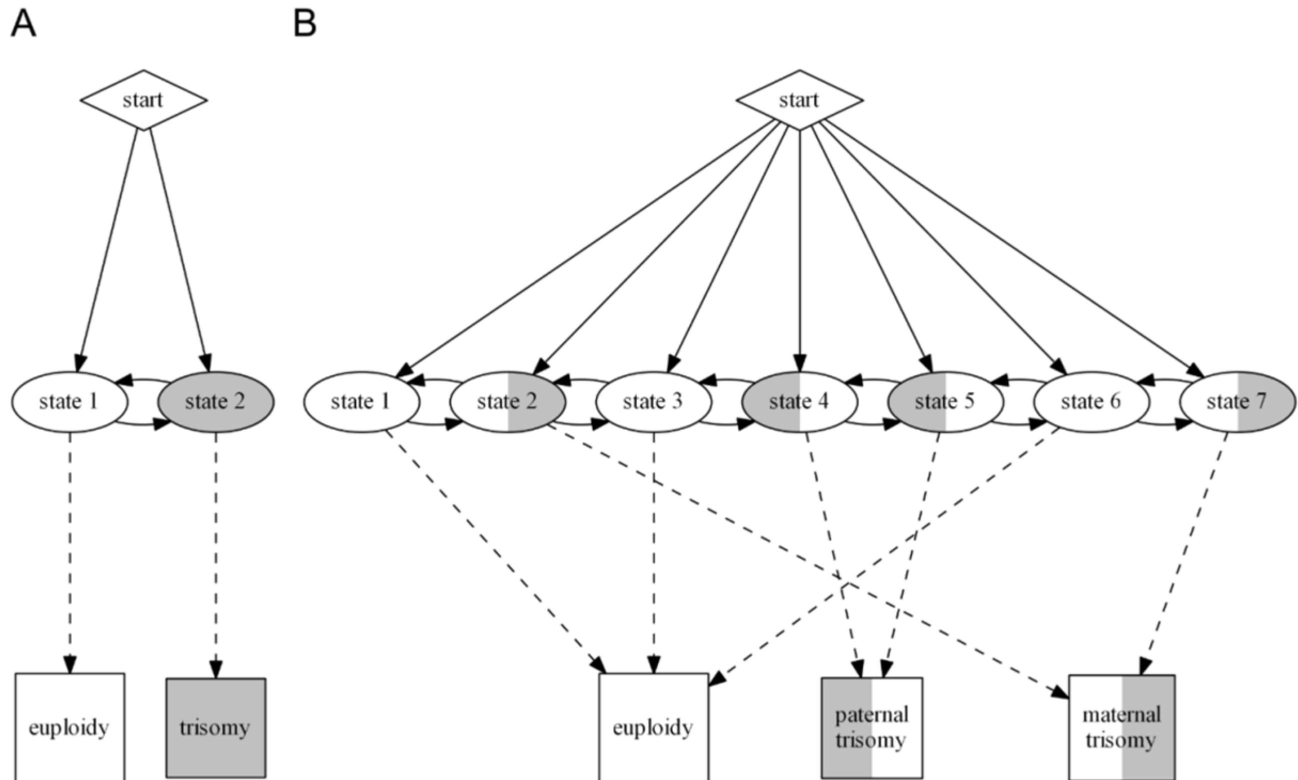

Figure 3. The architecture of 2-state and 7-state hidden Markov models (HMMs). (A) The 2-state HMM classified sequential targeted loci into two pre-defined states representing fetal euploidy (white) and trisomy (grey) by read count. (B) The 7-state HMM classified loci into seven pre-defined states representing fetal euploidy (white), maternally (white-grey) and paternally originated trisomy (grey-white) by allelic ratio with or without read count. Adopted from (Teder et al., 2019).

For model parameter estimation, no prior distribution of the initial state was assumed in any of the three HMMs. Each possible fetal condition had an equal likelihood of occurrence. The HMM transition probability was fixed at a state-to-state stay-switch ratio of 10, which means that any studied locus stays ten times more likely in the previous state instead of switching to a state describing a different fetal condition. The HMM emissions were assumed to follow the Gaussian distribution. The distribution parameters were estimated for each state using the simulated training dataset. The mean and variance of sequencing read counts and allelic ratios were calculated over the observed parameter space at continuous intervals using 100 simulated samples with corresponding fetal fraction and sequencing coverage. The developed models with parameters are publicly available (Teder, 2018).

6.2.3. Supplemental machine learning methods (Paper II)

Decision tree (DT) and support vector machine (SVM) methods were used to improve the chromosome classification, especially in the case of paternally originating trisomy. Both models used read depth, fetal fraction, and state proportions of HMM-classified loci as input for chromosome classification. The DT and the SVM methods were implemented in Python version 3.5.5 (https://www.python.org/) using the *scikit-learn* package version 0.19.1 (https://scikit-learn.org/). The DT model was created using the default parameters, except the maximum depth of the tree (*max_depth = 3*) and random state generator (*random_state = 123*) were changed after some systematic testing. The SVM model was also created using the default parameters with a fixed random state generator (*random_state = 123*). As the DT and SVM are both supervised learning models, the simulated training dataset (with known fetal fraction and sequencing coverage) was used to train both models.

6.2.4. Fetal fraction estimation for targeted sequencing approach (Paper II)

Fetal fraction is defined as the proportion of cffDNA in total cfDNA. In a TAC-seq assay, a fetal fraction of a sample was estimated using the allelic counts of the reference chromosomes (chr2 and chr3). First, the informative loci on the reference chromosome were filtered (allelic ratio > 2.5). In this subset, the major allele count was the sum of maternal allele counts and 1/2 of the fetal allele count. The minor allele count was proportional to 1/2 of the fetal allele count. The estimated fetal fraction was the median value of 2 times minor allele count over the sum of major and minor allele counts over all informative loci (Formula 1).

$$fetal\ fraction = median_{i=0}^{n}\left(\frac{2 \times min_i}{max_i + min_i}\right)$$

Formula 1. Fetal fraction calculation for targeted sequencing data. max_i − the major allele count of locus i, and min_i − the minor allele count of locus i. Adopted and modified from (Teder et al., 2019).

6.2.5. Fetal fraction estimation for whole-genome sequencing approach (Paper III)

Fetal fraction estimation was carried out on clinical patient samples in the validation set using SeqFF software and the computational pipeline developed for this purpose (Kim et al., 2015). First, raw sequencing reads were mapped with Bowtie2 version 2.3.4.1 against the human reference genome (GRCh37) with the preset option (*--very-sensitive*) in end-to-end mode (Langmead & Salzberg, 2012). Next, mapped reads were filtered by mapping quality (*-q 30*) using Samtools (Danecek et al., 2021). Finally, the remaining high-quality mapped reads were used to estimate fetal fraction using SeqFF software.

6.2.6. Detection of fetal trisomies (Paper III)

Fetal aneuploidy detection was carried out on clinical patient samples in the validation set using a NIPTmer software (Sauk et al., 2018) based pipeline. First, raw sequencing reads were converted to unique k-mers and counted by chromosome using GenomeTester4 software (Kaplinski et al., 2015). K-mer length of 25 bp was used. Next, a linear regression model was trained for each chromosome of interest using k-mer counts of other autosomes and the GC percentage of the reference set as predictors. Finally, samples were analyzed by standardizing (z-scoring) the normalized difference between observed and predicted k-mer counts. Obtained z-scores were used as indicators of aneuploidy risk for each chromosome of interest.

6.2.7. Statistical analysis

Statistical analysis was performed using R statistical software (https://www.r-project.org/). Pearson linear correlation was used to evaluate the relationship between the input and detected molecules. One-tailed Welch's t-test was used to assess the differences between molecule counts of the reference and target chromosomes. Statistical modeling was implemented in Python (https://www.python.org/) using *hmmlearn* (https://hmmlearn.readthedocs.io/) and *scikit-learn* (https://scikit-learn.org/) packages.

6.3. Ethics

The study conducted with clinical patients (Paper III) was approved by the Research Ethics Committee of the University of Tartu (#246/T-21) and written informed consent was obtained from all 408 participants whose blood samples were collected at Tartu University Hospital (Tartu, Estonia) and East-Tallinn Central Hospital (Tallinn, Estonia) during 2015 to 2018.

7. RESULTS

7.1. Paper I

7.1.1. TAC-seq assay

TAC-seq assay enables precise biomarker molecule counting. The assay uses oligonucleotide detector probes to quantify targeted biomarkers such as DNA or RNA molecules. Each targeted molecule is identified by a pair of detector probes, which incorporate a detector sequence specifically designed to be complementary only to the targeted region of the molecule under interest. In addition, detector probes also include unique molecular identifiers (UMIs), which enable deduplication and error correction introduced by PCR amplification and sequencing. As a result, targeted molecules are quantified to absolute molecule counts (Figure 4).

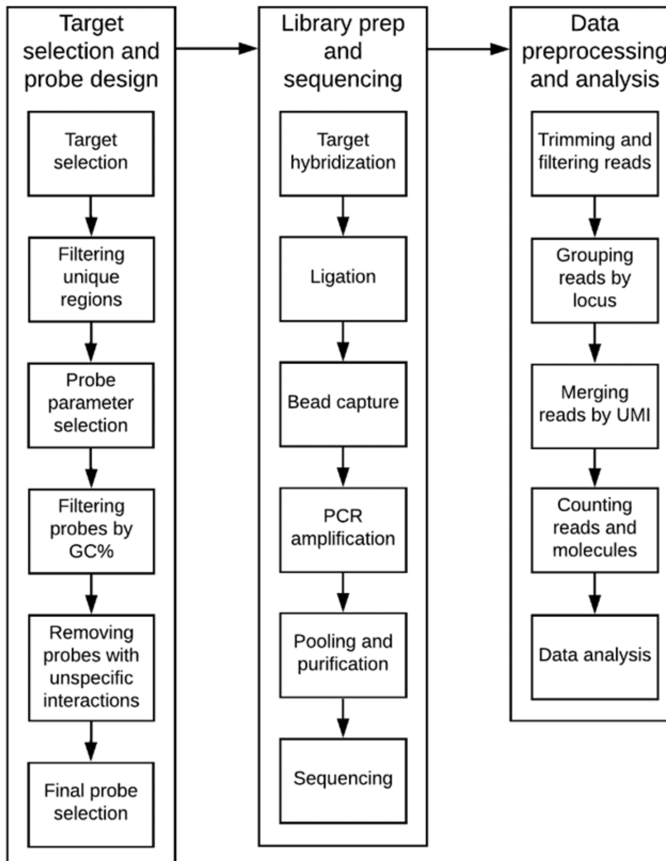

Figure 4. Overview of TAC-seq workflow. The workflow of TAC-seq assay includes three main steps: (1) target selection and probe design, (2) library preparation and sequencing, and (3) data preprocessing and analysis.

7.1.1.1. Target selection and probe design

TAC-seq detector oligonucleotide probes consist of a left and a right probe, which are nearly structurally symmetrical. Both detector probes include (1) a 27 bp target-specific detector sequence, (2) a 4 bp UMI motif, and (3) a universal sequence for Illumina sequencing technology. Also, the right detector oligonucleotide is 5'-phosphorylated to catalyze the formation of a phosphodiester bond between the pair of probes in library preparation. TAC-seq detector probes work as a pair. The left and the right probe need to hybridize on the same molecule so that the left detector 3'-end and the right detector 5'-end are located one after another for the target to be detectable.

7.1.1.2. Library preparation and sequencing

TAC-seq is a next-generation sequencing based assay, which uses specifically designed oligonucleotide detector probes for target identification and quantification. Once the detector probes are hybridized to the target molecule, a thermostable ligase catalyzes the formation of a phosphodiester bond between the 5′-phosphate and the 3′-hydroxyl group of the probe pair. Next, ligated detector-target complexes are captured using magnetic beads and amplified by PCR to introduce the sample-specific indices and other common motifs that are required for single-read next-generation sequencing. TAC-seq is a single tube assay, where ligated detector probes are captured, amplified, and identified by sequencing. The design minimizes the risk of losing studied biomolecules (Figure 5).

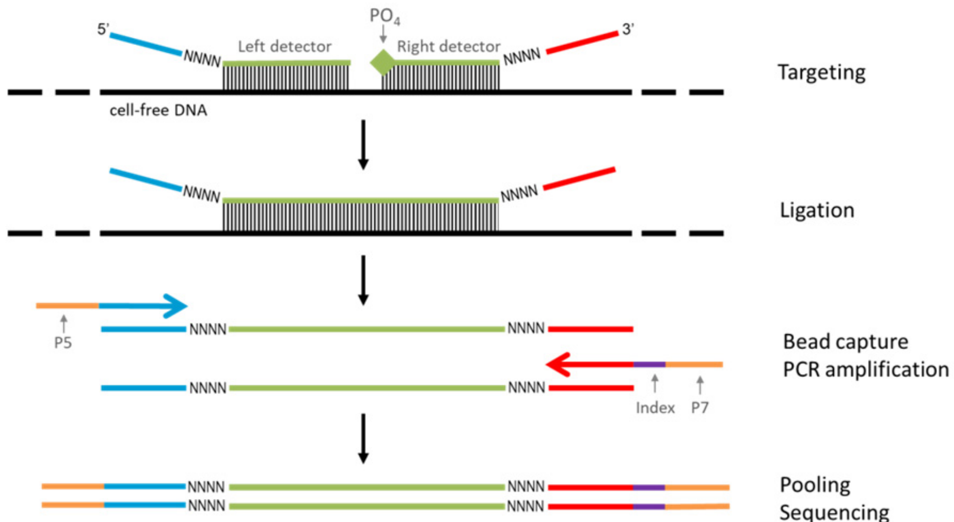

Figure 5. Library preparation of TAC-seq assay. The target-specific DNA oligonucleotide detector probes hybridize to the studied cfDNA molecule. Both detector probes include a 27 bp target-specific detector sequence (green), a 4 bp unique mole-

cular identifier motif (NNNN), and a universal sequence for Illumina sequencing technology (blue and red). The right oligonucleotide detector probe is 5′-phosphorylated for ligation. After hybridization, the detector probe pairs are ligated using a thermostable ligase. Next, the ligated detector complexes are captured with magnetic beads. Captured complexes are PCR amplified to introduce sample-specific indices for pooling and other common motifs that are required for single-read sequencing. Adopted and modified from (Teder et al., 2018).

7.1.1.3. Data preprocessing and analysis

In data preprocessing, sequencing reads of the targeted biomolecules are converted to read and molecule counts. An open-source tool called TAC-seq data analysis was developed for this purpose (Teder & Paluoja, 2018). TAC-seq data analysis command-line interface wraps FASTX-Toolkit for sequencing data operations with FASTQ and FASTA files (Hannon, 2010). Each sequencing read is trimmed, keeping only the target-specific detector sequence and UMI motif. Short reads with incomplete length are discarded. The remaining sequencing reads are grouped by target loci while tolerating a user-defined number of mismatches. Grouped reads with identical UMI sequences likely originate from the same molecule and, therefore, are merged. In the final step, grouped and merged reads are counted as read and molecule counts, respectively. The resulting read and molecule counts for each target locus can be analyzed according to the study purpose.

7.1.2. Experimental evaluation of TAC-seq

The External RNA Controls Consortium RNA spike-in controls were used to validate the technical sensitivity and accuracy of the TAC-seq method. For the experiment, 22 spike-in control sequences with known concentrations, ranging from 1 to 3×10^5 molecules, were targeted and quantified using TAC-seq assay. The experiment demonstrated a high correlation (Pearson correlation coefficient of log-transformed molecule counts > 0.98) between the input and the detected number of molecules at different UMI thresholds (0 to 10) (Figure 6).

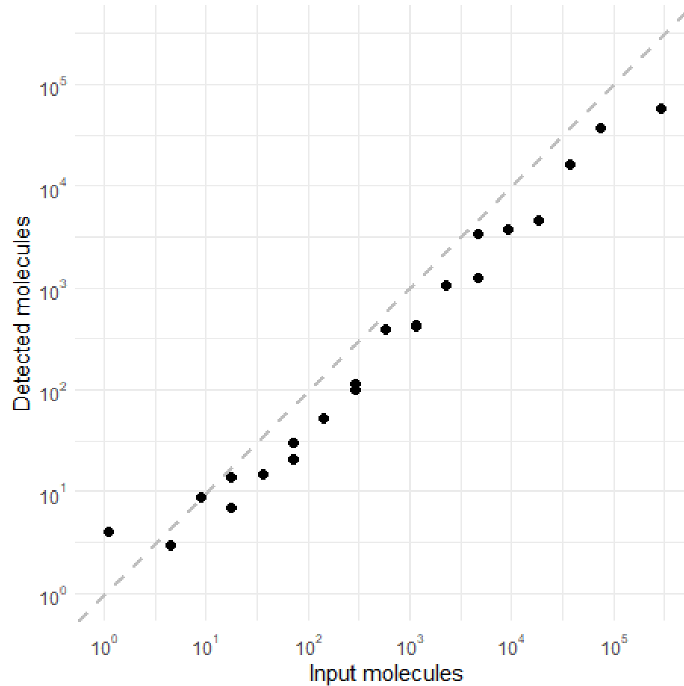

Figure 6. Detecting input molecules with TAC-seq assay. The experiment included 22 synthetic RNA spike-in control sequences with known concentrations ranging from 1 to 3×10^5 molecules. Unique molecule count was set to four to remove PCR amplification and sequencing error, which is mostly relevant when targeting highly similar sequences. The number of input and detected molecules with different concentrations show high correlation. Adopted and modified from (Teder et al., 2018).

With a UMI threshold of four, the average read count decreased from 1.5×10^6 to 5.7×10^3 molecules, demonstrating over 100-fold PCR redundancy. These results suggested that conservative UMI thresholds, which tolerate sequencing errors better, are justified and applicable for high-coverage sequencing, in which the read count is significantly higher than the UMI-corrected molecule count.

7.1.3. Detection of trisomy using TAC-seq assay

A proof-of-principle experiment was carried out to demonstrate the trisomy detection assay with the TAC-seq methodology. For the corresponding *in vitro* experiment, different proportions of genomic DNA of non-T21 and T21 cell lines were mixed to imitate cfDNA samples with predetermined fetal fractions. Multiple loci of the artificially mixed samples were targeted and quantified using a TAC-seq assay. Molecule counts with various UMI thresholds (0 to 3) were used to detect putative trisomies in each sample. If the mean molecule

count of the targeted loci on chr21 was significantly higher compared to the loci on euploid reference chromosome(s), then the sample was classified as trisomic ($p < 0.05$, one-tailed Welch's t-test).

Two independent experiments were carried out. Experiment 1 included 16 samples and 114 targeted loci on chr2 and chr21. Experiment 2 included 12 samples and 224 targets on chr2, chr3, and chr21. In both experiments, all euploid samples were correctly classified as non-trisomic. A statistically significant difference was observed in samples with 10% or higher trisomic proportion with few exceptions in different UMI thresholds. As expected, the difference was positively correlated with the trisomic DNA proportion of the synthetic samples. The higher the proportion of T21 cell line in a sample, the greater the difference between the mean molecule counts of chr21 and reference chromosome(s) (Figure 7).

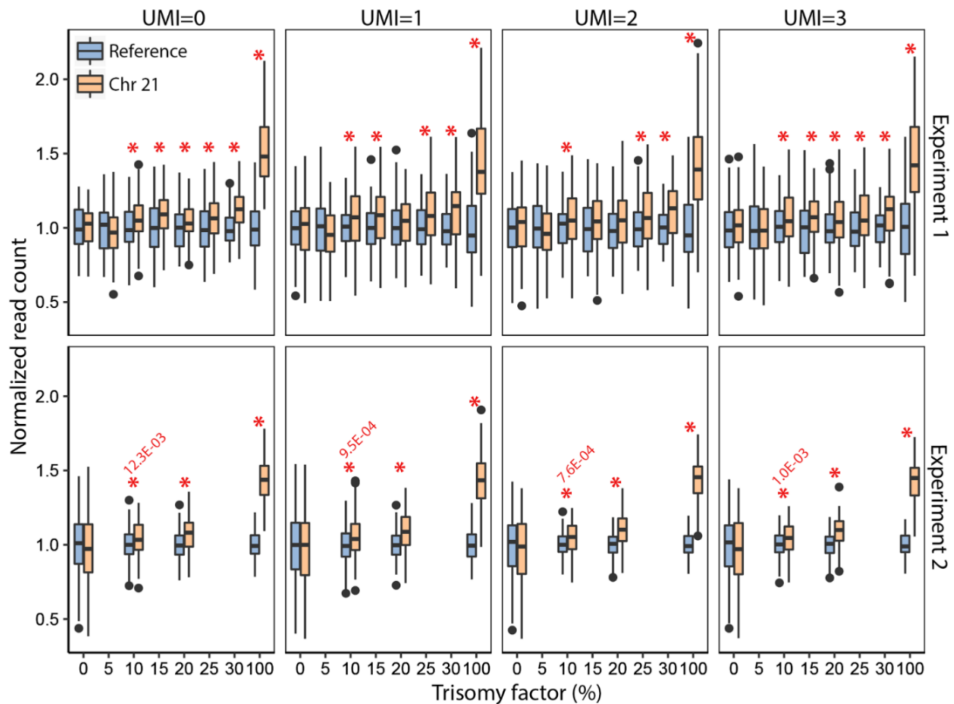

Figure 7. Trisomy detection with in vitro developed synthetic samples. Boxplots of normalized molecule counts of TAC-seq experiments at various trisomy factor percentages (trisomic cell proportions). Experiment 1 (upper four plots) involved 114 targeted loci along chr2 and chr21. Experiment 2 (lower four plots) involved extended TAC-seq probe set with 224 targeted loci along chr2, chr3, and chr21. One biological replica is depicted at various UMI thresholds. The red asterisk indicates a significant difference between the normalized molecule counts of the reference chromosomes (2 and 3) and the chromosome of interest (21) of a studied sample ($p < 0.05$, one-tailed Welch's t-test). Adopted from (Teder et al., 2018).

7.2. Paper II

7.2.1. Computational NIPT analysis framework for targeted sequencing assays

The developed computational analysis framework enables the detection of fetal chromosomal trisomy from targeted high-coverage sequencing data. This computational framework incorporates a 2-phase analysis. Firstly, sequentially targeted loci are classified with an appropriate HMM to find the most likely fetal condition behind each locus. Secondly, the sequentially classified loci are in turn used to classify the entire chromosome by additional machine learning (ML) methods. As a result, the developed framework enables to detect fetal trisomy and the parental origin of an extra chromosome (Figure 8).

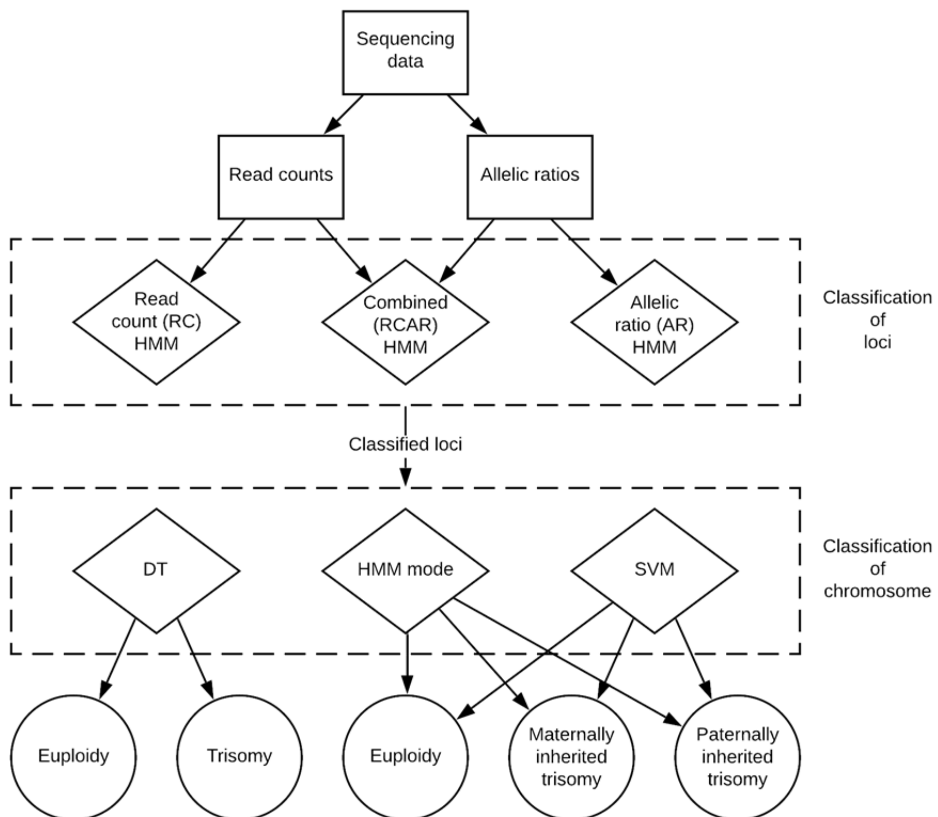

Figure 8. Overview of the developed 2-phase computational framework. The computational framework included two steps: (1) classification of loci by chromosome using hidden Markov model (HMM) based read count (RC), allelic ratio (AR), and combined (RCAR) models; and (2) classification of a chromosome using HMM mode, decision tree (DT), or support vector machine (SVM) as euploidy or (maternally or paternally inherited) trisomy. Adopted and modified from (Teder et al., 2019).

To evaluate the developed computational analysis framework for fetal trisomy detection, comprehensive simulation-based experiments were conducted. The *in silico* simulated datasets covered a wide variety of relevant variables such as fetal chromosomal condition (euploidy, maternally and paternally inherited trisomy), read depth (500 to 15,000 reads per locus), and different fetal fraction proportions (1 to 20% of total cfDNA).

7.2.1.1. Locus classification

In order to classify sequential targeted loci into predefined states representing different fetal conditions like euploidy and trisomy (maternally and paternally inherited trisomy), we developed three different HMMs, considering possible input data from currently available targeted sequencing assays. The developed models considered either read count data, allele-informative count data for polymorphic loci, or both aforementioned data types. The targeted loci were classified by chromosome into predefined states using the Viterbi algorithm, which can find the most likely underlying state of each locus that matches best to a given HMM.

The read count (RC) model is a 2-state HMM, which uses sequencing read counts to classify sequential loci into two predefined fetal conditions – euploidy or trisomy. The RC model is not able to determine the parental origin of the trisomy nor estimate the fetal fraction as maternally and paternally inherited alleles are indistinguishable using sequencing read count data. To take this limitation into account, an RC model with a fixed fetal fraction of 10% was used for comparison. The RC model with fixed fetal fraction showed excellent results detecting fetal euploidy on simulated datasets (accuracy > 0.99). On the other hand, this method was ineffective for detecting fetal trisomy if the fetal fraction was lower than 6% (accuracy = 0.11). Increasing the read depth induced only a minor increase in detection accuracy (Table 8).

The allelic ratio (AR) based model is a 7-state HMM, which uses allelic ratios (calculated using allele-informative count data for polymorphic loci) to classify the sequential loci into seven predefined states representing three fetal conditions – euploidy, maternally and parentally inherited trisomy. The AR model showed excellent results detecting fetal euploidy on simulated datasets even at a fetal fraction of 1% and read depth of 500 (accuracy > 0.99) and reasonable accuracy detecting maternally originated trisomy if fetal fraction was equal to or greater than 6% and read depth was over 10,000. The model was unable to distinguish paternally inherited trisomy from maternally inherited trisomy in a given range of read depth and fetal fraction (Table 8).

The combined (RCAR) model is a 7-state HMM, which uses both read count and allelic ratio information to classify sequential loci into seven predefined states representing three fetal conditions – euploidy, maternally and paternally inherited trisomy. The RCAR model showed excellent results detecting fetal euploidy on simulated datasets (accuracy > 0.99). However, the model was

unable to detect paternally originated trisomy in a given range of read depth and fetal fraction (Table 8).

7.2.1.2. Chromosome classification

Sequential loci classified by HMM in the previous step were used to classify the entire chromosome between three studied fetal conditions – euploidy, maternally, and paternally inherited trisomy. In addition to HMM mode, where the decision was based on the most frequently observed state, two supplemental ML methods for multi-class classification were also applied – decision tree (DT) and support vector machine (SVM). The supplemental ML models were integrated into the framework to improve chromosome classification results of the HMMs, especially in the case of paternally inherited trisomy. The results of each locus classification model (RC, AR, and RCAR) on *in silico* simulated datasets were used as input to supplemental ML model for the evaluation of fetal trisomy detection.

Fetal trisomy detection accuracy improved with the increase of fetal fraction and read depth for all three approaches. Also, using the supplemental ML methods for chromosome classification significantly improved the detection accuracy. The gain in classification accuracy was more significant in extreme situations, such as low fetal fraction and/or read depth (Table 8).

Table 8. Summarized fetal euploidy and trisomy classification accuracy with different computational models. Each value represents an average classification accuracy across 3,600,000 simulated cell-free DNA samples within a given fetal condition (euploidy, maternally and paternally inherited trisomy), fetal fraction (1 to 20%), and sequencing depth (500 to 15,000 reads). RC – read count model; AR – allelic ratio model; RCAR – combined model; HMM – hidden Markov model; DT – decision tree model; SVM – support vector machine model. Adopted and modified from (Teder et al., 2019).

Fetal fraction	RC (fixed fetal fraction)			AR			RCAR		
	HMM	DT	SVM	HMM	DT	SVM	HMM	DT	SVM
1–5%	0.56	0.93	0.95	0.34	0.69	0.69	0.49	0.91	0.90
6–10%	1.00	1.00	1.00	0.54	0.92	0.92	0.65	1.00	1.00
11–15%	1.00	1.00	1.00	0.63	0.98	0.98	0.67	1.00	1.00
16–20%	1.00	1.00	1.00	0.65	0.99	0.99	0.67	1.00	1.00
Overall	0.89	0.98	0.99	0.54	0.89	0.90	0.62	0.98	0.98

The read count based chromosome classification uses locus classification results of the RC model to classify chromosomes into two predefined conditions representing fetal euploidy and trisomy. The RC model with fixed fetal fraction offered the highest HMM-based chromosome classification accuracy (accuracy = 0.89). Applying supplemental ML methods significantly improved the classification at the lower range of fetal fraction (1 to 5%). The DT based model allowed accurate detection of fetal condition at a fetal fraction as low as 3%.

The SVM-based model had an even lower limit as the model allowed accurate detection even at 2% fetal fraction (Figure 9, Table 8).

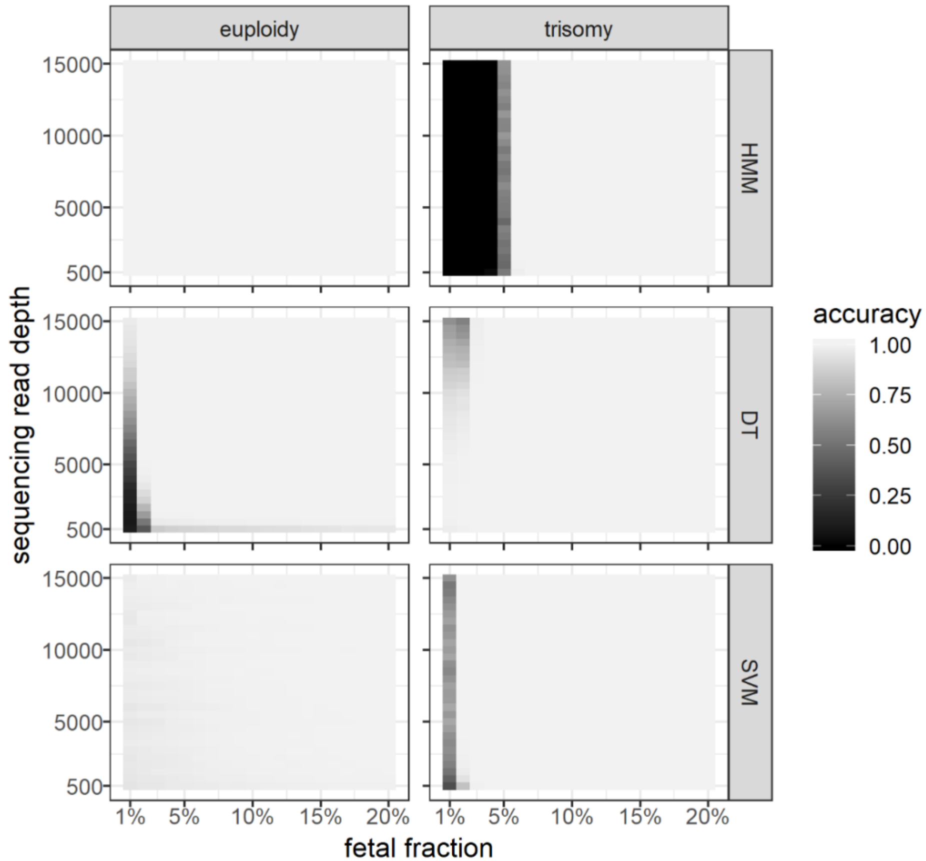

Figure 9. Classification accuracy of the read count (RC) model with fixed fetal fraction in conjunction with the machine learning models on simulated datasets. The simulated datasets of fetal euploidy and trisomy (vertical panels) were first classified by the RC model and the resulting class frequencies were further classified by the decision tree (DT) and support vector machine (SVM) and compared against the hidden Markov model (HMM) results (horizontal panels). Each panel includes cells with different fetal fractions (x-axis) and sequencing read depths (y-axis). Each cell includes 10,000 cell-free DNA samples and the color represents the model classification accuracy. Adopted from (Teder et al., 2019).

The allelic ratio based chromosome classification uses locus classification results of the AR model to classify chromosomes into three predefined fetal conditions – euploidy, maternally, and paternally inherited trisomy. The AR model offered the lowest HMM-based chromosome classification accuracy

(accuracy = 0.54). The model was not able to detect maternally inherited trisomy below 6% fetal fraction and paternally inherited trisomy at a given range of read depth. Supplemental ML methods increased the detection accuracy of the AR model results significantly in most cases. Both the DT and SVM based models allowed maternally inherited trisomy detection at 5% fetal fraction and paternally inherited trisomy detection at a given parameter space. However, the SVM-based method was unable to detect fetal euploidy at a read depth of 500 (Figure 10, Table 8).

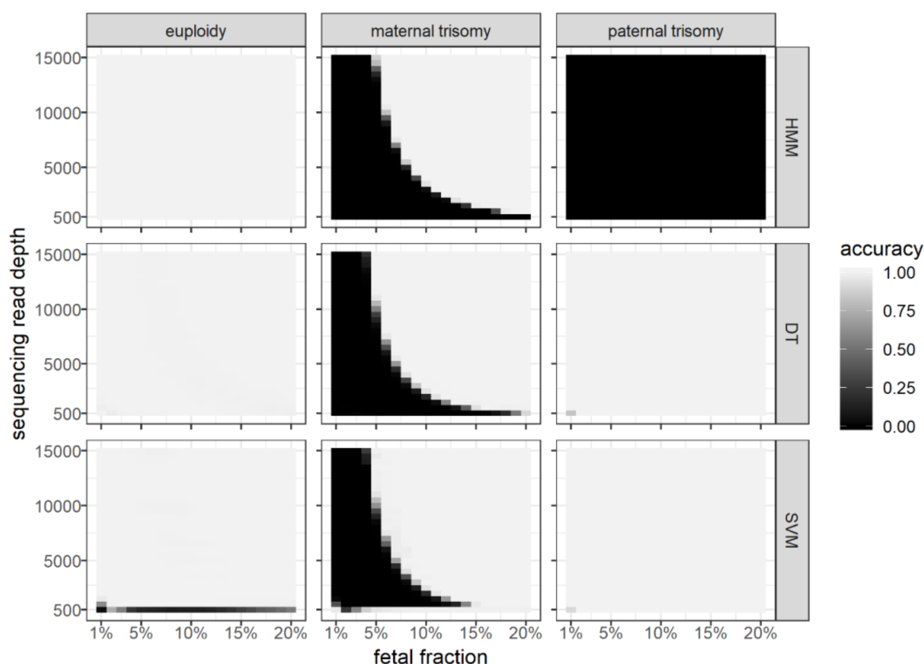

Figure 10. Classification accuracy of the allelic ratio (AR) model in conjunction with the machine learning models on simulated datasets. The simulated datasets of fetal euploidy, maternally and paternally trisomy (vertical panels) were first classified by the AR model and the resulting class frequencies were further classified by the decision tree (DT) and support vector machine (SVM) and compared against the hidden Markov model (HMM) results (horizontal panels). Each panel includes cells with different fetal fractions (x-axis) and sequencing read depths (y-axis). Each cell includes 10,000 cell-free DNA samples and the color represents the model classification accuracy. Adopted from (Teder et al., 2019).

The read count and allelic ratio based chromosome classification use locus classification results of the RCAR model to classify chromosomes into three predefined fetal conditions – euploidy, maternally, and paternally inherited trisomy. The RCAR model offered slightly improved chromosome classification results (accuracy = 0.62) compared to the AR model (accuracy = 0.54). The

model was not able to detect maternally inherited trisomy below 3% fetal fraction and paternally inherited trisomy at a given range of read depth. Supplemental ML methods increased the detection accuracy of the RCAR model results in most cases. Both the DT and SVM based models allowed maternally inherited trisomy detection at 2% fetal fraction and paternally inherited trisomy detection at a given parameter space. However, the supplemental methods decreased fetal euploidy detection results at low fetal fraction and read depth compared to the RCAR model (Figure 11, Table 8).

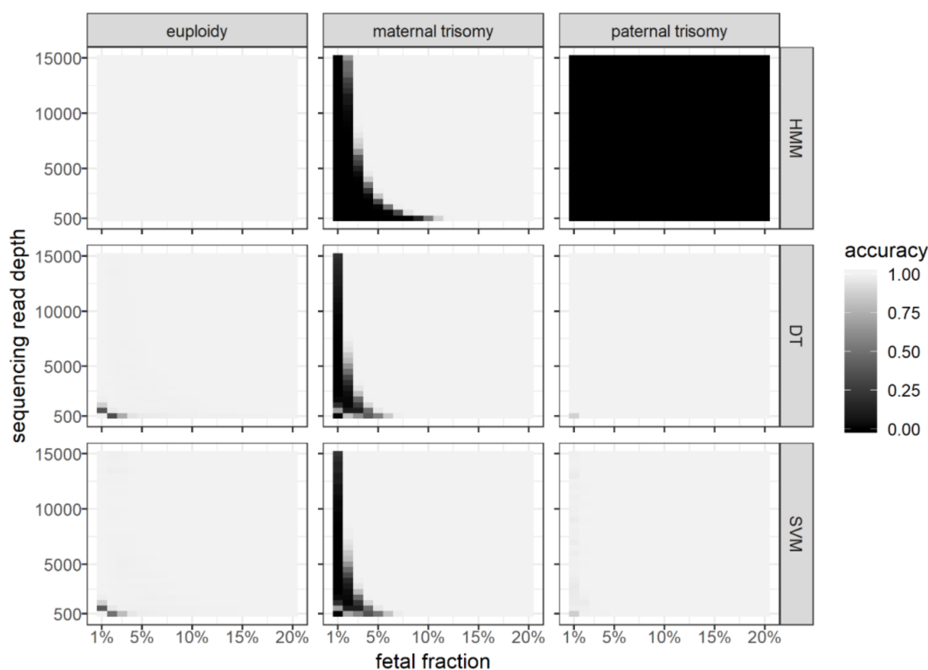

Figure 11. Classification accuracy of the combined (RCAR) model in conjunction with the machine learning models on simulated datasets. The simulated datasets of fetal euploidy, maternally and paternally trisomy (vertical panels) were first classified by the RCAR model and the resulting class frequencies were further classified by the decision tree (DT) and support vector machine (SVM) and compared against the hidden Markov model (HMM) results (horizontal panels). Each panel includes cells with different fetal fractions (x-axis) and sequencing read depths (y-axis). Each cell includes 10,000 cell-free DNA samples and the color represents the model classification accuracy. Adopted from (Teder et al., 2019).

7.2.2. *In vitro* analysis

To demonstrate the developed computational analysis framework in practice, the RC model was applied on an experimentally controlled dataset corresponding to fetal T21 pregnancy. The Experiment 2 dataset included 12 samples

with different proportions of DNA from a T21 cell line sample. First, the molecule counts of the targeted loci were cleaned. From the initial 99 targeted loci on chromosome 21, only 69 to 86 loci with 527 reads per sample remained on average after locus-wise outlier removal. Second, the molecule counts of sequential loci on chromosome 21 were classified with the developed RC model as either euploid or trisomic. In conclusion, the RC model showed high classification accuracy even with less than 100 targeted loci per chromosome of interest with a sequencing depth close to 500 reads per target (Figure 12).

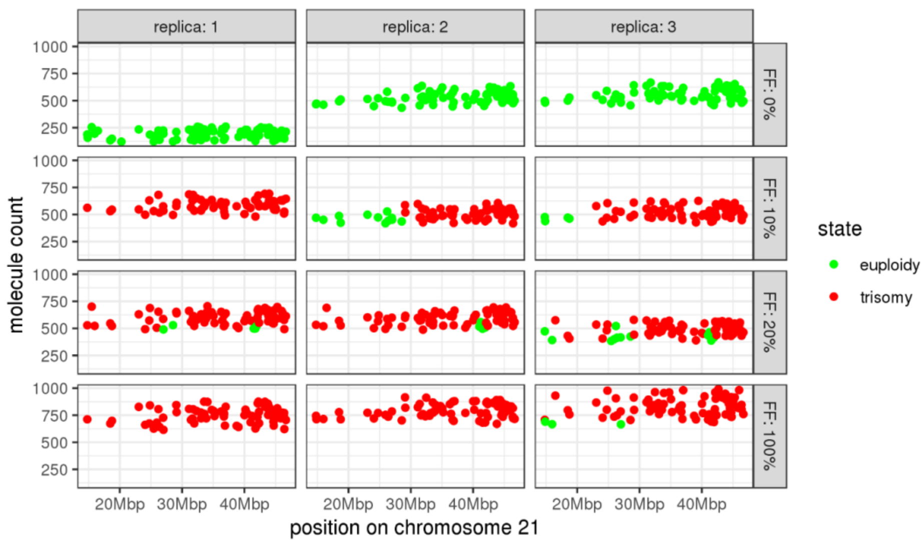

Figure 12. Classification of targeted loci using the read count (RC) model on experimental targeted sequencing dataset. Experimentally controlled synthetic trisomy samples were created by mixing different proportions of genomic DNA from a non-trisomy 21 cell line with a genomic DNA from a trisomy 21 cell line to imitate different proportions of fetal fraction (FF), each in three replicas. Genomic DNA was sheared to 150 to 200 bp fragments by sonication to mimic cfDNA. DNA fragments were hybridized and sequenced using TAC-seq detector probes specifically designed to target only the reference chromosome 2 (68 targeted loci) and chromosome 3 (60 targeted loci), and the studied chromosome 21 (99 targeted loci). Read counts of targeted loci were converted to absolute molecular counts using unique molecular identifiers with the threshold of one to reduce the amplification bias. Obtained absolute molecule counts of each sample were filtered by interquartile range to remove outliers and used as input to RC model to classify each targeted locus either euploidy (green color) or trisomy (red color). Sequentially classified targeted loci per sample are visualized as colored dots. Adopted from (Teder et al., 2019).

7.3. Paper III

A total of 424 clinical patient samples were collected to validate the low-coverage WGS-based NIPT analysis pipeline for Estonian prenatal clinical care. The initial sample set included 259 high-risk and 149 general population samples from the Estonian cohort and 16 samples from the Belgium cohort (enriched for aneuploidy cases). Four samples diagnosed with rare chromosomal aberrations not included in the validation study were removed from the analysis. Therefore, the final test set included 255 high-risk and 16 Belgium samples that were invasively confirmed either by CVS or AC. Additional 149 general population samples were used as the initial reference set.

7.3.1. Analysis pipeline

A computational analysis pipeline was developed to integrate NIPT into prenatal clinical care in Estonia. The developed NIPT pipeline includes sequencing data quality control, fetal aneuploidy detection, fetal fraction estimation, and fetal sex determination. Prior to the computational NIPT analysis, the sequenced cfDNA samples of pregnant women undergo routine quality control. Samples passing this inspection are further analyzed by two separate workflows. First, the NIPTmer workflow is used to detect fetal aneuploidy and determine fetal sex. In sequencing data preprocessing, raw sequencing reads are used to create a list of k-mers for each chromosome. The k-mer lists are cleaned from non-unique k-mers, potentially polymorphic k-mers (overlapping with SNPs), and k-mers from difficult-to-interrogate regions of the human genome (e.g., centromeres, telomeres, and pseudo-autosomal regions). NIPTmer uses a linear regression model trained on reference samples to predict the expected k-mer counts for each chromosome. The difference between the measured and predicted k-mer counts is used to calculate the z-scores for each chromosome of interest. The z-scores with pre-defined threshold values are used to detect fetal aneuploidy and determine fetal sex. Second, the SeqFF workflow is used to estimate the fetal fraction of a sample. For the preprocessing, raw sequencing reads are aligned against the human reference genome. Next, aligned reads are filtered by mapping quality and sorted by coordinates. Last, the preprocessed reads are analyzed by SeqFF software to estimate the fetal fraction of a sample (Figure 13).

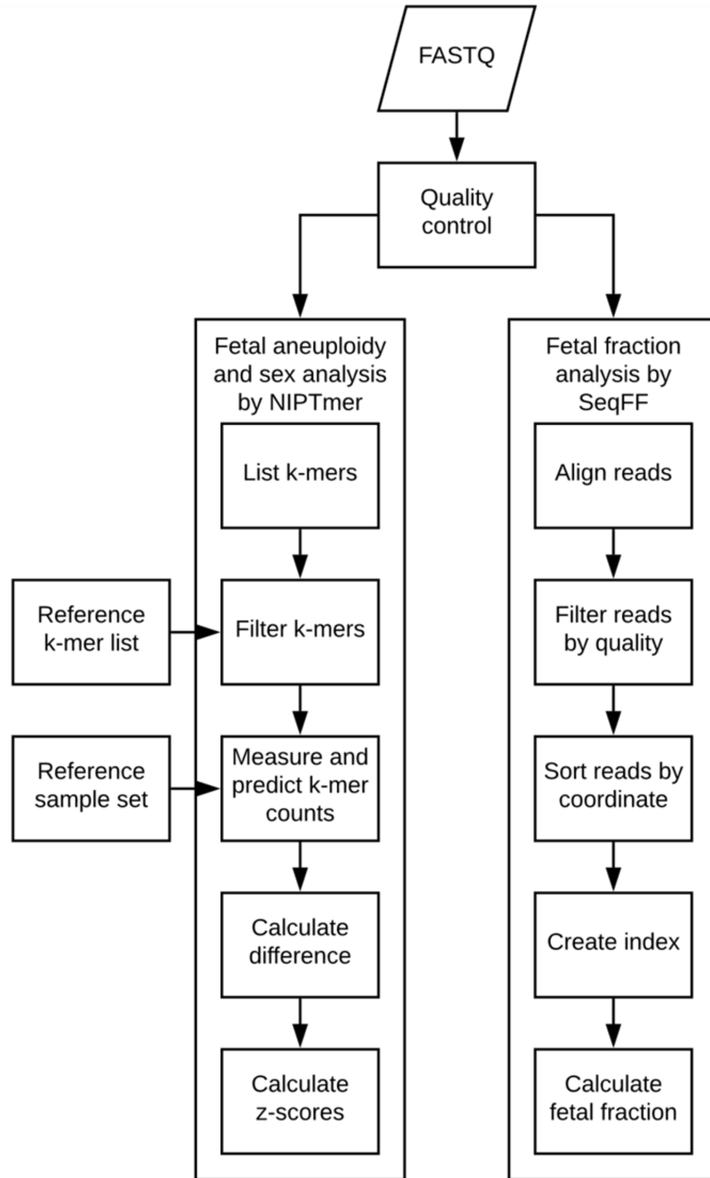

Figure 13. Developed non-invasive prenatal testing (NIPT) analysis pipeline. The developed NIPT analysis workflow uses a sequenced cell-free DNA sample of a pregnant woman as input. that have passed quality control as an input to detect fetal aneuploidy, estimate fetal fraction, and determine fetal gender.

7.3.2. Fetal fraction analysis

Fetal fraction was estimated using SeqFF software. The mean fetal fractions were 10.1% (SD = 6.7%) and 16.7% (SD = 8.4%) for the test and reference sets, respectively. Approximately 3% (8/271) of the test samples had fetal fraction below 4%, a widely accepted minimum for reliable detection of common trisomies (Canick et al., 2013; Ehrich et al., 2011; Gregg et al., 2016; Norton et al., 2012) (Figure 14). Despite that, all samples with low fetal fraction, including two aneuploidy cases (T13 and T18), were correctly identified as euploid or aneuploid.

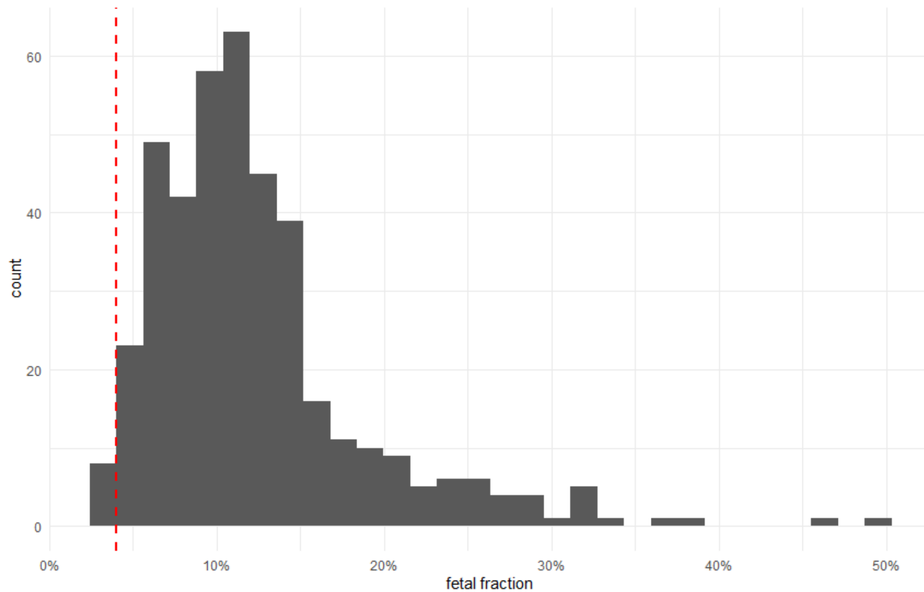

Figure 14. Fetal fraction distribution. A total of 271 test samples in a validation study were analyzed. The mean fetal fraction was 10.1%. About 3.0% (8/271) of the samples had fetal fraction below the lower detection boundary of 4% and were eliminated from the following fetal aneuploidy analysis.

7.3.3. Fetal aneuploidy analysis

Fetal aneuploidy was estimated on the test set using NIPTmer software. The three most common fetal autosomal trisomies, such as T21, T18, T13, and an allosomal 45,X were validated in this study. For autosomal trisomies, z-score of 3.0 was used as a threshold value for trisomy detection, which should theoretically include approximately 99.9% of all euploid samples. This allowed correct identification of all known non-mosaic T21 (15/15), T18 (9/9), and T13 (4/4) cases. 0.7% (2/271) of the test set samples had false positive test results. These include two false positive T18 cases, while no false positive cases were observed in the case of T21 and T13. These false positive T18 cases were

confirmed healthy by an invasive procedure. The sensitivity of NIPTmer was 100% for T21, non-mosaic T18, and T13, and the specificity was 100% for T21 and T13, and 99.2% for T18 (Table 9).

In addition to autosomal trisomies, the test set included four fetal 45,X cases. Female fetus with chromosome X deficit (z-score < -6.5) was the high-risk indicator for 45,X. All 45,X cases (4/4) were correctly identified (z-scores ranging from -8.1 to -16.4), and there were no false positive results. The sensitivity and specificity were 100% for fetal 45,X (Table 9).

Table 9. Non-invasive prenatal testing validation study for common fetal aneuploidies (T21, T18, T13, and 45,X). The table includes 255 high-risk Estonian and 16 Belgium cohort samples. 149 general population samples, used as a reference set, are excluded from the results, as well as 4 samples diagnosed with other chromosomal aberrations were not included in the final validation set.

Fetal condition	NIPT result					Total	Sensitivity	Specificity	Positive predictive value
	Low-risk	High-risk							
		T21	T18	T13	45,X				
Euploidy	237		2			239			
T21		15				15	100%	100%	100%
T18			9			9	100%	99.2%	81.8%
T13				4		4	100%	100%	100%
45,X					4	4	100%	100%	100%
Overall	237	15	11	4	4	271	100%	99.2%	94.1%

7.3.4. Fetal sex determination

Fetal sex was estimated on the invasively confirmed Estonian samples using NIPTmer. The test sample set included 123 female (48.2%) and 132 male (51.8%) fetuses. Four different sex chromosome (allosome) models were used to determine the fetal sex. Unlike the autosomal models, the allosomal models used reference samples of corresponding sex to estimate the model parameters for the female and male models separately. As a result, fetal sex was determined correctly for all tested samples (255/255), including 45,X cases (4/4).

8. DISCUSSION

Prenatal screening helps to detect congenital abnormalities at early stage. Detected high risk for chromosomal aneuploidy is confirmed by an invasive diagnostic procedure, such as AC or CVS followed by karyotyping. Unfortunately, the invasive procedure poses a small risk of miscarriage. To avoid unnecessary risk of miscarriage and stress for pregnant women, prenatal screening methods have to be accurate and cost-effective at the same time. A relatively new non-invasive approach called NIPT has been recently introduced to prenatal screening practice. Compared to conventional non-invasive prenatal screening methods, such as maternal serum biomarker analysis and fetal NT scan, NIPT is more accurate at detecting fetal aneuploidy and produces significantly fewer false positive test results. Unfortunately, NIPT is more expensive due to the laboratory and sequencing costs. Therefore, different combinations of prenatal methods, such as FCT and NIPT, should be used for optimal balance between accuracy and cost-effectiveness in a clinical setting. Continuous advancements in throughput and availability of sequencing platforms have significantly reduced the sequencing cost in recent years. These developments have allowed a wider introduction of NIPT. Currently, many European countries have adopted NIPT into their national healthcare programs. For example, in Belgium and Netherlands, NIPT is offered for all pregnant women. Then again, most other European countries (e.g., Estonia), offer NIPT only for high-risk patients after the FCT (Gadsbøll et al., 2020).

This dissertation introduces two different NIPT approaches for fetal aneuploidy detection based on targeted sequencing and low-coverage WGS. To use the targeted sequencing approach, a novel method called TAC-seq was developed, which enables precise nucleic-acid based biomarker quantification. Similar to other targeted assays, TAC-seq uses oligonucleotides, which are specifically designed to detect pre-defined genomic loci of interest. In addition, TAC-seq locus-specific oligonucleotides also incorporate UMI sequences, which enable to reduce possible quantitative bias introduced by the PCR amplification step. Applying TAC-seq on a cfDNA sample of a pregnant woman enables to differentiate imbalances of targeted chromosomal regions by taking reference chromosome into account as a baseline. Compared to the WGS-based approach, TAC-seq quantifies only targeted regions, which has its pros and cons. By targeting specific chromosomes or genomic regions of interest, one can be confident that selected regions are sufficiently covered, which might not be the case with low-coverage WGS. Also, targeting allows local high-resolution investigation even at the gene exon level. This advantage allows targeted approaches to be used for detecting small clinically important regions which could be missed by a low-coverage WGS-based approach. Most importantly, targeted approaches need less cfDNA and sequencing reads per sample, and thus more samples can be sequenced in parallel, therefore, reducing the sequencing cost per sample. As sequencing cost and coverage are inversely

proportional and play a major role in successful and accurate analysis, it significantly affects the overall cost of NIPT. The limitation of TAC-seq is that only selected regions, which are unique in the human genome, can be targeted. Also, there is a theoretical upper limiting number of regions that can be targeted in a single reaction as the risk of cross-hybridization between detector oligonucleotide probes increase with a larger probe set. To date, TAC-seq assay with up to a few hundred target loci has been tested in practice (Teder et al., 2018, 2019; Wedenoja et al., 2020). Experiments with an upper limiting number of targets are yet to be carried out. For further research, it would be interesting to see if TAC-seq assay with variant-specific detector oligonucleotide probes can distinguish maternal and fetal genotypes. Targeting clinically relevant genetic variants would offer new opportunities for prenatal screening in addition to fetal aneuploidy detection. Additionally, TAC-seq assay and its application in NIPT should be clinically validated on cfDNA samples of pregnant women.

A novel computational analysis framework was developed for targeted high-coverage sequencing based NIPT. Besides wider usage, this work was motivated by the development of TAC-seq aneuploidy screening assay. The framework incorporates 2-layer analysis to detect fetal trisomy and the parental origin of an extra chromosome. In the first phase of analysis, sequentially targeted loci are classified by HMM to find the most likely fetal condition behind each locus. The developed models use either read counts or allelic ratios of sequentially targeted loci as input. Although such naïve classification models work relatively well in optimal conditions, an accurate fetal fraction estimation is necessary, especially in edge cases (e.g., low fetal fraction and read depth), in order to control the precision and uncertainty of fetal trisomy detection. The fetal fraction of a studied cfDNA sample can be estimated using models, which include allelic ratios as one of the inputs. Natively, the read count based model is not able to distinguish maternal and fetal genotypes and thereby cannot be used to estimate fetal fraction. On the other hand, the read count model is robust as it only requires sequencing read counts as its input. As separate allele counts are not supported by all targeted approaches, a read count based model can be incorporated into currently used targeted approaches allowing wider adoption compared to allelic ratio based models. In the second phase of analysis, the sequentially classified loci (output of the first phase) are used to classify the entire studied chromosome as either euploid or trisomic by additional ML methods, such as DT or SVM. The developed computational NIPT analysis framework was effectively applied on *in silico* simulated euploid and trisomic samples (Table 8). The read count based model in conjunction with the SVM showed the best results in detecting fetal trisomy over a wider range of fetal fraction and read depth. It can detect fetal trisomy even at 2% of fetal fraction, which is lower than the widely used minimum threshold of 4% (Canick et al., 2013; Ehrich et al., 2011; Gregg et al., 2016; Norton et al., 2012). Although the read count model can be easily incorporated into all targeted sequencing based NIPT workflows, the combined model of read count and allelic ratio data is the recommended choice with allele counts, as it enables to estimate fetal fraction,

which is an important metric for analytical uncertainty assessment. In addition, the combined model allows to investigate the parental origin of extra chromosomes using allelic patterns (Table 7). For the proof-of-principle study, the read count based model was applied on *in vitro* developed synthetical samples to demonstrate the feasibility of the approach. TAC-seq assay was used to target a few hundred loci on experimentally controlled samples corresponding to fetal T21 pregnancy with different fetal fraction intervals. The results with a relatively small number of targets (69 to 86 per chromosome of interest) and read depth (on average 527 reads per locus) demonstrate that the read count based model on TAC-seq assay can be used to successfully detect fetal trisomy (Figure 12). Further analysis and possible optimization with actual cfDNA samples are needed before the developed computational analysis framework can be used in practice. Furthermore, a systematic search for a suitable target set and optimal threshold values for different fetal conditions is needed. This is followed by validation studies with clinical samples before this novel approach can be applied for aneuploidy screening in clinical settings.

For low-coverage WGS-based NIPT, a dedicated NIPT pipeline, including laboratory and analytical workflows, was set up. The analytical workflow is based on NIPTmer software developed by our colleagues. NIPTmer takes the advantage of k-mers to detect fetal aneuploidies using WGS data. Compared to the read mapping based methods, the k-mer based approach is computationally less expensive and significantly faster. NIPTmer outputs z-scores for each chromosome, including allosomes, based on the difference between the observed and predicted k-mer counts using a linear regression model trained on the reference set of euploid samples. Calculated z-scores are used to detect fetal aneuploidy for chromosomes of interest (chr13, chr18, chr21, and chrX) and to determine fetal sex. The developed NIPTmer workflow was applied on clinical patients of Estonian (N = 255) and Belgium (N = 16) cohorts. The sensitivity of fetal trisomy was 100% for common aneuploidies (T21, T18, T13, and 45,X), and specificity was 100% for T21, T13, and 45,X and 99.2% for T18. The sensitivity and specificity of NIPTmer workflow for rare trisomies (T13 and T18) were higher than expected (Gil et al., 2017). This might be due to the low number of cases included in the validation study. For further optimization of NIPTmer, a more extensive reference set allows to estimate population parameters more precisely, which is important for decisions made on the basis of z-score threshold criteria.

Overall, both targeted and WGS-based NIPT approaches are one way to make prenatal screening for fetal aneuploidies more accurate and cost-effective.

9. CONCLUSION

First, we designed and implemented a new targeted sequencing assay called TAC-seq for precise biomarker molecule counting (Paper I). Although there are already several targeting sequencing methods in clinical use, TAC-seq assay takes a new approach by incorporating UMI sequences into specifically designed oligonucleotide detector probes, which alleviate quantification bias introduced by PCR amplification in the library preparation step. As a result, TAC-seq assay quantifies targeted biomarkers, such as DNA or RNA sequences, with a molecule-level precision. We applied TAC-seq on synthetic T21 control samples and evaluated trisomy detection accuracy at different concentrations of trisomic DNA. Also, I developed a command-line tool suitable for TAC-seq data analysis. TAC-seq data analysis tool is freely available software, which mediates the conversion of targeted sequencing data into read and molecule counts per targeted loci and sample with user-defined UMI threshold.

I also developed a computational NIPT analysis framework for targeted high-coverage sequencing data (Paper II). The 2-phase framework uses HMM in conjunction with DT or SVM model to detect chromosomal trisomy. In addition to fetal trisomy detection, the allelic ratio based model is able to distinguish the parental origin of an extra chromosome and supports fetal fraction estimation, which is an important quality metric for uncertainty assessment in NIPT analysis. A comprehensive evaluation of the developed computational analysis framework was conducted to determine the capabilities and limitations of fetal trisomy detection. For a proof-of-principle, we also used *in vitro* generated synthetic control samples corresponding to trisomic pregnancies with different fetal fractions. By targeting cfDNA samples of pregnant patients, TAC-seq assay can potentially be used in NIPT for fetal aneuploidy screening. Using TAC-seq with the developed computational NIPT analysis framework enables accurate fetal trisomy detection while being more cost-effective compared to the low-coverage WGS-based NIPT approaches, which after clinical validation can be used for routine clinical practice.

We set up, implemented, optimized, and validated a low-coverage WGS-based NIPT analysis pipeline for Estonian prenatal clinical care (Paper III). Based on our validation study results with Estonian and Belgium cohorts, the sensitivity of the developed NIPTmer-based aneuploidy detection workflow was 100% for common fetal aneuploidies (T21, T18, T13, and 45,X), and specificity was 100% for T21, T13, 45,X, and 99.2% for T18. As a result, we believe that the developed NIPT analysis pipeline can successfully be used as an efficient prenatal screening test in combination with an ultrasound scan for the first trimester fetal examination. Compared to the conventionally used FCT, NIPT has superior detection rate of common fetal aneuploidies and yields fewer false positive test results, which will significantly reduce the number of invasive procedures in prenatal screening.

10. SUMMARY IN ESTONIAN

Suunatud ja ülegenoomsel sekveneerimisel põhinevate mitteinvasiivsete sünnieelsete testide arvutusmeetodite ja töövoogude väljatöötamine

Aneuploidsus on patoloogiline seisund, kus rakus esineb üks või mitu üleliigset või puuduolevat kromosoomi. See on tingitud kromosoomide ebaühtlasest jagunemisest tütarrakkude vahel. Arvatakse, et rohkem kui 90% loote aneuploidiatest on tingitud emapoolsetest meiootilistest vigadest oogeneesis. Loote aneuploidiad on üks peamisi raseduse katkemise ja kaasasündinud arenguhäirete põhjuseid. Aneuploidsed rasedused moodustavad vähemalt 10% kõigist kliiniliselt diagnoositud rasedustest. Paljud neist katkevad juba raseduse varases staadiumis (Nagaoka et al., 2012). Trisoomiate esinemissagedust mõjutavad mitmed faktori, millest peamine on raseda vanus. Näiteks moodustavad trisoomsed rasedused alla 25-aastaste naiste kõigist rasedustest 2%, samas kui üle 40-aastaste rasedate seas on vastav näit 35% (Hassold & Hunt, 2001).

Loote sõleuuring võimaldab avastada lootel esinevaid arenguhäireid ja erinevaid kromosoomhaiguseid nagu näiteks Down'i, Edwards'i ja Patau' sündroomid. Varajane teave võimaliku loote kromosoomihaiguse kohta võimaldab tulevasi vanemaid paremini ette valmistada ja langetada informeeritud otsus raseduse jätkamise osas.

Laialdaselt on kasutusel esimese trimestri kombineeritud sõeluuring, mis sisaldab loote ultraheli- ja vereseerumi uuringuid. Kombineeritud sõeluuringu abil tuvastatakse ligikaudu 90% loote aneuploidiatest (Ridnõi et al., 2021; van Elslande et al., 2019). Lõpliku diagnoosi saamiseks suunatakse kõrge riski saanud patsient edasi invasiivsele protseduurile nagu amniotsentees ja koorionibiopsia, millele järgneb karüotüpiseerimine. Eelnimetatud sõeluuringute puuduseks on arvestatav valepositiivsete testitulemuste hulk, mistõttu üle 90% positiivse tulemuse saanud patsientidest kannab täiesti tervet loodet (Russo & Blakemore, 2014). Ebavajalik invasiivne uuring põhjustab rasedatele asjatut stressi ja sellega kaasneb ka väike protseduurist tingitud oht raseduse katkemiseks (Akolekar et al., 2015).

Viimane täiendus loote aneuploidiate sõeluuringule on mitteinvasiivne sünnieelne testimine (ingl k *non-invasive prenatal testing* ehk NIPT), mis põhineb loote päritolu rakuvaba DNA (ingl k *cell-free DNA* ehk cfDNA) analüüsil. Loote cfDNA avastati raseda vereproovist 1997. aastal (Dennis Lo et al., 1997). Loote cfDNA pärineb peamiselt platsenta trofoblastidest, kust apoptoosi teel vabanenud DNA fragmendid suunatakse raseda vereringesse ja sedakaudu utiliseerimisse. Loote cfDNA moodustab ligikaudu 10% kogu raseda cfDNA-st ja selle osakaal kasvab raseduse edenemisel (Ashoor, Syngelaki, Poon, et al., 2013). Esimese trimestri kombineeritud sõeluuringuga võrreldes on NIPT tundlikkus ja spetsiifilisus oluliselt kõrgemad. NIPT tuvastamise määr sagedasemate loote aneuploidiate korral on üle 95% ja valepositiivsete

testitulemuste hulk jääb alla 0,5% (Gil et al., 2017). Lisaks loote aneuploidiate tuvastamisele võimaldab NIPT määrata loote sugu ja ka sub-kromosomaalseid aberratsioone. NIPT laialdasemat rakendamist kliinilises praktikas takistab sekveneerimistehnoloogiast tingitud kõrgem kulu.

Uurimistöö eesmärgid

1. Töötada välja TAC-seq analüüsi jaoks sobiv sekveneerimisandmete töötluse ja analüüsi töövoog. Rakendada väljatöötatud TAC-seq analüüsitarkvara sünteetiliste trisoomia proovide analüüsimiseks ning hinnata tulemuste täpsust.
2. Töötada välja arvutuslik analüüsiraamistik ja sellega seotud statistilised meetodid suunatud kõrge katvusega NIPT sekveneerimisanalüüsi jaoks. Uurida nende meetodite bioloogilisi ja tehnilisi piiranguid ja võimekust, kasutades erinevaid *in silico* simuleeritud andmekogumeid. Lisaks rakendada neid arvutusvahendeid *in vitro* loodud sünteetiliste trisoomia proovide NIPT analüüsiks.
3. Seadistada üles, optimeerida ning juurutada madala katvusega ülegenoomsel sekveneerimisel põhinev NIPT töövoog Eesti tervishoiusüsteemis. Viia läbi valideerimisuuring kasutades rasedate patsientide kliinilisi proove.

Materjal ja metoodika

Esimeses ja teises *in vitro* uurimuses loodi eksperimentaalsed trisoomiaga rasedust imiteerivad kontrollproovid. Selleks kasutati trisoomset ja mitte-trisoomset rakuliini. Esmalt eraldati rakuliinist genoomne DNA, mis lõhuti sonikeerimise abil 150 kuni 200 bp juppideks, nagu on iseloomulik cfDNA fragmentidele. Seejärel segati kahest rakuliinidest saadud fragmendid kokku erinevates vahekordades, mis esindas laia loote fraktsiooni vahemikku. Eksperimentaalseid *in vitro* proove kasutati suunatud sekveneerimisel põhineva NIPT metoodika katsetes. Selleks disainiti referentskromosoomide 2 ja 3 ning uuritava kromosoomi 21 kvantiseerimiseks spetsiaalsed TAC-seq oligonukleotiidid. Eksperiment 1 sisaldas 16 eksperimentaalselt loodud proovi ja 114 sihtmärklookust. Eksperiment 2 sisaldas 12 eksperimentaalselt loodud proovi ja 224 sihtmärklookust. Iga proovi puhul hinnati referentskromosoomide ja uuritava kromosoomi erinevust kasutades sihtmärklookuste kvantitatiivseid väärtuseid.

Teises *in silico* uurimuses simuleeriti suunatud sekveneerimist modelleerides raseda cfDNA andmestikud. Kokku loodi 1 800 andmestikku, mis erinesid üksteisest loote kromosomaalse oleku (euploida, ema päritolu trisoomia ja isa päritolu trisoomia), loote fraktsiooni (1% kuni 20%) ja sekveneerimissügavuse (500 kuni 15 000 lugemit lookuse kohta) osas. Iga andmestik sisaldas 10 000 indiviidi ja iga indiviid sisaldas 1 000 lookust. Simuleeritud *in silico* andmes-

tikel rakendati loodud varjatud Markovi mudeleid (ingl k *hidden Markov model* ehk HMM) lookuste ja kromosoomide klassifitseerimiseks loote kromosomaalse oleku kategooriatesse. Täiendavalt rakendati klassifitseeritud lookustele ka masinõppe meetodeid nagu otsustuspuu (ingl k *decision tree* ehk DT) ja tugivektor masin (ingl k *support-vector machine* ehk SVM) kromosoomide täiustatud klassifitseerimiseks loote kromosomaalse oleku kategooriatesse.

Kolmas uurimus hõlmas Eesti ja Belgia rasedate kohorte, kus oli kokku 424 üksikrasedusega patsienti. Eesti kohort koosnes vastavalt 259 kõrge ja 149 madala riskiga raseda vereproovidest. Vereproovid võeti Tartu Ülikooli Kliinikumis ja Ida-Tallinna Keskhaiglas aastatel 2015 kuni 2018. Eesti kohort sisaldas 28 aneuploidia juhtu – 13 trisoomia 21 (T21), 7 trisoomia 18 (T18), 4 monosoomia X (45,X), 1 mosaiik T18, 1 trisoomia 13 (T13), 1 mosaiik trisoomia 8 ja 1 trisoomia X. Belgia kohort hõlmas 16 proovi UZ Leuveni inimesegeneetika keskusest (Leuven, Belgia), mille hulgas esines 8 aneuploidia juhtu – 2 T21, 3 T18 ja 3 T13. Kogutud raseda cfDNA proovid sekveneeriti madala katvusega ülegenoomselt. Saadud sekveneerimisandmete põhjal hinnati iga proovi puhul sagedasemate aneuploidiate (T21, T18, T13 ja 45,X) esinemise riski kasutades NIPTmer tarkvara.

Uurimistöö tulemused

Doktoritöö esimese uurimuse raames töötati välja uus suunatud sekveneerimisel põhinev TAC-seq meetod, mis võimaldab DNA ja RNA algmolekulide täpset loendamist. Kuigi kliinilises kasutuses on juba mitmeid analoogseid lahendusi, on TAC-seq eripäraks detektsiooni oligonukleotiididesse kavandatud unikaalsed molekulide identifikaatorid (ingl k *unique molecular identifier* ehk UMI). See võimaldab algmolekulide loendamist ka pärast amplifikatsiooni sekveneerimisraamatukogu ettevalmistamise etapis. Lisaks töötati välja TAC-seq laboratoorsele protokollile vastav analüüsi tarkvara, mis vahendab suunatud sekveneerimisandmete konverteerimist proovide ja lookuste kaupa lugemite ja molekulide hulgaks etteantud UMI lävendi alusel. TAC-seq analüüsi töövoo hindamiseks kasutati etteantud kontsentratsiooniga sünteetilisi kontrollproove, mille puhul võrreldi sisendmolekulide hulka detekteeritud molekulide arvuga. Selle tulemusel selgus, et TAC-seq on kõrge sensitiivsusega meetod algmolekulide täpseks kvantiseerimiseks laias kontsentratsioonivahemikus (1 kuni 3×10^5 molekuli sihtmärklookuse kohta). Lisaks hinnati TAC-seq meetodi tundlikkust trisoomia määramisel erinevate loote fraktsioonide korral. Nii Eksperiment 1 kui ka Eksperiment 2 tulemuste põhjal selgus, et referentskromosoomide ja uuritava kromosoomi sihtmärklookuste kvantitatiivsete väärtuste vahel esineb statistiliselt oluline erinevus alates 10% loote fraktsioonist.

Teise uurimuse käigus arendati välja arvutuslik analüüsiraamistik loote trisoomiate määramiseks kõrge katvusega suunatud sekveneerimisandmetest. Analüüsi esimeses etapis klassifitseeritakse ühel kromosoomil asuvad järjestikused lookused kromosomaalse oleku kategooriatesse kasutades HMM-i.

Klassifitseeritud lookused jaotatakse omakorda kromosoomi oleku kategooriatesse, kasutades täiendavaid masinõppe meetodeid nagu DT ja SVM, mille tulemuseks on riskihinnang iga uuritava kromosoomi kohta. Arendatud analüüsiraamistiku rakendamisel *in silico* simuleeritud andmestikul selgus, et lugemite hulgal põhinev HMM koostöös SVM mudeliga võimaldab loote trisoomia määramist alates 2% loote fraktsioonist. Samal ajal võimaldab lugemite hulgal ja alleelide suhtel põhinev mudel lisaks määrata trisoomse kromosoomi vanemlikku päritolu ja loote fraktsiooni, kuid trisoomia määramisel vajab sama loote fraktsiooni korral oluliselt kõrgemat sekveneerimise sügavust. Lisaks rakendati lugemite hulgal põhinevat HMM mudelit ka eksperimentaalsetel *in vitro* proovidel, mille tulemusena suudeti õigesti tuvastada kõik trisoomsed ja mitte-trisoomsed kromosoomid.

Kolmanda uurimuse eesmärgiks oli madala katvusega ülegenoomse NIPT meetodi valideerimine ja kasutuselevõtt Eesti tervishoiusüsteemis. Selleks implementeeriti NIPTmer ja SeqFF tarkvaradel põhinevad analüüsi töövood loote aneuploidiate määramiseks ja loote fraktsiooni hindamiseks. Lisaks viidi läbi validatsiooniuuring, mis hõlmas ühtekokku 424 rasedat, kellest 271 moodustasid testgrupi ja 149 proovi referentsgrupi. Uuringust selgus, et NIPTmer töövoo sensitiivsus on kõikide sagedamini esinevate loote aneuploidiate T21 (15/15), T18 (9/9), T13 (4/4) ja 45,X (4/4) korral 100% ja spetsiifilisus on T21, T13 ja 45,X korral 100% ning T18 korral 99,2%. Vale positiivsete osakaal oli 0,7% (2/271).

Uurimistöö kokkuvõte ja järeldused

Antud doktoritöö keskseks teemaks oli NIPT, mis põhineb ema veres leiduva loote päritolu rakuvaba DNA analüüsil. Võrreldes traditsionaalsete sõeluuringu meetoditega, pakub NIPT oluliselt kõrgemat sensitiivsust ja spetsiifilisust sagemini esinevate kromosoomihäirete suhtes. Doktoritöö tutvustab kahte erinevat NIPT lähenemist loote aneuploidsuse tuvastamiseks, mis põhinevad suunatud ja madala katvusega ülegenoomsel sekveneerimisel.

Suunatud sekveneerimismeetodi kasutamiseks töötati välja uudne meetod nimega TAC-seq, mis võimaldab nukleiinhappejärjestustel põhinevate biomarkerite täpset kvantifitseerimist. TAC-seq rakendamine raseda naise cfDNA proovile võimaldab kvantitatiivselt hinnata uuritava kromosoomi osakaalu referentskromosoomi(de) suhtes. Võrreldes ülegenoomse lähenemisviisiga, kvantifitseerib TAC-seq ainult etteantud sihtmärkpiirkonnad genoomis. Konkreetsete huvipakkuvate piirkondade sekveneerimisel võib olla kindel, et valitud piirkonnad on piisavalt kaetud, mida ei pruugi alati saavutada madala katvusega ülegenoomse sekveneerimisega. See eelis võimaldab kasutada suunatud lähenemisviise lühemate, kuid kliiniliselt oluliste, piirkondade tuvastamiseks. Lisaks on TAC-seq eeliseks ka väiksem bioloogilise sisendmaterjali vajadus ja sekveneerimislugemite hulk proovi kohta, mis võimaldab suuremal hulgal proovide üheaegset sekveneerimist ja langetab analüüsi kulu proovi kohta.

Sellest lähtuvalt võib eeldada, et suunatud sekveneerimisel põhinev lahendus võimaldab oluliselt vähendada NIPT analüüsi kogukulusid ja seeläbi laiendada NIPT kättesaadavust suuremale hulgale rasedatele. TAC-seq ja teiste suunatud meetodite piiranguks on see, et sihtida saab ainult selliseid etteantud piirkondi genoomis, millel on piisavalt unikaalne järjestus. Samuti on olemas teoreetiline ülempiir sihtmärkide hulgal, mida saab korraga ühe reaktsiooni käigus vaadelda. Sihtmärkide hulga tõustes kasvab ka detektsiooni oligonukleotiidide vahelise risthübridiseerumise oht. Praeguseks on praktikas testitud kuni mõnesaja sihtmärklookusega TAC-seq analüüsi (Wedenoja et al., 2020). Katsed sihtmärkide hulga ülempiiri osas on veel läbi viimata. Vaja oleks läbi viia edasisi uuringuid, et kontrollida, kas TAC-seq suudab eristada ka ema ja loote genotüüpe cfDNA proovist. Kliiniliselt oluliste variantide sihtimine pakuks lisaks loote aneuploidsuse tuvastamisele ka uusi võimalusi sünnieelseks sõeluuringuks. TAC-seq põhise testi rakendamiseks NIPT analüüsiks oleks samuti vaja läbi uuring kliiniliselt valideeritud raseda naiste cfDNA proovidega.

TAC-seq laboratoorsest meetodist inspireerituna töötati välja sobiv arvutuslik NIPT analüüsiraamistik. Kahekihiline analüüsiraamistik võimaldab tuvastada loote trisoomia. Lisaks võimaldavad alleelide suhtel põhinevad mudelid määrata loote fraktsiooni ja trisoomse kromosoomi vanemlikku päritolu. Lugemite hulgal põhinev HMM koos SVM mudeliga näitas parimaid tulemusi loote trisoomia tuvastamisel madala loote fraktsiooni ja sekveneerimise sügavuse vahemikus. See suudab tuvastada loote trisoomiat isegi 2% loote fraktsioonist, mis on madalam kui laialdaselt kasutatav minimaalne lävi 4%. Kuna kasutusel olevad suunatud sekveneerimisel põhinevad lähenemisviisid ei toeta alleelivariantide eristamist, siis saab lugemite arvul põhinevat mudelit rakendada ka olemasolevatel kõrge katvusega töövoogudel. Samas võimaldab lugemite arvul ja alleelide suhte andmete kombineeritud mudel loote fraktsiooni määramist, mis on oluline meetrik määramatuse hindamiseks. Enne väljatöötatud arvutusliku analüüsiraamistiku praktikas kasutamist on vaja täiendavaid uuringuid kliiniliselt valideeritud rasedate cfDNA proovidega.

Madala katvusega ülegenoomsel sekveneerimisel põhineva NIPT analüüsi jaoks loodi analüüsi töövoog, mis sisaldas meie varasemas publikatsioonis loodud NIPTmer tarkvara põhist loote aneuploidia tuvastamist ja SeqFF tarkvaral põhist loote fraktsiooni hindamist (Sauk et al., 2018). NIPTmer kasutab unikaalseid k-meeri järjestusi loote aneuploidsuse tuvastamiseks ülegenoomsetest sekveneerimisandmetest. Võrreldes joondamispõhiste meetoditega, vajab k-meeridel töötav lahendus oluliselt vähem arvutusvõimsust ja on sama jõudluse juures oluliselt kiirem. Väljatöötatud töövoogu rakendati Eesti ja Belgia kliiniliselt valideeritud rasedate cfDNA proovi andmetel. Validatsiooniuuringu tulemusel on sagedamini esinevate loote aneuploidiate (T21, T18, T13 ja 45,X) sensitiivsus 100% ja spetsiifilisus T21, T13 ja 45,X ja korral 100% ja T18 korral 99,2%. Valepositiivsete osakaal oli 0,7%. NIPTmer sensitiivsus ja spetsiifilisus loote trisoomiate puhul olid oodatust suuremad. Samas oli valepositiivsete osakaal oodatule sarnane (Gil et al., 2017). See võib olla tingitud valideerimisuuringusse kaasatud juhtumite piiratud arvust. Edasised

uuringud võiksid keskenduda madala katvusega ülegenoomsel sekveneerimisel põhineva NIPT meetodi edasisele optimiseerimisele, seda nii tehnilises kui ka analüütilises võtmes. Pidev tehnoloogiline areng ja soodsam sekveneerimise hind võimaldaks kulutõhusamat NIPT analüüsi ja selle laialdasemat kättesaadavust. Lisaks võimaldab kõrgem sekveneerimise sügavus lühemate kliiniliselt oluliste piirkondade usaldusväärsemat määramist.

Üldiselt on nii suunatud kui ka ülegenoomsel sekveneerimisel põhinevad NIPT meetodid muutnud rasedate sõeluuringu varasemast veel täpsemaks. Kui suunatud sekveneerimise suureks eeliseks on kulutõhusus, siis ülegenoomne lähenemine tuvastab valimatult kõikvõimalikke geneetilisi aberratsioone üle kogu genoomi.

11. REFERENCES

Akolekar, R., Beta, J., Picciarelli, G., Ogilvie, C., & D'Antonio, F. (2015). Procedure-related risk of miscarriage following amniocentesis and chorionic villus sampling: A systematic review and meta-analysis. *Ultrasound in Obstetrics and Gynecology*, *45*(1), 16–26. https://doi.org/10.1002/uog.14636

Alberry, M., Maddocks, D., Jones, M., Abdel Hadi, M., Abdel-Fattah, S., Avent, N., & Soothill, P. W. (2007). Free fetal DNA in maternal plasma in anembryonic pregnancies: Confirmation that the origin is the trophoblast. *Prenatal Diagnosis*, *27*(5), 415–418. https://doi.org/10.1002/pd.1700

Alldred, S. K., Takwoingi, Y., Guo, B., Pennant, M., Deeks, J. J., Neilson, J. P., & Alfirevic, Z. (2017). First trimester ultrasound tests alone or in combination with first trimester serum tests for Down's syndrome screening. *Cochrane Database of Systematic Reviews*, *2017*(3). https://doi.org/10.1002/14651858.CD012600

Antonarakis, S. E., Lyle, R., Dermitzakis, E. T., Reymond, A., & Deutsch, S. (2004). Chromosome 21 and Down syndrome: From genomics to pathophysiology. *Nature Reviews Genetics*, *5*(10), 725–738. https://doi.org/10.1038/nrg1448

Ashoor, G., Syngelaki, A., Poon, L. C. Y., Rezende, J. C., & Nicolaides, K. H. (2013). Fetal fraction in maternal plasma cell-free DNA at 11–13 weeks' gestation: Relation to maternal and fetal characteristics. *Ultrasound in Obstetrics and Gynecology*, *41*(1), 26–32. https://doi.org/10.1002/uog.12331

Ashoor, G., Syngelaki, A., Wang, E., Struble, C., Oliphant, A., Song, K., & Nicolaides, K. H. (2013). Trisomy 13 detection in the first trimester of pregnancy using a chromosome-selective cell-free DNA analysis method. *Ultrasound in Obstetrics and Gynecology*, *41*(1), 21–25. https://doi.org/10.1002/uog.12299

Asim, A., Kumar, A., Muthuswamy, S., Jain, S., & Agarwal, S. (2015). "Down syndrome: an insight of the disease." *Journal of Biomedical Science*, *22*(1). https://doi.org/10.1186/S12929-015-0138-Y

Aucamp, J., Bronkhorst, A. J., Badenhorst, C. P. S., & Pretorius, P. J. (2018). The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature. *Biological Reviews*, *93*(3), 1649–1683. https://doi.org/10.1111/BRV.12413

Ayse Kirbas, Korkut Daglar, N. D. (2016). Non-Invasive Prenatal Testing for Aneuploidy: A Review of the Literature. *Medicine Science | International Medical Journal*, 1. https://doi.org/10.5455/medscience.2016.05.8421

Bakker, M., Pajkrt, E., & Bilardo, C. M. (2014). Increased nuchal translucency with normal karyotype and anomaly scan: What next? *Best Practice and Research: Clinical Obstetrics and Gynaecology*, *28*(3), 355–366. https://doi.org/10.1016/ j.bpobgyn.2013.10.004

Baum, L. E., & Petrie, T. (1966). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *Https://Doi.Org/10.1214/Aoms/1177699147*, *37*(6), 1554–1563. https://doi.org/10.1214/AOMS/1177699147

Bayindir, B., Dehaspe, L., Brison, N., Brady, P., Ardui, S., Kammoun, M., Van Der Veken, L., Lichtenbelt, K., Van Den Bogaert, K., Van Houdt, J., Peeters, H., Van Esch, H., De Ravel, T., Legius, E., Devriendt, K., & Vermeesch, J. R. (2015). Non-invasive prenatal testing using a novel analysis pipeline to screen for all autosomal fetal aneuploidies improves pregnancy management. *European Journal of Human Genetics*, *23*(10), 1286–1293. https://doi.org/10.1038/ejhg.2014.282

Bianchi, D. W. (1999). Fetal cells in the maternal circulation: Feasibility for prenatal diagnosis. *British Journal of Haematology*, *105*(3), 574–583. https://doi.org/10.1046/j.1365-2141.1999.01383.x

Bianchi, D. W. (2004). Circulating fetal DNA: Its origin and diagnostic potential – A review. *Placenta*, *25*(SUPPL. A), S93–S101. https://doi.org/10.1016/j.placenta.2004.01.005

Bogart, M. H., Pandian, M. R., & Jones, O. W. (1987). Abnormal maternal serum chorionic gonadotropin levels in pregnancies with fetal chromosome abnormalities. *Prenatal Diagnosis*, *7*(9), 623–630. https://doi.org/10.1002/pd.1970070904

Bugge, M., Collins, A., Petersen, M. B., Fisher, J., Brandt, G., Hertz, J. M., Tranebjærg, L., De Lozier-Blanchet, C., Nicolaides, P., Brøndum-Nielsen, K., Morton, N., & Mikkelsen, M. (1998). Non-disjunction of chromosome 18. *Human Molecular Genetics*, *7*(4), 661–669. https://doi.org/10.1093/hmg/7.4.661

Cammarata-Scalisi, F., Araque, D., Ramírez, R., Guaran, L., & da Silva, G. (2019). Trisomy 13 mosaicism. *Boletin Medico Del Hospital Infantil de Mexico*, *76*(5), 246–250. https://doi.org/10.24875/BMHIM.19000003

Canick, J. A., Palomaki, G. E., Kloza, E. M., Lambert-Messerlian, G. M., & Haddow, J. E. (2013). The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. *Prenatal Diagnosis*, *33*(7), 667–674. https://doi.org/10.1002/pd.4126

Carlson, L. M., & Vora, N. L. (2017). Prenatal Diagnosis: Screening and Diagnostic Tools. *Obstetrics and Gynecology Clinics of North America*, *44*(2), 245–256. https://doi.org/10.1016/j.ogc.2017.02.004

Cereda, A., & Carey, J. C. (2012). The trisomy 18 syndrome. *Orphanet Journal of Rare Diseases*, *7*(1), 81. https://doi.org/10.1186/1750-1172-7-81

Chiang, T., Schultz, R. M., & Lampson, M. A. (2012). Meiotic Origins of Maternal Age-Related Aneuploidy1. *Biology of Reproduction*, *86*(1), 1–7. https://doi.org/10.1095/biolreprod.111.094367

Chitayat, D., Langlois, S., & Wilson, R. D. (2011). Prenatal screening for fetal aneuploidy in singleton pregnancies. *J Obstet Gynaecol Can*, *33*(7), 736–750. https://doi.org/10.1016/S1701-2163(16)34961-1

Chitty, L. S. (1998). Antenatal screening for aneuploidy. *Current Opinion in Obstetrics and Gynecology*, *10*(2), 91–96. https://doi.org/10.1097/00001703-199804000-00003

Chiu, R. W. K., Chan, K. C. A., Gao, Y., Lau, V. Y. M., Zheng, W., Leung, T. Y., Foo, C. H. F., Xie, B., Tsui, N. B. Y., Lun, F. M. F., Zee, B. C. Y., Lau, T. K., Cantor, C. R., & Lo, Y. M. D. (2008). Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proceedings of the National Academy of Sciences*, *105*(51), 20458–20463. https://doi.org/10.1073/PNAS.0810641105

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), 1–4. https://doi.org/10.1093/GIGASCIENCE/GIAB008

Dennis Lo, Y. M., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. G., & Wainscoat, J. S. (1997). Presence of fetal DNA in maternal plasma and serum. *Lancet*, *350*(9076), 485–487. https://doi.org/10.1016/S0140-6736(97)02174-0

Dennis Lo, Y. M., Zhang, J., Leung, T. N., Lau, T. K., Chang, A. M. Z., & Magnus Hjelm, N. (1999). Rapid clearance of fetal DNA from maternal plasma. *American Journal of Human Genetics*, *64*(1), 218–224. https://doi.org/10.1086/302205

Ehrich, M., Deciu, C., Zwiefelhofer, T., Tynan, J. A., Cagasan, L., Tim, R., Lu, V., McCullough, R., McCarthy, E., Nygren, A. O. H., Dean, J., Tang, L., Hutchison, D., Lu, T., Wang, H., Angkachatchai, V., Oeth, P., Cantor, C. R., Bombard, A., & van den Boom, D. (2011). Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *American Journal of Obstetrics and Gynecology*, *204*(3), 205.e1-205.e11. https://doi.org/10.1016/j.ajog.2010.12.060

Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L., & Quake, S. R. (2008). Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proceedings of the National Academy of Sciences*, *105*(42), 16266–16271. https://doi.org/10.1073/pnas.0808319105

Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L., & Quake, S. R. (2010). Analysis of the Size Distributions of Fetal and Maternal Cell-Free DNA by Paired-End Sequencing. *Clinical Chemistry*, *56*(8), 1279–1286. https://doi.org/10.1373/ CLINCHEM.2010.144188

Ferreira, J. C. P., Grati, F. R., Bajaj, K., Malvestiti, F., Grimi, M. B., Trotta, A., Liuti, R., Milani, S., Branca, L., Hartman, J., Maggi, F., Simoni, G., & Gross, S. J. (2016). Frequency of fetal karyotype abnormalities in women undergoing invasive testing in the absence of ultrasound and other high-risk indications. *Prenatal Diagnosis*, *36*(12), 1146–1155. https://doi.org/10.1002/PD.4951

Gadsbøll, K., Petersen, O. B., Gatinois, V., Strange, H., Jacobsson, B., Wapner, R., Vermeesch, J. R., & Vogel, I. (2020). Current use of noninvasive prenatal testing in Europe, Australia and the USA: A graphical presentation. *Acta Obstetricia et Gynecologica Scandinavica*, *99*(6), 722–730. https://doi.org/10.1111/aogs.13841

Gahan, P. (2013). Circulating nucleic acids: possible inherited effects. *Biological Journal of the Linnean Society*, *110*(4), 931–948. https://doi.org/10.1111/BIJ.12160

Geipel, A., Willruth, A., Vieten, J., Gembruch, U., & Berg, C. (2010). Nuchal fold thickness, nasal bone absence or hypoplasia, ductus venosus reversed flow and tricuspid valve regurgitation in screening for trisomies 21, 18 and 13 in the early second trimester. *Ultrasound in Obstetrics and Gynecology*, *35*(5), 535–539. https://doi.org/10.1002/uog.7597

Ghosh, S., Feingold, E., Chakraborty, S., & Dey, S. K. (2010). Telomere length is associated with types of chromosome 21 nondisjunction: A new insight into the maternal age effect on Down syndrome birth. *Human Genetics*, *127*(4), 403–409. https://doi.org/10.1007/s00439-009-0785-8

Gil, M. M., Accurti, V., Santacruz, B., Plana, M. N., & Nicolaides, K. H. (2017). Analysis of cell-free DNA in maternal blood in screening for aneuploidies: updated meta-analysis. *Ultrasound in Obstetrics & Gynecology*, *50*(3), 302–314. https://doi.org/10.1002/uog.17484

Goel, N., Morris, J. K., Tucker, D., Walle, H. E. K. de, Bakker, M. K., Kancherla, V., Marengo, L., Canfield, M. A., Kallen, K., Lelong, N., Camelo, J. L., Stallings, E. B., Jones, A. M., Nance, A., Huynh, M.-P., Martínez-Fernández, M.-L., Sipek, A., Pierini, A., Nembhard, W. N., … Morgan, M. (2019). Trisomy 13 and 18—Prevalence and mortality – A multi-registry population based analysis. *American Journal of Medical Genetics. Part A*, *179*(12), 2382. https://doi.org/10.1002/ AJMG.A.61365

Gregg, A. R., Skotko, B. G., Benkendorf, J. L., Monaghan, K. G., Bajaj, K., Best, R. G., Klugman, S., & Watson, M. S. (2016). Noninvasive prenatal screening for fetal aneuploidy, 2016 update: A position statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine*, *18*(10), 1056–1065. https://doi.org/-10.1038/gim.2016.97

Gunther, D. F., Eugster, E., Zagar, A. J., Bryant, C. G., Davenport, M. L., & Quigley, C. A. (2004). Ascertainment bias in turner syndrome: new insights from girls who were diagnosed incidentally in prenatal life. *Pediatrics*, *114*(3), 640–644. https://doi.org/10.1542/peds.2003-1122-L

Hannon, G. J. (2010). *FASTX-Toolkit*. http://hannonlab.cshl.edu/fastx_toolkit

Hassold, T., Hall, H., & Hunt, P. (2007). The origin of human aneuploidy: where we have been, where we are going. *Human Molecular Genetics*, *16 Spec No. 2*(R2), R203-8. https://doi.org/10.1093/hmg/ddm243

Hassold, T., & Hunt, P. (2001). To err (meiotically) is human: The genesis of human aneuploidy. *Nature Reviews Genetics*, *2*(4), 280–291. https://doi.org/10.1038/35066065

Hassold, T., Merrill, M., Adkins, K., Freeman, S., & Sherman2, S. (1995). Recombination and Maternal Age-Dependent Nondisjunction: Molecular Studies of Trisomy 16. *Am. J. Hum. Genet*, *57*(4), 867–874.

Holmes, G. (2014). Gastrointestinal disorders in Down syndrome. *Gastroenterology and Hepatology From Bed to Bench*, *7*(1), 6.

Hultén, M. A., Patel, S., Jonasson, J., & Iwarsson, E. (2010). On the origin of the maternal age effect in trisomy 21 Down syndrome: The oocyte mosaicism selection model. *Reproduction*, *139*(1), 1–9. https://doi.org/10.1530/REP-09-0088

Iwarsson, E., Jacobsson, B., Dagerhamn, J., Davidson, T., Bernabé, E., & Heibert Arnlind, M. (2017). Analysis of cell-free fetal DNA in maternal blood for detection of trisomy 21, 18 and 13 in a general pregnant population and in a high risk population – a systematic review and meta-analysis. *Acta Obstetricia et Gynecologica Scandinavica*, *96*(1), 7–18. https://doi.org/10.1111/aogs.13047

Iwarsson, E., Kvist, U., & Hultén, M. A. (2015). Disomy 21 in spermatozoa and the paternal origin of trisomy 21 Down syndrome. *Molecular Cytogenetics*, *8*(1), 67. https://doi.org/10.1186/s13039-015-0155-2

Jia, C. W., Wang, L., Lan, Y. L., Song, R., Zhou, L. Y., Yu, L., Yang, Y., Liang, Y., Li, Y., Ma, Y. M., & Wang, S. Y. (2015). Aneuploidy in early miscarriage and its related factors. *Chinese Medical Journal*, *128*(20), 2772–2776. https://doi.org/10.4103/0366-6999.167352

Juneau, K., Bogard, P. E., Huang, S., Mohseni, M., Wang, E. T., Ryvkin, P., Kingsley, C., Struble, C. A., Oliphant, A., & Zahn, J. M. (2014). Microarray-based cell-free DNA analysis improves noninvasive prenatal testing. *Fetal Diagnosis and Therapy*, *36*(4), 282–286. https://doi.org/10.1159/000367626

Kaplinski, L., Lepamets, M., & Remm, M. (2015). GenomeTester4: a toolkit for performing basic set operations - union, intersection and complement on k-mer lists. *GigaScience*, *4*(1). https://doi.org/10.1186/S13742-015-0097-Y

Kikkeri, N. S., & Nagalli, S. (2021). Turner Syndrome. *StatPearls*.

Kim, S. K., Hannum, G., Geis, J., Tynan, J., Hogg, G., Zhao, C., Jensen, T. J., Mazloom, A. R., Oeth, P., Ehrich, M., van den Boom, D., & Deciu, C. (2015). Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. *Prenatal Diagnosis*, *35*(8), 810–815. https://doi.org/10.1002/pd.4615

Kinnings, S. L., Geis, J. A., Almasri, E., Wang, H., Guan, X., Mccullough, R. M., Bombard, A. T., Saldivar, J. S., Oeth, P., & Deciu, C. (2015). Factors affecting levels of circulating cell-free fetal DNA in maternal plasma and their implications for non-invasive prenatal testing. *Prenatal Diagnosis*, *35*(8), 816–822. https://doi.org/10.1002/PD.4625

Kotsopoulou, I., Tsoplou, P., Mavrommatis, K., & Kroupis, C. (2015). Non-invasive prenatal testing (NIPT): limitations on the way to become diagnosis. *Diagnosis*, *2*(3), 141–158. https://doi.org/10.1515/dx-2015-0002

Kusre, G., Sarma, M., Nirmolia, T., & Shankarishan, P. (2015). Robertsonian Translocation t (21; 21) in a Female Born to Normal Parents: A Case Report. *Journal of Clinical and Diagnostic Research : JCDR*, *9*(1), GD01. https://doi.org/10.7860/JCDR/2015/10638.5372

Lamb, N. E., Sherman, S. L., & Hassold, T. J. (2005). Effect of meiotic recombination on the production of aneuploid gametes in humans. *Cytogenetic and Genome Research*, *111*(3–4), 250–255. https://doi.org/10.1159/000086896

Langlois, S., Yong, P. J., Yong, S. L., Barrett, I., Kalousek, D. K., Miny, P., Exeler, R., Morris, K., & Robinson, W. P. (2006). Postnatal follow-up of prenatally diagnosed trisomy 16 mosaicism. *Prenatal Diagnosis*, *26*(6), 548–558. https://doi.org/10.1002/pd.1457

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Li, Y., Zimmermann, B., Rusterholz, C., Kang, A., Holzgreve, W., & Hahn, S. (2004). Size Separation of Circulatory DNA in Maternal Plasma Permits Ready Detection of Fetal DNA Polymorphisms. *Clinical Chemistry*, *50*(6), 1002–1011. https://doi.org/10.1373/CLINCHEM.2003.029835

Liao, G. J. W., Chan, K. C. A., Jiang, P., Sun, H., Leung, T. Y., Chiu, R. W. K., & Dennis Lo, Y. M. (2012). Noninvasive prenatal diagnosis of fetal trisomy 21 by allelic ratio analysis using targeted massively parallel sequencing of maternal plasma DNA. *PLoS ONE*, *7*(5), e38154. https://doi.org/10.1371/journal.pone.0038154

Lo, Y. M. D., Tein, M. S. C., Lau, T. K., Haines, C. J., Leung, T. N., Poon, P. M. K., Wainscoat, J. S., Johnson, P. J., Chang, A. M. Z., & Hjelm, N. M. (1998). Quantitative Analysis of Fetal DNA in Maternal Plasma and Serum: Implications for Noninvasive Prenatal Diagnosis. *The American Journal of Human Genetics*, *62*(4), 768–775. https://doi.org/10.1086/301800

Loane, M., Morris, J. K., Addor, M. C., Arriola, L., Budd, J., Doray, B., Garne, E., Gatt, M., Haeusler, M., Khoshnood, B., Klungsoyr Melve, K., Latos-Bielenska, A., McDonnell, B., Mullaney, C., O'Mahony, M., Queißer-Wahrendorf, A., Rankin, J., Rissmann, A., Rounding, C., … Dolk, H. (2013). Twenty-year trends in the prevalence of Down syndrome and other trisomies in Europe: impact of maternal age and prenatal screening. *European Journal of Human Genetics*, *21*(1), 27. https://doi.org/10.1038/EJHG.2012.94

Lunding, S. A., Aksglaede, L., Anderson, R. A., Main, K. M., Juul, A., Hagen, C. P., & Pedersen, A. T. (2015). AMH as predictor of premature ovarian insufficiency: A longitudinal study of 120 turner syndrome patients. *Journal of Clinical Endocrinology and Metabolism*, *100*(7), E1030–E1038. https://doi.org/10.1210/jc.2015-1621

Macri, J. N., Spencer, K., Aitken, D., Garver, K., Buchanan, P. D., Muller, F., & Boue, A. (1993). First-trimester free beta (hCG) screening for Down syndrome. *Prenatal Diagnosis*, *13*(7), 557–562. https://doi.org/10.1002/pd.1970130704

Mandel, P., & Métais, P. (1948). Les acides nucléiques du plasma sanguin chez l'homme. *Comptes Rendus Des Séances de La Société de Biologie et de Ses Filiales*, *142*(3–4), 241–243.

Miller, C. A., Hampton, O., Coarfa, C., & Milosavljevic, A. (2011). ReadDepth: A parallel R package for detecting copy number alterations from short sequencing reads. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0016327

Miryounesi, M., Diantpour, M., Motevaseli, E., & Ghafouri-Fard, S. (2016). Homozygosity for a Robertsonian Translocation (13q;14q) in a Phenotypically Normal 44, XX Female with a History of Recurrent Abortion and a Normal Pregnancy Outcome. *Journal of Reproduction & Infertility*, *17*(3), 184.

Muller, F., Rebiffé, M., Taillandier, A., Oury, J. F., & Mornet, E. (2000). Parental origin of the extra chromosome in prenatally diagnosed fetal trisomy 21. *Human Genetics*, *106*(3), 340–344. https://doi.org/10.1007/s004390051047

Nagaoka, S. I., Hassold, T. J., & Hunt, P. A. (2012). Human aneuploidy: Mechanisms and new insights into an age-old problem. *Nature Reviews Genetics*, *13*(7), 493–504. https://doi.org/10.1038/nrg3245

Norton, M. E., Brar, H., Weiss, J., Karimi, A., Laurent, L. C., Caughey, A. B., Rodriguez, M. H., Williams 3rd, J., Mitchell, M. E., Adair, C. D., Lee, H., Jacobsson, B., Tomlinson, M. W., Oepkes, D., Hollemon, D., Sparks, A. B., Oliphant, A., Song, K., Williams, J., … Song, K. (2012). Non-Invasive Chromosomal Evaluation (NICE) Study: results of a multicenter prospective cohort study for detection of fetal trisomy 21 and trisomy 18. *American Journal of Obstetrics and Gynecology*, *207*(2), 137 e1-8. https://doi.org/10.1016/j.ajog.2012.05.021

Norton, M. E., Jacobsson, B., Swamy, G. K., Laurent, L. C., Ranzini, A. C., Brar, H., Tomlinson, M. W., Pereira, L., Spitz, J. L., Hollemon, D., Cuckle, H., Musci, T. J., & Wapner, R. J. (2015). Cell-free DNA Analysis for Noninvasive Examination of Trisomy. *New England Journal of Medicine*, *372*(17), 1589–1597. https://doi.org/10.1056/NEJMOA1407349/SUPPL_FILE/NEJMOA1407349_DISCLOSURES.PDF

Norwitz, E. R., & Levy, B. (2013). Noninvasive Prenatal Testing: The Future Is Now. *Reviews in Obsetrics and Gynecology*, *6*(2), 48–62. https://doi.org/10.3909/riog0201

Pajkrt, E., Bilardo, caterina M., Van Lith, J. M., Mol, B. W. J., & Bleker, O. P. (1995). Nuchal translucency measurement in normal fetuses. *Obstetrics and Gynecology*, *86*(6), 994–997. https://doi.org/10.1016/0029-7844(95)00310-N

Palomaki, G. E., Kloza, E. M., Lambert-Messerlian, G. M., Haddow, J. E., Neveux, L. M., Ehrich, M., van den Boom, D., Bombard, A. T., Deciu, C., Grody, W. W., Nelson, S. F., & Canick, J. a. (2011). DNA sequencing of maternal plasma to detect Down syndrome: An international clinical validation study. *Genetics in Med.*, *XX*(Xx), 1–8. https://doi.org/10.1097/GIM.0b013e3182368a0e

Paluoja, P., Teder, H., Ardeshirdavani, A., Bayindir, B., Vermeesch, J., Salumets, A., Krjutškov, K., & Palta, P. (2021). Systematic evaluation of NIPT aneuploidy detection software tools with clinically validated NIPT samples. *PLOS Computational Biology*, *17*(12), e1009684. https://doi.org/10.1371/JOURNAL.PCBI.1009684

Papageorgiou, E. A., Fiegler, H., Rakyan, V., Beck, S., Hulten, M., Lamnissou, K., Carter, N. P., & Patsalis, P. C. (2009). Sites of Differential DNA Methylation between Placenta and Peripheral Blood: Molecular Markers for Noninvasive

Prenatal Diagnosis of Aneuploidies. *The American Journal of Pathology*, *174*(5), 1609–1618. https://doi.org/10.2353/AJPATH.2009.081038

Papavassiliou, P., Charalsawadi, C., Rafferty, K., & Jackson-Cook, C. (2015). Mosaicism for trisomy 21: A review. *American Journal of Medical Genetics Part A*, *167*(1), 26–39. https://doi.org/10.1002/AJMG.A.36861

Pellestor, F., Anahory, T., & Hamamah, S. (2005). The chromosomal analysis of human oocytes. An overview of established procedures. *Human Reproduction Update*, *11*(1), 15–32. https://doi.org/10.1093/humupd/dmh051

Ramasamy, R., Scovell, J. M., Kovac, J. R., Cook, P. J., Lamb, D. J., & Lipshultz, L. I. (2015). Fluorescence in situ hybridization detects increased sperm aneuploidy in men with recurrent pregnancy loss. *Fertility and Sterility*, *103*(4), 906-909.e1. https://doi.org/10.1016/j.fertnstert.2015.01.029

Rasmussen, S. A., Wong, L. Y. C., Yang, Q., May, K. M., & Friedman, J. M. (2003). Population-Based Analyses of Mortality in Trisomy 13 and Trisomy 18. *Pediatrics*, *111*(4), 777–784. https://doi.org/10.1542/PEDS.111.4.777

Renna, M. D., Pisani, P., Conversano, F., Perrone, E., Casciaro, E., Renzo, G. C. di, Paola, M. di, Perrone, A., & Casciaro, S. (2013). Sonographic markers for early diagnosis of fetal malformations. *World Journal of Radiology*, *5*(10), 356. https://doi.org/10.4329/wjr.v5.i10.356

Ridnõi, K., Muru, K., Ustav, E.-L., Tammur, P., Mölter-Väär, T., Asser, K., Szirko, F., Reimand, T., & Õunap, K. (2021). Screening for fetal chromosomal disease in Estonia. Results of a two-year prospective study (available in Estonian). *Eesti Arst*.

Roper, R. J., & Reeves, R. H. (2006). Understanding the Basis for Down Syndrome Phenotypes. *PLoS Genetics*, *2*(3), 0231–0236. https://doi.org/10.1371/JOURNAL.PGEN.0020050

Russo, M. L., & Blakemore, K. J. (2014). A historical and practical review of first trimester aneuploidy screening. In *Seminars in Fetal and Neonatal Medicine* (Vol. 19, Issue 3, pp. 183–187). https://doi.org/10.1016/j.siny.2013.11.013

Saenger, P. (2001). Recommendations for the diagnosis and management of Turner Syndrome (TS). *Revista Argentina de Endocrinologia y Metabolismo*, *38*(SUPPL.), 54–55. https://doi.org/10.1210/jcem.86.7.7683

Samango-Sprouse, C., Banjevic, M., Ryan, A., Sigurjonsson, S., Zimmermann, B., Hill, M., Hall, M. P., Westemeyer, M., Saucier, J., Demko, Z., & Rabinowitz, M. (2013). SNP-based non-invasive prenatal testing detects sex chromosome aneuploidies with high accuracy. *Prenatal Diagnosis*, *33*(7), 643–649. https://doi.org/10.1002/pd.4159

Satgé, D., Nishi, M., Sirvent, N., Vekemans, M., Chenard, M. P., & Barnes, A. (2017). A tumor profile in Patau syndrome (trisomy 13). *American Journal of Medical Genetics, Part A*, *173*(8), 2088–2096. https://doi.org/10.1002/ajmg.a.38294

Sauk, M., Žilina, O., Kurg, A., Ustav, E.-L., Peters, M., Paluoja, P., Roost, A. M., Teder, H., Palta, P., Brison, N., Vermeesch, J. R., Krjutškov, K., Salumets, A., & Kaplinski, L. (2018). NIPTmer: rapid k-mer-based software package for detection of fetal aneuploidies. *Scientific Reports*, *8*(1), 5616. https://doi.org/10.1038/s41598-018-23589-8

Sävendahl, L., & Davenport, M. L. (2000). Delayed diagnoses of Turner's syndrome: Proposed guidelines for change. *Journal of Pediatrics*, *137*(4), 455–459. https://doi.org/10.1067/mpd.2000.107390

Sersinger, N. A., Zakher, A., Huber, U., Pescia, G., & Schneider, H. (1995). A sensitive enzyme immunoassay for pregnancy-associated plasma protein A (PAPP-A): a possible first trimester method of screening for down syndrome and other trisomies.

Archives of Gynecology and Obstetrics, *256*(4), 185–192. https://doi.org/10.1007/BF00634490

Sherman, S. L., Lamb, N. E., & Feingold, E. (2006). Relationship of recombination patterns and maternal age among non-disjoined chromosomes 21. *Biochemical Society Transactions*, *34*(4), 578–580. https://doi.org/10.1042/BST0340578

Shiefa, S., Amargandhi, M., Bhupendra, J., Moulali, S., & Kristine, T. (2013). First Trimester Maternal Serum Screening Using Biochemical Markers PAPP-A and Free β-hCG for Down Syndrome, Patau Syndrome and Edward Syndrome. *Indian Journal of Clinical Biochemistry*, *28*(1), 3. https://doi.org/10.1007/S12291-012-0269-9

Sitska, M. (2018). Pärilike haiguste sünnieelne diagnoosimine Eestis aastatel 1990–2016. *Eesti Arst*.

Sparks, A. B., Struble, C. A., Wang, E. T., Song, K., & Oliphant, A. (2012). Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: Evaluation for trisomy 21 and trisomy 18. *American Journal of Obstetrics and Gynecology*, *206*(4), 319.e1-319.e9. https://doi.org/10.1016/j.ajog.2012.01.030

Sparks, A. B., Wang, E. T., Struble, C. A., Barrett, W., Stokowski, R., Mcbride, C., Zahn, J., Lee, K., Shen, N., Doshi, J., Sun, M., Garrison, J., Sandler, J., Hollemon, D., Pattee, P., Tomita-Mitchell, A., Mitchell, M., Stuelpnagel, J., Song, K., & Oliphant, A. (2012). Selective analysis of cell-free DNA in maternal blood for evaluation of fetal trisomy. *Prenatal Diagnosis*, *32*(1), 3–9. https://doi.org/10.1002/pd.2922

Spencer, K., Souter, V., Tul, N., Snijders, R., & Nicolaides, K. H. (1999). A screening program for trisomy 21 at 10-14 weeks using fetal nuchal translucency, maternal serum free beta-human chorionic gonadotropin and pregnancy-associated plasma protein-A. *Ultrasound in Obstetrics & Gynecology : The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, *13*(4), 231–237. https://doi.org/10.1046/J.1469-0705.1999.13040231.X

Springett, A., Wellesley, D., Greenlees, R., Loane, M., Addor, M. C., Arriola, L., Bergman, J., Cavero-Carbonell, C., Csaky-Szunyogh, M., Draper, E. S., Garne, E., Gatt, M., Haeusler, M., Khoshnood, B., Klungsoyr, K., Lynch, C., Dias, C. M., Mcdonnell, R., Nelen, V., … Morris, J. K. (2015). Congenital anomalies associated with trisomy 18 or trisomy 13: A registry-based study in 16 european countries, 2000-2011. *American Journal of Medical Genetics, Part A*, *167*(12), 3062–3069. https://doi.org/10.1002/ajmg.a.37355

Stochholm, K., Juul, S., Juel, K., Naeraa, R. W., & Gravholt, C. H. (2006). Prevalence, incidence, diagnostic delay, and mortality in Turner syndrome. *The Journal of Clinical Endocrinology and Metabolism*, *91*(10), 3897–3902. https://doi.org/10.1210/JC.2006-0558

Stokowski, R., Wang, E., White, K., Batey, A., Jacobsson, B., Brar, H., Balanarasimha, M., Hollemon, D., Sparks, A., Nicolaides, K., & Musci, T. J. (2015). Clinical performance of non-invasive prenatal testing (NIPT) using targeted cell-free DNA analysis in maternal plasma with microarrays or next generation sequencing (NGS) is consistent across multiple controlled clinical studies. *Prenatal Diagnosis*, *35*(12), 1243–1246. https://doi.org/10.1002/pd.4686

Surerus, E., Huggon, I. C., & Allan, L. D. (2003). Turner's syndrome in fetal life. *Ultrasound in Obstetrics & Gynecology*, *22*(3), 264–267. https://doi.org/10.1002/UOG.151

Taylor-Phillips, S., Freeman, K., Geppert, J., Agbebiyi, A., Uthman, O. A., Madan, J., Clarke, A., Quenby, S., & Clarke, A. (2016). Accuracy of non-invasive prenatal testing using cell-free DNA for detection of Down, Edwards and Patau syndromes: a systematic review and meta-analysis. *BMJ Open*, *6*(1), e010002. https://doi.org/10.1136/bmjopen-2015-010002

Teder, H. (2018). *cfDNA simulation*. https://github.com/seqinfo/cfdna-simulation

Teder, H., Koel, M., Paluoja, P., Jatsenko, T., Rekker, K., Laisk-Podar, T., Kukuškina, V., Velthut-Meikas, A., Fjodorova, O., Peters, M., Kere, J., Salumets, A., Palta, P., & Krjutškov, K. (2018). TAC-seq: targeted DNA and RNA sequencing for precise biomarker molecule counting. *Npj Genomic Medicine*, *3*(1), 34. https://doi.org/10.1038/s41525-018-0072-5

Teder, H., & Paluoja, P. (2018). *TAC-seq data analysis*. https://github.com/seqinfo/tac-seq-data-analysis

Teder, H., Paluoja, P., Rekker, K., Salumets, A., Krjutškov, K., & Palta, P. (2019). Computational framework for targeted high-coverage sequencing based NIPT. *PLOS ONE*, *14*(7), e0209139. https://doi.org/10.1371/journal.pone.0209139

Templado, C., Uroz, L., & Estop, A. (2013). New insights on the origin and relevance of aneuploidy in human spermatozoa. *Molecular Human Reproduction*, *19*(10), 634–643. https://doi.org/10.1093/molehr/gat039

Uroz, L., & Templado, C. (2012). Meiotic non-disjunction mechanisms in human fertile males. *Human Reproduction*, *27*(5), 1518–1524. https://doi.org/10.1093/humrep/des051

Ustav, E.-L., Haldre, K., Asser, K., Muru, K., Ridnõi, K., Buts, K., Rajasalu, L., & Reimand, T. (2016). *Prenatal diagnostics guideline: screening and diagnostics for chromosomal disease. Fetal ultrasound examination (available in Estonian)*.

van Elslande, J., Brison, N., Vermeesch, J. R., Devriendt, K., van den Bogaert, K., Legius, E., van Ranst, M., Vermeersch, P., & Billen, J. (2019). The sudden death of the combined first trimester aneuploidy screening, a single centre experience in Belgium. *Clinical Chemistry and Laboratory Medicine*, *57*(11), e294–e297. https://doi.org/10.1515/CCLM-2019-0231/PDF

Wang, E., Batey, A., Struble, C., Musci, T., Song, K., & Oliphant, A. (2013). Gestational age and maternal weight effects on fetal cell-free DNA in maternal plasma. *Prenatal Diagnosis*, *33*(7), 662–666. https://doi.org/10.1002/pd.4119

Wartosch, L., Schindler, K., Schuh, M., Gruhn, J. R., Hoffmann, E. R., McCoy, R. C., & Xing, J. (2021). Origins and mechanisms leading to aneuploidy in human eggs. *Prenatal Diagnosis*, *41*(5), 620–630. https://doi.org/10.1002/PD.5927

Watson, W. J., Miller, R. C., Wax, J. R., Hansen, W. F., Yamamura, Y., & Polzin, W. J. (2007). Sonographic detection of trisomy 13 in the first and second trimesters of pregnancy. *Journal of Ultrasound in Medicine*, *26*(9), 1209–1214. https://doi.org/10.7863/jum.2007.26.9.1209

Wedenoja, S., Yoshihara, M., Teder, H., Sariola, H., Gissler, M., Katayama, S., Wedenoja, J., Häkkinen, I. M., Ezer, S., Linder, N., Lundin, J., Skoog, T., Sahlin, E., Iwarsson, E., Pettersson, K., Kajantie, E., Mokkonen, M., Heinonen, S., Laivuori, H., … Kere, J. (2020). Fetal HLA-G mediated immune tolerance and interferon response in preeclampsia. *EBioMedicine*, *59*. https://doi.org/10.1016/j.ebiom.2020.102872

Weijerman, M. E., & De Winter, J. P. (2010). Clinical practice: The care of children with Down syndrome. *European Journal of Pediatrics*, *169*(12), 1445–1452. https://doi.org/10.1007/s00431-010-1253-0

Welker, N. C., Lee, A. K., Kjolby, R. A. S., Wan, H. Y., Theilmann, M. R., Jeon, D., Goldberg, J. D., Haas, K. R., Muzzey, D., & Chu, C. S. (2020). High-throughput fetal fraction amplification increases analytical performance of noninvasive prenatal screening. *Genetics in Medicine 2020 23:3*, *23*(3), 443–450. https://doi.org/10.1038/s41436-020-01009-5

Wolff, D. J., Van Dyke, D. L., & Powell, C. M. (2010). Laboratory guideline for turner syndrome. *Genetics in Medicine*, *12*(1), 52–55. https://doi.org/10.1097/GIM.0b013e3181c684b2

Wright, C. F., & Chitty, L. S. (2009). Science, medicine, and the future. Cell-free fetal DNA and RNA in maternal blood: Implications for safer antenatal testing. *BMJ (Online)*, *339*(7713), 161–164. https://doi.org/10.1136/bmj.b2451

Yoon, B.-J. (2009). Hidden Markov Models and their Applications in Biological Sequence Analysis. *Current Genomics*, *10*(6), 402–415. https://doi.org/10.2174/138920209789177575

Zimmermann, B., Hill, M., Gemelos, G., Demko, Z., Banjevic, M., Baner, J., Ryan, A., Sigurjonsson, S., Chopra, N., Dodd, M., Levy, B., & Rabinowitz, M. (2012). Non-invasive prenatal aneuploidy testing of chromosomes 13, 18, 21, X, and Y, using targeted sequencing of polymorphic loci. *Prenatal Diagnosis*, *32*(13), 1233–1241. https://doi.org/10.1002/pd.3993

12. ACKNOWLEDGMENTS

The research presented in this thesis was performed at the Competence Centre on Health Technologies and Institute of Biomedicine and Translational Medicine, University of Tartu. I would like to express my most sincere gratitude to the people who supported me throughout my doctoral studies:

- My supervisors, Priit Palta, Kaarel Krjutškov, and Andres Salumets, who have guided and encouraged me along my academic path. Especially Priit Palta, who has been my intellectual mentor for years.
- My wonderful colleagues from Competence Centre on Health Technologies for their help and extensive scientific knowledge.
- My referees Dr. Kai Muru and Dr. Sander Pajusalu for their valuable feedback.
- My opponent Dr. Erik Sistermans for his time to provide scientific and intellectual criticism.
- All the co-authors of my scientific publications and collaborators for their valuable contributions.
- My dear family and friends for being always there for me. Especially Epp Kaleviste, who motivated me throughout the journey.

13. PUBLICATIONS

14. CURRICULUM VITAE

Name: Hindrek Teder
Born: January 19th, 1989, Rapla, Estonia
Citizenship: Estonia
Address: Competence Centre on Health Technologies, Teaduspargi 13, 50411, Tartu, Estonia
Telephone: +372 733 0401
Email: hindrek.teder@gmail.com

Education:

2005–2008 Rapla Co-Educational Gymnasium
2008–2013 University of Tartu, Genetic Engineering, BSc
2013–2016 University of Tartu, Genetic Engineering (Bioinformatics), MSc
2016–2022 University of Tartu, Medicine, PhD

Institution and position:

2014–2015 Competence Centre on Health Technologies, Laboratory Specialist
2015−2016 Competence Centre on Health Technologies, Bioinformatics Specialist
2016−… Competence Centre on Health Technologies, Bioinformatics Researcher

R&D related managerial and administrative work:

2013−2016 Bioscience Students' Association

Publications:

1. Krjutškov, K., Katayama, S., Saare, M., Vera-Rodriguez, M., Lubenets, D., Samuel, K., Laisk-Podar, T., **Teder, H.**, Einarsdottir, E., Salumets, A., & Kere, J. (2016). Single-cell transcriptome analysis of endometrial tissue. *Human Reproduction*, *31*(4), 844–853. https://doi.org/10.1093/humrep/dew008
2. Sõritsa, D., **Teder, H.**, Roosipuu, R., Tamm, H., Laisk-Podar, T., Soplepmann, P., Altraja, A., Salumets, A., & Peters, M. (2018). Whole exome sequencing of benign pulmonary metastasizing leiomyoma reveals mutation in the BMP8B gene. *BMC Medical Genetics*, *19*(1), 20. https://doi.org/10.1186/s12881-018-0537-5
3. Sauk, M., Žilina, O., Kurg, A., Ustav, E.-L., Peters, M., Paluoja, P., Roost, A. M., **Teder, H.**, Palta, P., Brison, N., Vermeesch, J. R., Krjutškov, K., Salumets, A., & Kaplinski, L. (2018). NIPTmer: rapid k-mer-based software package for detection of fetal aneuploidies. *Scientific Reports*, *8*(1), 5616. https://doi.org/10.1038/s41598-018-23589-8

4. **Teder, H.**, Koel, M., Paluoja, P., Jatsenko, T., Rekker, K., Laisk-Podar, T., Kukuškina, V., Velthut-Meikas, A., Fjodorova, O., Peters, M., Kere, J., Salumets, A., Palta, P., & Krjutškov, K. (2018). TAC-seq: targeted DNA and RNA sequencing for precise biomarker molecule counting. *Npj Genomic Medicine*, *3*(1), 34. https://doi.org/10.1038/s41525-018-0072-5
5. Saare, M., Laisk, T., **Teder, H.**, Paluoja, P., Palta, P., Koel, M., Kirss, F., Karro, H., Sõritsa, D., Salumets, A., Krjutškov, K., & Peters, M. (2019). A molecular tool for menstrual cycle phase dating of endometrial samples in endometriosis transcriptome studies†. *Biology of Reproduction*, *0*(April), 1–3. https://doi.org/10.1093/biolre/ioz072
6. Žilina, O., Rekker, K., Kaplinski, L., Sauk, M., Paluoja, P., **Teder, H.**, Ustav, E., Tõnisson, N., Reimand, T., Ridnõi, K., Palta, P., Vermeesch, J. R., Krjutškov, K., Kurg, A., & Salumets, A. (2019). Creating basis for introducing non-invasive prenatal testing in the Estonian public health setting. *Prenatal Diagnosis*, *39*(13), 1262–1268. https://doi.org/10.1002/pd.5578
7. **Teder, H.**, Paluoja, P., Rekker, K., Salumets, A., Krjutškov, K., & Palta, P. (2019). Computational framework for targeted high-coverage sequencing based NIPT. *PLOS ONE*, *14*(7), e0209139. https://doi.org/10.1371/journal.pone.0209139
8. Wedenoja, S., Yoshihara, M., **Teder, H.**, Sariola, H., Gissler, M., Katayama, S., Wedenoja, J., Häkkinen, I. M., Ezer, S., Linder, N., Lundin, J., Skoog, T., Sahlin, E., Iwarsson, E., Pettersson, K., Kajantie, E., Mokkonen, M., Heinonen, S., Laivuori, H., Krjutkov, K., Kere, J. (2020). Fetal HLA-G mediated immune tolerance and interferon response in preeclampsia. *EBioMedicine*, *59*. https://doi.org/10.1016/j.ebiom.2020.102872
9. Paluoja, P., **Teder, H.**, Ardeshirdavani, A., Bayindir, B., Vermeesch, J., Salumets, A., Krjutškov, K., & Palta, P. (2021). Systematic evaluation of NIPT aneuploidy detection software tools with clinically validated NIPT samples. *PLOS Computational Biology*, *17*(12), e1009684. https://doi.org/10.1371/JOURNAL.PCBI.1009684

15. ELULOOKIRJELDUS

Nimi: Hindrek Teder
Sündinud: 19. jaanuar 1989, Rapla, Eesti
Kodakondsus: Eesti
Aadress: Tervisetehnoloogiate Arenduskeskus AS,
Teaduspargi 13, 50411, Tartu, Eesti
Telefon: +372 733 0401
E-mail: hindrek.teder@gmail.com

Hariduskäik:
2005–2008 Rapla Ühisgümnaasium
2008–2013 Tartu Ülikool, geenitehnoloogia, BSc
2013–2016 Tartu Ülikool, geenitehnoloogia (bioinformaatika), MSc
2016–2022 Tartu Ülikool, arstiteadus, PhD

Teenistuskäik:
2014–2015 laborispetsialist, Tervisetehnoloogiate Arenduskeskus AS
2015−2016 bioinformaatika spetsialist, Tervisetehnoloogiate Arenduskeskus AS
2016−… bioinformaatika teadur, Tervisetehnoloogiate Arenduskeskus AS

Teadusorganisatsiooniline ja -administratiivne tegevus:
2013–2016 Bioteaduste Üliõpilaste Selts

Publikatsioonid:

1. Krjutškov, K., Katayama, S., Saare, M., Vera-Rodriguez, M., Lubenets, D., Samuel, K., Laisk-Podar, T., **Teder, H.**, Einarsdottir, E., Salumets, A., & Kere, J. (2016). Single-cell transcriptome analysis of endometrial tissue. *Human Reproduction*, *31*(4), 844–853. https://doi.org/10.1093/humrep/dew008
2. Sõritsa, D., **Teder, H.**, Roosipuu, R., Tamm, H., Laisk-Podar, T., Soplepmann, P., Altraja, A., Salumets, A., & Peters, M. (2018). Whole exome sequencing of benign pulmonary metastasizing leiomyoma reveals mutation in the BMP8B gene. *BMC Medical Genetics*, *19*(1), 20. https://doi.org/10.1186/s12881-018-0537-5
3. Sauk, M., Žilina, O., Kurg, A., Ustav, E.-L., Peters, M., Paluoja, P., Roost, A. M., **Teder, H.**, Palta, P., Brison, N., Vermeesch, J. R., Krjutškov, K., Salumets, A., & Kaplinski, L. (2018). NIPTmer: rapid k-mer-based software package for detection of fetal aneuploidies. *Scientific Reports*, *8*(1), 5616. https://doi.org/10.1038/s41598-018-23589-8
4. **Teder, H.**, Koel, M., Paluoja, P., Jatsenko, T., Rekker, K., Laisk-Podar, T., Kukuškina, V., Velthut-Meikas, A., Fjodorova, O., Peters, M., Kere, J., Salumets, A., Palta, P., & Krjutškov, K. (2018). TAC-seq: targeted DNA and

RNA sequencing for precise biomarker molecule counting. *Npj Genomic Medicine*, *3*(1), 34. https://doi.org/10.1038/s41525-018-0072-5

5. Saare, M., Laisk, T., **Teder, H.**, Paluoja, P., Palta, P., Koel, M., Kirss, F., Karro, H., Sõritsa, D., Salumets, A., Krjutškov, K., & Peters, M. (2019). A molecular tool for menstrual cycle phase dating of endometrial samples in endometriosis transcriptome studies†. *Biology of Reproduction*, *0*(April), 1–3. https://doi.org/10.1093/biolre/ioz072
6. Žilina, O., Rekker, K., Kaplinski, L., Sauk, M., Paluoja, P., **Teder, H.**, Ustav, E., Tõnisson, N., Reimand, T., Ridnõi, K., Palta, P., Vermeesch, J. R., Krjutškov, K., Kurg, A., & Salumets, A. (2019). Creating basis for introducing non-invasive prenatal testing in the Estonian public health setting. *Prenatal Diagnosis*, *39*(13), 1262–1268. https://doi.org/10.1002/pd.5578
7. **Teder, H.**, Paluoja, P., Rekker, K., Salumets, A., Krjutškov, K., & Palta, P. (2019). Computational framework for targeted high-coverage sequencing based NIPT. *PLOS ONE*, *14*(7), e0209139. https://doi.org/10.1371/journal.pone.0209139
8. Wedenoja, S., Yoshihara, M., **Teder, H.**, Sariola, H., Gissler, M., Katayama, S., Wedenoja, J., Häkkinen, I. M., Ezer, S., Linder, N., Lundin, J., Skoog, T., Sahlin, E., Iwarsson, E., Pettersson, K., Kajantie, E., Mokkonen, M., Heinonen, S., Laivuori, H., Krjutkov, K., Kere, J. (2020). Fetal HLA-G mediated immune tolerance and interferon response in preeclampsia. *EBioMedicine*, *59*. https://doi.org/10.1016/j.ebiom.2020.102872
9. Paluoja, P., **Teder, H.**, Ardeshirdavani, A., Bayindir, B., Vermeesch, J., Salumets, A., Krjutškov, K., & Palta, P. (2021). Systematic evaluation of NIPT aneuploidy detection software tools with clinically validated NIPT samples. *PLOS Computational Biology*, *17*(12), e1009684. https://doi.org/10.1371/JOURNAL.PCBI.1009684

DISSERTATIONES MEDICINAE UNIVERSITATIS TARTUENSIS

1. **Heidi-Ingrid Maaroos**. The natural course of gastric ulcer in connection with chronic gastritis and *Helicobacter pylori*. Tartu, 1991.
2. **Mihkel Zilmer**. Na-pump in normal and tumorous brain tissues: Structural, functional and tumorigenesis aspects. Tartu, 1991.
3. **Eero Vasar**. Role of cholecystokinin receptors in the regulation of behaviour and in the action of haloperidol and diazepam. Tartu, 1992.
4. **Tiina Talvik**. Hypoxic-ischaemic brain damage in neonates (clinical, biochemical and brain computed tomographical investigation). Tartu, 1992.
5. **Ants Peetsalu**. Vagotomy in duodenal ulcer disease: A study of gastric acidity, serum pepsinogen I, gastric mucosal histology and *Helicobacter pylori*. Tartu, 1992.
6. **Marika Mikelsaar**. Evaluation of the gastrointestinal microbial ecosystem in health and disease. Tartu, 1992.
7. **Hele Everaus**. Immuno-hormonal interactions in chronic lymphocytic leukaemia and multiple myeloma. Tartu, 1993.
8. **Ruth Mikelsaar**. Etiological factors of diseases in genetically consulted children and newborn screening: dissertation for the commencement of the degree of doctor of medical sciences. Tartu, 1993.
9. **Agu Tamm**. On metabolic action of intestinal microflora: clinical aspects. Tartu, 1993.
10. **Katrin Gross**. Multiple sclerosis in South-Estonia (epidemiological and computed tomographical investigations). Tartu, 1993.
11. **Oivi Uibo**. Childhood coeliac disease in Estonia: occurrence, screening, diagnosis and clinical characterization. Tartu, 1994.
12. **Viiu Tuulik**. The functional disorders of central nervous system of chemistry workers. Tartu, 1994.
13. **Margus Viigimaa**. Primary haemostasis, antiaggregative and anticoagulant treatment of acute myocardial infarction. Tartu, 1994.
14. **Rein Kolk**. Atrial versus ventricular pacing in patients with sick sinus syndrome. Tartu, 1994.
15. **Toomas Podar**. Incidence of childhood onset type 1 diabetes mellitus in Estonia. Tartu, 1994.
16. **Kiira Subi**. The laboratory surveillance of the acute respiratory viral infections in Estonia. Tartu, 1995.
17. **Irja Lutsar**. Infections of the central nervous system in children (epidemiologic, diagnostic and therapeutic aspects, long term outcome). Tartu, 1995.
18. **Aavo Lang**. The role of dopamine, 5-hydroxytryptamine, sigma and NMDA receptors in the action of antipsychotic drugs. Tartu, 1995.
19. **Andrus Arak**. Factors influencing the survival of patients after radical surgery for gastric cancer. Tartu, 1996.

20. **Tõnis Karki**. Quantitative composition of the human lactoflora and method for its examination. Tartu, 1996.
21. **Reet Mändar**. Vaginal microflora during pregnancy and its transmission to newborn. Tartu, 1996.
22. **Triin Remmel**. Primary biliary cirrhosis in Estonia: epidemiology, clinical characterization and prognostication of the course of the disease. Tartu, 1996.
23. **Toomas Kivastik**. Mechanisms of drug addiction: focus on positive reinforcing properties of morphine. Tartu, 1996.
24. **Paavo Pokk**. Stress due to sleep deprivation: focus on $GABA_A$ receptor-chloride ionophore complex. Tartu, 1996.
25. **Kristina Allikmets**. Renin system activity in essential hypertension. Associations with atherothrombogenic cardiovascular risk factors and with the efficacy of calcium antagonist treatment. Tartu, 1996.
26. **Triin Parik**. Oxidative stress in essential hypertension: Associations with metabolic disturbances and the effects of calcium antagonist treatment. Tartu, 1996.
27. **Svetlana Päi**. Factors promoting heterogeneity of the course of rheumatoid arthritis. Tartu, 1997.
28. **Maarike Sallo**. Studies on habitual physical activity and aerobic fitness in 4 to 10 years old children. Tartu, 1997.
29. **Paul Naaber**. *Clostridium difficile* infection and intestinal microbial ecology. Tartu, 1997.
30. **Rein Pähkla**. Studies in pinoline pharmacology. Tartu, 1997.
31. **Andrus Juhan Voitk**. Outpatient laparoscopic cholecystectomy. Tartu, 1997.
32. **Joel Starkopf**. Oxidative stress and ischaemia-reperfusion of the heart. Tartu, 1997.
33. **Janika Kõrv**. Incidence, case-fatality and outcome of stroke. Tartu, 1998.
34. **Ülla Linnamägi**. Changes in local cerebral blood flow and lipid peroxidation following lead exposure in experiment. Tartu, 1998.
35. **Ave Minajeva**. Sarcoplasmic reticulum function: comparison of atrial and ventricular myocardium. Tartu, 1998.
36. **Oleg Milenin**. Reconstruction of cervical part of esophagus by revascularised ileal autografts in dogs. A new complex multistage method. Tartu, 1998.
37. **Sergei Pakriev**. Prevalence of depression, harmful use of alcohol and alcohol dependence among rural population in Udmurtia. Tartu, 1998.
38. **Allen Kaasik**. Thyroid hormone control over β-adrenergic signalling system in rat atria. Tartu, 1998.
39. **Vallo Matto**. Pharmacological studies on anxiogenic and antiaggressive properties of antidepressants. Tartu, 1998.
40. **Maire Vasar**. Allergic diseases and bronchial hyperreactivity in Estonian children in relation to environmental influences. Tartu, 1998.
41. **Kaja Julge**. Humoral immune responses to allergens in early childhood. Tartu, 1998.

42. **Heli Grünberg**. The cardiovascular risk of Estonian schoolchildren. A cross-sectional study of 9-, 12- and 15-year-old children. Tartu, 1998.
43. **Epp Sepp**. Formation of intestinal microbial ecosystem in children. Tartu, 1998.
44. **Mai Ots**. Characteristics of the progression of human and experimental glomerulopathies. Tartu, 1998.
45. **Tiina Ristimäe**. Heart rate variability in patients with coronary artery disease. Tartu, 1998.
46. **Leho Kõiv**. Reaction of the sympatho-adrenal and hypothalamo-pituitary-adrenocortical system in the acute stage of head injury. Tartu, 1998.
47. **Bela Adojaan**. Immune and genetic factors of childhood onset IDDM in Estonia. An epidemiological study. Tartu, 1999.
48. **Jakov Shlik**. Psychophysiological effects of cholecystokinin in humans. Tartu, 1999.
49. **Kai Kisand**. Autoantibodies against dehydrogenases of α-ketoacids. Tartu, 1999.
50. **Toomas Marandi**. Drug treatment of depression in Estonia. Tartu, 1999.
51. **Ants Kask**. Behavioural studies on neuropeptide Y. Tartu, 1999.
52. **Ello-Rahel Karelson**. Modulation of adenylate cyclase activity in the rat hippocampus by neuropeptide galanin and its chimeric analogs. Tartu, 1999.
53. **Tanel Laisaar**. Treatment of pleural empyema — special reference to intrapleural therapy with streptokinase and surgical treatment modalities. Tartu, 1999.
54. **Eve Pihl**. Cardiovascular risk factors in middle-aged former athletes. Tartu, 1999.
55. **Katrin Õunap**. Phenylketonuria in Estonia: incidence, newborn screening, diagnosis, clinical characterization and genotype/phenotype correlation. Tartu, 1999.
56. **Siiri Kõljalg**. *Acinetobacter* – an important nosocomial pathogen. Tartu, 1999.
57. **Helle Karro**. Reproductive health and pregnancy outcome in Estonia: association with different factors. Tartu, 1999.
58. **Heili Varendi**. Behavioral effects observed in human newborns during exposure to naturally occurring odors. Tartu, 1999.
59. **Anneli Beilmann**. Epidemiology of epilepsy in children and adolescents in Estonia. Prevalence, incidence, and clinical characteristics. Tartu, 1999.
60. **Vallo Volke**. Pharmacological and biochemical studies on nitric oxide in the regulation of behaviour. Tartu, 1999.
61. **Pilvi Ilves**. Hypoxic-ischaemic encephalopathy in asphyxiated term infants. A prospective clinical, biochemical, ultrasonographical study. Tartu, 1999.
62. **Anti Kalda**. Oxygen-glucose deprivation-induced neuronal death and its pharmacological prevention in cerebellar granule cells. Tartu, 1999.
63. **Eve-Irene Lepist**. Oral peptide prodrugs – studies on stability and absorption. Tartu, 2000.

64. **Jana Kivastik**. Lung function in Estonian schoolchildren: relationship with anthropometric indices and respiratory symptomas, reference values for dynamic spirometry. Tartu, 2000.
65. **Karin Kull**. Inflammatory bowel disease: an immunogenetic study. Tartu, 2000.
66. **Kaire Innos**. Epidemiological resources in Estonia: data sources, their quality and feasibility of cohort studies. Tartu, 2000.
67. **Tamara Vorobjova**. Immune response to *Helicobacter pylori* and its association with dynamics of chronic gastritis and epithelial cell turnover in antrum and corpus. Tartu, 2001.
68. **Ruth Kalda**. Structure and outcome of family practice quality in the changing health care system of Estonia. Tartu, 2001.
69. **Annika Krüüner**. *Mycobacterium tuberculosis* – spread and drug resistance in Estonia. Tartu, 2001.
70. **Marlit Veldi**. Obstructive Sleep Apnoea: Computerized Endopharyngeal Myotonometry of the Soft Palate and Lingual Musculature. Tartu, 2001.
71. **Anneli Uusküla**. Epidemiology of sexually transmitted diseases in Estonia in 1990–2000. Tartu, 2001.
72. **Ade Kallas**. Characterization of antibodies to coagulation factor VIII. Tartu, 2002.
73. **Heidi Annuk**. Selection of medicinal plants and intestinal lactobacilli as antimicrobil components for functional foods. Tartu, 2002.
74. **Aet Lukmann**. Early rehabilitation of patients with ischaemic heart disease after surgical revascularization of the myocardium: assessment of health-related quality of life, cardiopulmonary reserve and oxidative stress. A clinical study. Tartu, 2002.
75. **Maigi Eisen**. Pathogenesis of Contact Dermatitis: participation of Oxidative Stress. A clinical – biochemical study. Tartu, 2002.
76. **Piret Hussar**. Histology of the post-traumatic bone repair in rats. Elaboration and use of a new standardized experimental model – bicortical perforation of tibia compared to internal fracture and resection osteotomy. Tartu, 2002.
77. **Tõnu Rätsep**. Aneurysmal subarachnoid haemorrhage: Noninvasive monitoring of cerebral haemodynamics. Tartu, 2002.
78. **Marju Herodes**. Quality of life of people with epilepsy in Estonia. Tartu, 2003.
79. **Katre Maasalu**. Changes in bone quality due to age and genetic disorders and their clinical expressions in Estonia. Tartu, 2003.
80. **Toomas Sillakivi**. Perforated peptic ulcer in Estonia: epidemiology, risk factors and relations with *Helicobacter pylori*. Tartu, 2003.
81. **Leena Puksa**. Late responses in motor nerve conduction studies. F and A waves in normal subjects and patients with neuropathies. Tartu, 2003.
82. **Krista Lõivukene**. *Helicobacter pylori* in gastric microbial ecology and its antimicrobial susceptibility pattern. Tartu, 2003.

83. **Helgi Kolk**. Dyspepsia and *Helicobacter pylori* infection: the diagnostic value of symptoms, treatment and follow-up of patients referred for upper gastrointestinal endoscopy by family physicians. Tartu, 2003.
84. **Helena Soomer**. Validation of identification and age estimation methods in forensic odontology. Tartu, 2003.
85. **Kersti Oselin**. Studies on the human MDR1, MRP1, and MRP2 ABC transporters: functional relevance of the genetic polymorphisms in the *MDR1* and *MRP1* gene. Tartu, 2003.
86. **Jaan Soplepmann**. Peptic ulcer haemorrhage in Estonia: epidemiology, prognostic factors, treatment and outcome. Tartu, 2003.
87. **Margot Peetsalu**. Long-term follow-up after vagotomy in duodenal ulcer disease: recurrent ulcer, changes in the function, morphology and *Helico-bacter pylori* colonisation of the gastric mucosa. Tartu, 2003.
88. **Kersti Klaamas**. Humoral immune response to *Helicobacter pylori* a study of host-dependent and microbial factors. Tartu, 2003.
89. **Pille Taba**. Epidemiology of Parkinson's disease in Tartu, Estonia. Pre-valence, incidence, clinical characteristics, and pharmacoepidemiology. Tartu, 2003.
90. **Alar Veraksitš**. Characterization of behavioural and biochemical pheno-type of cholecystokinin-2 receptor deficient mice: changes in the function of the dopamine and endopioidergic system. Tartu, 2003.
91. **Ingrid Kalev**. CC-chemokine receptor 5 (CCR5) gene polymorphism in Estonians and in patients with Type I and Type II diabetes mellitus. Tartu, 2003.
92. **Lumme Kadaja**. Molecular approach to the regulation of mitochondrial function in oxidative muscle cells. Tartu, 2003.
93. **Aive Liigant**. Epidemiology of primary central nervous system tumours in Estonia from 1986 to 1996. Clinical characteristics, incidence, survival and prognostic factors. Tartu, 2004.
94. **Andres, Kulla**. Molecular characteristics of mesenchymal stroma in human astrocytic gliomas. Tartu, 2004.
95. **Mari Järvelaid**. Health damaging risk behaviours in adolescence. Tartu, 2004.
96. **Ülle Pechter**. Progression prevention strategies in chronic renal failure and hypertension. An experimental and clinical study. Tartu, 2004.
97. **Gunnar Tasa**. Polymorphic glutathione S-transferases – biology and role in modifying genetic susceptibility to senile cataract and primary open angle glaucoma. Tartu, 2004.
98. **Tuuli Käämbre**. Intracellular energetic unit: structural and functional aspects. Tartu, 2004.
99. **Vitali Vassiljev**. Influence of nitric oxide syntase inhibitors on the effects of ethanol after acute and chronic ethanol administration and withdrawal. Tartu, 2004.

100. **Aune Rehema**. Assessment of nonhaem ferrous iron and glutathione redox ratio as markers of pathogeneticity of oxidative stress in different clinical groups. Tartu, 2004.
101. **Evelin Seppet**. Interaction of mitochondria and ATPases in oxidative muscle cells in normal and pathological conditions. Tartu, 2004.
102. **Eduard Maron**. Serotonin function in panic disorder: from clinical experiments to brain imaging and genetics. Tartu, 2004.
103. **Marje Oona**. *Helicobacter pylori* infection in children: epidemiological and therapeutic aspects. Tartu, 2004.
104. **Kersti Kokk**. Regulation of active and passive molecular transport in the testis. Tartu, 2005.
105. **Vladimir Järv**. Cross-sectional imaging for pretreatment evaluation and follow-up of pelvic malignant tumours. Tartu, 2005.
106. **Andre Õun**. Epidemiology of adult epilepsy in Tartu, Estonia. Incidence, prevalence and medical treatment. Tartu, 2005.
107. **Piibe Muda**. Homocysteine and hypertension: associations between homocysteine and essential hypertension in treated and untreated hypertensive patients with and without coronary artery disease. Tartu, 2005.
108. **Külli Kingo**. The interleukin-10 family cytokines gene polymorphisms in plaque psoriasis. Tartu, 2005.
109. **Mati Merila**. Anatomy and clinical relevance of the glenohumeral joint capsule and ligaments. Tartu, 2005.
110. **Epp Songisepp**. Evaluation of technological and functional properties of the new probiotic *Lactobacillus fermentum* ME-3. Tartu, 2005.
111. **Tiia Ainla**. Acute myocardial infarction in Estonia: clinical characteristics, management and outcome. Tartu, 2005.
112. **Andres Sell**. Determining the minimum local anaesthetic requirements for hip replacement surgery under spinal anaesthesia – a study employing a spinal catheter. Tartu, 2005.
113. **Tiia Tamme**. Epidemiology of odontogenic tumours in Estonia. Pathogenesis and clinical behaviour of ameloblastoma. Tartu, 2005.
114. **Triine Annus**. Allergy in Estonian schoolchildren: time trends and characteristics. Tartu, 2005.
115. **Tiia Voor**. Microorganisms in infancy and development of allergy: comparison of Estonian and Swedish children. Tartu, 2005.
116. **Priit Kasenõmm**. Indicators for tonsillectomy in adults with recurrent tonsillitis – clinical, microbiological and pathomorphological investigations. Tartu, 2005.
117. **Eva Zusinaite**. Hepatitis C virus: genotype identification and interactions between viral proteases. Tartu, 2005.
118. **Piret Kõll**. Oral lactoflora in chronic periodontitis and periodontal health. Tartu, 2006.
119. **Tiina Stelmach**. Epidemiology of cerebral palsy and unfavourable neurodevelopmental outcome in child population of Tartu city and county, Estonia Prevalence, clinical features and risk factors. Tartu, 2006.

120. **Katrin Pudersell**. Tropane alkaloid production and riboflavine excretion in the field and tissue cultures of henbane (*Hyoscyamus niger* L.). Tartu, 2006.
121. **Külli Jaako**. Studies on the role of neurogenesis in brain plasticity. Tartu, 2006.
122. **Aare Märtson**. Lower limb lengthening: experimental studies of bone regeneration and long-term clinical results. Tartu, 2006.
123. **Heli Tähepõld**. Patient consultation in family medicine. Tartu, 2006.
124. **Stanislav Liskmann**. Peri-implant disease: pathogenesis, diagnosis and treatment in view of both inflammation and oxidative stress profiling. Tartu, 2006.
125. **Ruth Rudissaar**. Neuropharmacology of atypical antipsychotics and an animal model of psychosis. Tartu, 2006.
126. **Helena Andreson**. Diversity of *Helicobacter pylori* genotypes in Estonian patients with chronic inflammatory gastric diseases. Tartu, 2006.
127. **Katrin Pruus**. Mechanism of action of antidepressants: aspects of serotoninergic system and its interaction with glutamate. Tartu, 2006.
128. **Priit Põder**. Clinical and experimental investigation: relationship of ischaemia/reperfusion injury with oxidative stress in abdominal aortic aneurysm repair and in extracranial brain artery endarterectomy and possibilities of protection against ischaemia using a glutathione analogue in a rat model of global brain ischaemia. Tartu, 2006.
129. **Marika Tammaru**. Patient-reported outcome measurement in rheumatoid arthritis. Tartu, 2006.
130. **Tiia Reimand**. Down syndrome in Estonia. Tartu, 2006.
131. **Diva Eensoo**. Risk-taking in traffic and Markers of Risk-Taking Behaviour in Schoolchildren and Car Drivers. Tartu, 2007.
132. **Riina Vibo**. The third stroke registry in Tartu, Estonia from 2001 to 2003: incidence, case-fatality, risk factors and long-term outcome. Tartu, 2007.
133. **Chris Pruunsild**. Juvenile idiopathic arthritis in children in Estonia. Tartu, 2007.
134. **Eve Õiglane-Šlik**. Angelman and Prader-Willi syndromes in Estonia. Tartu, 2007.
135. **Kadri Haller**. Antibodies to follicle stimulating hormone. Significance in female infertility. Tartu, 2007.
136. **Pille Ööpik**. Management of depression in family medicine. Tartu, 2007.
137. **Jaak Kals**. Endothelial function and arterial stiffness in patients with atherosclerosis and in healthy subjects. Tartu, 2007.
138. **Priit Kampus**. Impact of inflammation, oxidative stress and age on arterial stiffness and carotid artery intima-media thickness. Tartu, 2007.
139. **Margus Punab**. Male fertility and its risk factors in Estonia. Tartu, 2007.
140. **Alar Toom**. Heterotopic ossification after total hip arthroplasty: clinical and pathogenetic investigation. Tartu, 2007.

141. **Lea Pehme**. Epidemiology of tuberculosis in Estonia 1991–2003 with special regard to extrapulmonary tuberculosis and delay in diagnosis of pulmonary tuberculosis. Tartu, 2007.
142. **Juri Karjagin**. The pharmacokinetics of metronidazole and meropenem in septic shock. Tartu, 2007.
143. **Inga Talvik**. Inflicted traumatic brain injury shaken baby syndrome in Estonia – epidemiology and outcome. Tartu, 2007.
144. **Tarvo Rajasalu**. Autoimmune diabetes: an immunological study of type 1 diabetes in humans and in a model of experimental diabetes (in RIP-B7.1 mice). Tartu, 2007.
145. **Inga Karu**. Ischaemia-reperfusion injury of the heart during coronary surgery: a clinical study investigating the effect of hyperoxia. Tartu, 2007.
146. **Peeter Padrik**. Renal cell carcinoma: Changes in natural history and treatment of metastatic disease. Tartu, 2007.
147. **Neve Vendt**. Iron deficiency and iron deficiency anaemia in infants aged 9 to 12 months in Estonia. Tartu, 2008.
148. **Lenne-Triin Heidmets**. The effects of neurotoxins on brain plasticity: focus on neural Cell Adhesion Molecule. Tartu, 2008.
149. **Paul Korrovits**. Asymptomatic inflammatory prostatitis: prevalence, etiological factors, diagnostic tools. Tartu, 2008.
150. **Annika Reintam**. Gastrointestinal failure in intensive care patients. Tartu, 2008.
151. **Kristiina Roots**. Cationic regulation of Na-pump in the normal, Alzheimer's and CCK_2 receptor-deficient brain. Tartu, 2008.
152. **Helen Puusepp**. The genetic causes of mental retardation in Estonia: fragile X syndrome and creatine transporter defect. Tartu, 2009.
153. **Kristiina Rull**. Human chorionic gonadotropin beta genes and recurrent miscarriage: expression and variation study. Tartu, 2009.
154. **Margus Eimre**. Organization of energy transfer and feedback regulation in oxidative muscle cells. Tartu, 2009.
155. **Maire Link**. Transcription factors FoxP3 and AIRE: autoantibody associations. Tartu, 2009.
156. **Kai Haldre**. Sexual health and behaviour of young women in Estonia. Tartu, 2009.
157. **Kaur Liivak**. Classical form of congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Estonia: incidence, genotype and phenotype with special attention to short-term growth and 24-hour blood pressure. Tartu, 2009.
158. **Kersti Ehrlich**. Antioxidative glutathione analogues (UPF peptides) – molecular design, structure-activity relationships and testing the protective properties. Tartu, 2009.
159. **Anneli Rätsep**. Type 2 diabetes care in family medicine. Tartu, 2009.
160. **Silver Türk**. Etiopathogenetic aspects of chronic prostatitis: role of mycoplasmas, coryneform bacteria and oxidative stress. Tartu, 2009.

161. **Kaire Heilman**. Risk markers for cardiovascular disease and low bone mineral density in children with type 1 diabetes. Tartu, 2009.
162. **Kristi Rüütel**. HIV-epidemic in Estonia: injecting drug use and quality of life of people living with HIV. Tartu, 2009.
163. **Triin Eller**. Immune markers in major depression and in antidepressive treatment. Tartu, 2009.
164. **Siim Suutre**. The role of TGF-β isoforms and osteoprogenitor cells in the pathogenesis of heterotopic ossification. An experimental and clinical study of hip arthroplasty. Tartu, 2010.
165. **Kai Kliiman**. Highly drug-resistant tuberculosis in Estonia: Risk factors and predictors of poor treatment outcome. Tartu, 2010.
166. **Inga Villa**. Cardiovascular health-related nutrition, physical activity and fitness in Estonia. Tartu, 2010.
167. **Tõnis Org**. Molecular function of the first PHD finger domain of Autoimmune Regulator protein. Tartu, 2010.
168. **Tuuli Metsvaht**. Optimal antibacterial therapy of neonates at risk of early onset sepsis. Tartu, 2010.
169. **Jaanus Kahu**. Kidney transplantation: Studies on donor risk factors and mycophenolate mofetil. Tartu, 2010.
170. **Koit Reimand**. Autoimmunity in reproductive failure: A study on associated autoantibodies and autoantigens. Tartu, 2010.
171. **Mart Kull**. Impact of vitamin D and hypolactasia on bone mineral density: a population based study in Estonia. Tartu, 2010.
172. **Rael Laugesaar**. Stroke in children – epidemiology and risk factors. Tartu, 2010.
173. **Mark Braschinsky**. Epidemiology and quality of life issues of hereditary spastic paraplegia in Estonia and implemention of genetic analysis in everyday neurologic practice. Tartu, 2010.
174. **Kadri Suija**. Major depression in family medicine: associated factors, recurrence and possible intervention. Tartu, 2010.
175. **Jarno Habicht**. Health care utilisation in Estonia: socioeconomic determinants and financial burden of out-of-pocket payments. Tartu, 2010.
176. **Kristi Abram**. The prevalence and risk factors of rosacea. Subjective disease perception of rosacea patients. Tartu, 2010.
177. **Malle Kuum**. Mitochondrial and endoplasmic reticulum cation fluxes: Novel roles in cellular physiology. Tartu, 2010.
178. **Rita Teek**. The genetic causes of early onset hearing loss in Estonian children. Tartu, 2010.
179. **Daisy Volmer**. The development of community pharmacy services in Estonia – public and professional perceptions 1993–2006. Tartu, 2010.
180. **Jelena Lissitsina**. Cytogenetic causes in male infertility. Tartu, 2011.
181. **Delia Lepik**. Comparison of gunshot injuries caused from Tokarev, Makarov and Glock 19 pistols at different firing distances. Tartu, 2011.
182. **Ene-Renate Pähkla**. Factors related to the efficiency of treatment of advanced periodontitis. Tartu, 2011.

183. **Maarja Krass**. L-Arginine pathways and antidepressant action. Tartu, 2011.
184. **Taavi Lai**. Population health measures to support evidence-based health policy in Estonia. Tartu, 2011.
185. **Tiit Salum**. Similarity and difference of temperature-dependence of the brain sodium pump in normal, different neuropathological, and aberrant conditions and its possible reasons. Tartu, 2011.
186. **Tõnu Vooder**. Molecular differences and similarities between histological subtypes of non-small cell lung cancer. Tartu, 2011.
187. **Jelena Štšepetova**. The characterisation of intestinal lactic acid bacteria using bacteriological, biochemical and molecular approaches. Tartu, 2011.
188. **Radko Avi**. Natural polymorphisms and transmitted drug resistance in Estonian HIV-1 CRF06_cpx and its recombinant viruses. Tartu, 2011, 116 p.
189. **Edward Laane**. Multiparameter flow cytometry in haematological malignancies. Tartu, 2011, 152 p.
190. **Triin Jagomägi**. A study of the genetic etiology of nonsyndromic cleft lip and palate. Tartu, 2011, 158 p.
191. **Ivo Laidmäe**. Fibrin glue of fish (*Salmo salar*) origin: immunological study and development of new pharmaceutical preparation. Tartu, 2012, 150 p.
192. **Ülle Parm**. Early mucosal colonisation and its role in prediction of invasive infection in neonates at risk of early onset sepsis. Tartu, 2012, 168 p.
193. **Kaupo Teesalu**. Autoantibodies against desmin and transglutaminase 2 in celiac disease: diagnostic and functional significance. Tartu, 2012, 142 p.
194. **Maksim Zagura**. Biochemical, functional and structural profiling of arterial damage in atherosclerosis. Tartu, 2012, 162 p.
195. **Vivian Kont**. Autoimmune regulator: characterization of thymic gene regulation and promoter methylation. Tartu, 2012, 134 p.
196. **Pirje Hütt**. Functional properties, persistence, safety and efficacy of potential probiotic lactobacilli. Tartu, 2012, 246 p.
197. **Innar Tõru**. Serotonergic modulation of CCK-4- induced panic. Tartu, 2012, 132 p.
198. **Sigrid Vorobjov**. Drug use, related risk behaviour and harm reduction interventions utilization among injecting drug users in Estonia: implications for drug policy. Tartu, 2012, 120 p.
199. **Martin Serg**. Therapeutic aspects of central haemodynamics, arterial stiffness and oxidative stress in hypertension. Tartu, 2012, 156 p.
200. **Jaanika Kumm**. Molecular markers of articular tissues in early knee osteoarthritis: a population-based longitudinal study in middle-aged subjects. Tartu, 2012, 159 p.
201. **Kertu Rünkorg**. Functional changes of dopamine, endopioid and endocannabinoid systems in CCK2 receptor deficient mice. Tartu, 2012, 125 p.
202. **Mai Blöndal**. Changes in the baseline characteristics, management and outcomes of acute myocardial infarction in Estonia. Tartu, 2012, 127 p.

203. **Jana Lass**. Epidemiological and clinical aspects of medicines use in children in Estonia. Tartu, 2012, 170 p.
204. **Kai Truusalu**. Probiotic lactobacilli in experimental persistent *Salmonella* infection. Tartu, 2013, 139 p.
205. **Oksana Jagur**. Temporomandibular joint diagnostic imaging in relation to pain and bone characteristics. Long-term results of arthroscopic treatment. Tartu, 2013, 126 p.
206. **Katrin Sikk**. Manganese-ephedrone intoxication – pathogenesis of neurological damage and clinical symptomatology. Tartu, 2013, 125 p.
207. **Kai Blöndal**. Tuberculosis in Estonia with special emphasis on drug-resistant tuberculosis: Notification rate, disease recurrence and mortality. Tartu, 2013, 151 p.
208. **Marju Puurand**. Oxidative phosphorylation in different diseases of gastric mucosa. Tartu, 2013, 123 p.
209. **Aili Tagoma**. Immune activation in female infertility: Significance of autoantibodies and inflammatory mediators. Tartu, 2013, 135 p.
210. **Liis Sabre**. Epidemiology of traumatic spinal cord injury in Estonia. Brain activation in the acute phase of traumatic spinal cord injury. Tartu, 2013, 135 p.
211. **Merit Lamp**. Genetic susceptibility factors in endometriosis. Tartu, 2013, 125 p.
212. **Erik Salum**. Beneficial effects of vitamin D and angiotensin II receptor blocker on arterial damage. Tartu, 2013, 167 p.
213. **Maire Karelson**. Vitiligo: clinical aspects, quality of life and the role of melanocortin system in pathogenesis. Tartu, 2013, 153 p.
214. **Kuldar Kaljurand**. Prevalence of exfoliation syndrome in Estonia and its clinical significance. Tartu, 2013, 113 p.
215. **Raido Paasma**. Clinical study of methanol poisoning: handling large outbreaks, treatment with antidotes, and long-term outcomes. Tartu, 2013, 96 p.
216. **Anne Kleinberg**. Major depression in Estonia: prevalence, associated factors, and use of health services. Tartu, 2013, 129 p.
217. **Triin Eglit**. Obesity, impaired glucose regulation, metabolic syndrome and their associations with high-molecular-weight adiponectin levels. Tartu, 2014, 115 p.
218. **Kristo Ausmees**. Reproductive function in middle-aged males: Associations with prostate, lifestyle and couple infertility status. Tartu, 2014, 125 p.
219. **Kristi Huik**. The influence of host genetic factors on the susceptibility to HIV and HCV infections among intravenous drug users. Tartu, 2014, 144 p.
220. **Liina Tserel**. Epigenetic profiles of monocytes, monocyte-derived macrophages and dendritic cells. Tartu, 2014, 143 p.
221. **Irina Kerna**. The contribution of *ADAM12* and *CILP* genes to the development of knee osteoarthritis. Tartu, 2014, 152 p.

222. **Ingrid Liiv**. Autoimmune regulator protein interaction with DNA-dependent protein kinase and its role in apoptosis. Tartu, 2014, 143 p.
223. **Liivi Maddison**. Tissue perfusion and metabolism during intra-abdominal hypertension. Tartu, 2014, 103 p.
224. **Krista Ress**. Childhood coeliac disease in Estonia, prevalence in atopic dermatitis and immunological characterisation of coexistence. Tartu, 2014, 124 p.
225. **Kai Muru**. Prenatal screening strategies, long-term outcome of children with marked changes in maternal screening tests and the most common syndromic heart anomalies in Estonia. Tartu, 2014, 189 p.
226. **Kaja Rahu**. Morbidity and mortality among Baltic Chernobyl cleanup workers: a register-based cohort study. Tartu, 2014, 155 p.
227. **Klari Noormets**. The development of diabetes mellitus, fertility and energy metabolism disturbances in a Wfs1-deficient mouse model of Wolfram syndrome. Tartu, 2014, 132 p.
228. **Liis Toome**. Very low gestational age infants in Estonia. Tartu, 2014, 183 p.
229. **Ceith Nikkolo**. Impact of different mesh parameters on chronic pain and foreign body feeling after open inguinal hernia repair. Tartu, 2014, 132 p.
230. **Vadim Brjalin**. Chronic hepatitis C: predictors of treatment response in Estonian patients. Tartu, 2014, 122 p.
231. **Vahur Metsna**. Anterior knee pain in patients following total knee arthroplasty: the prevalence, correlation with patellar cartilage impairment and aspects of patellofemoral congruence. Tartu, 2014, 130 p.
232. **Marju Kase**. Glioblastoma multiforme: possibilities to improve treatment efficacy. Tartu, 2015, 137 p.
233. **Riina Runnel**. Oral health among elementary school children and the effects of polyol candies on the prevention of dental caries. Tartu, 2015, 112 p.
234. **Made Laanpere**. Factors influencing women's sexual health and reproductive choices in Estonia. Tartu, 2015, 176 p.
235. **Andres Lust**. Water mediated solid state transformations of a polymorphic drug – effect on pharmaceutical product performance. Tartu, 2015, 134 p.
236. **Anna Klugman**. Functionality related characterization of pretreated wood lignin, cellulose and polyvinylpyrrolidone for pharmaceutical applications. Tartu, 2015, 156 p.
237. **Triin Laisk-Podar**. Genetic variation as a modulator of susceptibility to female infertility and a source for potential biomarkers. Tartu, 2015, 155 p.
238. **Mailis Tõnisson**. Clinical picture and biochemical changes in blood in children with acute alcohol intoxication. Tartu, 2015, 100 p.
239. **Kadri Tamme**. High volume haemodiafiltration in treatment of severe sepsis – impact on pharmacokinetics of antibiotics and inflammatory response. Tartu, 2015, 133 p.

240. **Kai Part**. Sexual health of young people in Estonia in a social context: the role of school-based sexuality education and youth-friendly counseling services. Tartu, 2015, 203 p.
241. **Urve Paaver**. New perspectives for the amorphization and physical stabilization of poorly water-soluble drugs and understanding their dissolution behavior. Tartu, 2015, 139 p.
242. **Aleksandr Peet**. Intrauterine and postnatal growth in children with HLA-conferred susceptibility to type 1 diabetes. Tartu. 2015, 146 p.
243. **Piret Mitt**. Healthcare-associated infections in Estonia – epidemiology and surveillance of bloodstream and surgical site infections. Tartu, 2015, 145 p.
244. **Merli Saare**. Molecular Profiling of Endometriotic Lesions and Endometria of Endometriosis Patients. Tartu, 2016, 129 p.
245. **Kaja-Triin Laisaar**. People living with HIV in Estonia: Engagement in medical care and methods of increasing adherence to antiretroviral therapy and safe sexual behavior. Tartu, 2016, 132 p.
246. **Eero Merilind**. Primary health care performance: impact of payment and practice-based characteristics. Tartu, 2016, 120 p.
247. **Jaanika Kärner**. Cytokine-specific autoantibodies in AIRE deficiency. Tartu, 2016, 182 p.
248. **Kaido Paapstel**. Metabolomic profile of arterial stiffness and early biomarkers of renal damage in atherosclerosis. Tartu, 2016, 173 p.
249. **Liidia Kiisk**. Long-term nutritional study: anthropometrical and clinicolaboratory assessments in renal replacement therapy patients after intensive nutritional counselling. Tartu, 2016, 207 p.
250. **Georgi Nellis**. The use of excipients in medicines administered to neonates in Europe. Tartu, 2017, 159 p.
251. **Aleksei Rakitin**. Metabolic effects of acute and chronic treatment with valproic acid in people with epilepsy. Tartu, 2017, 125 p.
252. **Eveli Kallas**. The influence of immunological markers to susceptibility to HIV, HBV, and HCV infections among persons who inject drugs. Tartu, 2017, 138 p.
253. **Tiina Freimann**. Musculoskeletal pain among nurses: prevalence, risk factors, and intervention. Tartu, 2017, 125 p.
254. **Evelyn Aaviksoo**. Sickness absence in Estonia: determinants and influence of the sick-pay cut reform. Tartu, 2017, 121 p.
255. **Kalev Nõupuu**. Autosomal-recessive Stargardt disease: phenotypic heterogeneity and genotype-phenotype associations. Tartu, 2017, 131 p.
256. **Ho Duy Binh**. Osteogenesis imperfecta in Vietnam. Tartu, 2017, 125 p.
257. **Uku Haljasorg**. Transcriptional mechanisms in thymic central tolerance. Tartu, 2017, 147 p.
258. **Živile Riispere**. IgA Nephropathy study according to the Oxford Classification: IgA Nephropathy clinical-morphological correlations, disease progression and the effect of renoprotective therapy. Tartu, 2017, 129 p.

259. **Hiie Soeorg**. Coagulase-negative staphylococci in gut of preterm neonates and in breast milk of their mothers. Tartu, 2017, 216 p.
260. **Anne-Mari Anton Willmore**. Silver nanoparticles for cancer research. Tartu, 2017, 132 p.
261. **Ott Laius**. Utilization of osteoporosis medicines, medication adherence and the trend in osteoporosis related hip fractures in Estonia. Tartu, 2017, 134 p.
262. **Alar Aab**. Insights into molecular mechanisms of asthma and atopic dermatitis. Tartu, 2017, 164 p.
263. **Sander Pajusalu**. Genome-wide diagnostics of Mendelian disorders: from chromosomal microarrays to next-generation sequencing. Tartu, 2017, 146 p.
264. **Mikk Jürisson**. Health and economic impact of hip fracture in Estonia. Tartu, 2017, 164 p.
265. **Kaspar Tootsi**. Cardiovascular and metabolomic profiling of osteo-arthritis. Tartu, 2017, 150 p.
266. **Mario Saare**. The influence of AIRE on gene expression – studies of transcriptional regulatory mechanisms in cell culture systems. Tartu, 2017, 172 p.
267. **Piia Jõgi**. Epidemiological and clinical characteristics of pertussis in Estonia. Tartu, 2018, 168 p.
268. **Elle Põldoja**. Structure and blood supply of the superior part of the shoulder joint capsule. Tartu, 2018, 116 p.
269. **Minh Son Nguyen**. Oral health status and prevalence of temporo-mandibular disorders in 65–74-year-olds in Vietnam. Tartu, 2018, 182 p.
270. **Kristian Semjonov**. Development of pharmaceutical quench-cooled molten and melt-electrospun solid dispersions for poorly water-soluble indomethacin. Tartu, 2018, 125 p.
271. **Janne Tiigimäe-Saar**. Botulinum neurotoxin type A treatment for sialorrhea in central nervous system diseases. Tartu, 2018, 109 p.
272. **Veiko Vengerfeldt**. Apical periodontitis: prevalence and etiopathogenetic aspects. Tartu, 2018, 150 p.
273. **Rudolf Bichele**. TNF superfamily and AIRE at the crossroads of thymic differentiation and host protection against *Candida albicans* infection. Tartu, 2018, 153 p.
274. **Olga Tšuiko**. Unravelling Chromosomal Instability in Mammalian Pre-implantation Embryos Using Single-Cell Genomics. Tartu, 2018, 169 p.
275. **Kärt Kriisa**. Profile of acylcarnitines, inflammation and oxidative stress in first-episode psychosis before and after antipsychotic treatment. Tartu, 2018, 145 p.
276. **Xuan Dung Ho**. Characterization of the genomic profile of osteosarcoma. Tartu, 2018, 144 p.
277. **Karit Reinson**. New Diagnostic Methods for Early Detection of Inborn Errors of Metabolism in Estonia. Tartu, 2018, 201 p.

278. **Mari-Anne Vals**. Congenital N-glycosylation Disorders in Estonia. Tartu, 2019, 148 p.
279. **Liis Kadastik-Eerme**. Parkinson's disease in Estonia: epidemiology, quality of life, clinical characteristics and pharmacotherapy. Tartu, 2019, 202 p.
280. **Hedi Hunt**. Precision targeting of intraperitoneal tumors with peptide-guided nanocarriers. Tartu, 2019, 179 p.
281. **Rando Porosk**. The role of oxidative stress in Wolfram syndrome 1 and hypothermia. Tartu, 2019, 123 p.
282. **Ene-Ly Jõgeda**. The influence of coinfections and host genetic factor on the susceptibility to HIV infection among people who inject drugs. Tartu, 2019, 126 p.
283. **Kristel Ehala-Aleksejev**. The associations between body composition, obesity and obesity-related health and lifestyle conditions with male reproductive function. Tartu, 2019, 138 p.
284. **Aigar Ottas**. The metabolomic profiling of psoriasis, atopic dermatitis and atherosclerosis. Tartu, 2019, 136 p.
285. **Elmira Gurbanova**. Specific characteristics of tuberculosis in low default, but high multidrug–resistance prison setting. Tartu, 2019, 129 p.
286. **Van Thai Nguyeni**. The first study of the treatment outcomes of patients with cleft lip and palate in Central Vietnam. Tartu, 2019, 144 p.
287. **Maria Yakoreva**. Imprinting Disorders in Estonia. Tartu, 2019, 187 p.
288. **Kadri Rekker**. The putative role of microRNAs in endometriosis pathogenesis and potential in diagnostics. Tartu, 2019, 140 p.
289. **Ülle Võhma**. Association between personality traits, clinical characteristics and pharmacological treatment response in panic disorder. Tartu, 2019, 121 p.
290. **Aet Saar**. Acute myocardial infarction in Estonia 2001–2014: towards risk-based prevention and management. Tartu, 2019, 124 p.
291. **Toomas Toomsoo**. Transcranial brain sonography in the Estonian cohort of Parkinson's disease. Tartu, 2019, 114 p.
292. **Lidiia Zhytnik**. Inter- and intrafamilial diversity based on genotype and phenotype correlations of Osteogenesis Imperfecta. Tartu, 2019, 224 p.
293. **Pilleriin Soodla**. Newly HIV-infected people in Estonia: estimation of incidence and transmitted drug resistance. Tartu, 2019, 194 p.
294. **Kristiina Ojamaa**. Epidemiology of gynecological cancer in Estonia. Tartu, 2020, 133 p.
295. **Marianne Saard**. Modern Cognitive and Social Intervention Techniques in Paediatric Neurorehabilitation for Children with Acquired Brain Injury. Tartu, 2020, 168 p.
296. **Julia Maslovskaja**. The importance of DNA binding and DNA breaks for AIRE-mediated transcriptional activation. Tartu, 2020, 162 p.
297. **Natalia Lobanovskaya**. The role of PSA-NCAM in the survival of retinal ganglion cells. Tartu, 2020, 105 p.

298. **Madis Rahu**. Structure and blood supply of the postero-superior part of the shoulder joint capsule with implementation of surgical treatment after anterior traumatic dislocation. Tartu, 2020, 104 p.
299. **Helen Zirnask**. Luteinizing hormone (LH) receptor expression in the penis and its possible role in pathogenesis of erectile disturbances. Tartu, 2020, 87 p.
300. **Kadri Toome**. Homing peptides for targeting of brain diseases. Tartu, 2020, 152 p.
301. **Maarja Hallik**. Pharmacokinetics and pharmacodynamics of inotropic drugs in neonates. Tartu, 2020, 172 p.
302. **Raili Müller**. Cardiometabolic risk profile and body composition in early rheumatoid arthritis. Tartu, 2020, 133 p.
303. **Sergo Kasvandik**. The role of proteomic changes in endometrial cells – from the perspective of fertility and endometriosis. Tartu, 2020, 191 p.
304. **Epp Kaleviste**. Genetic variants revealing the role of STAT1/STAT3 signaling cytokines in immune protection and pathology. Tartu, 2020, 189 p.
305. **Sten Saar**. Epidemiology of severe injuries in Estonia. Tartu, 2020, 104 p.
306. **Kati Braschinsky**. Epidemiology of primary headaches in Estonia and applicability of web-based solutions in headache epidemiology research. Tartu, 2020, 129 p.
307. **Helen Vaher**. MicroRNAs in the regulation of keratinocyte responses in *psoriasis vulgaris* and atopic dermatitis. Tartu, 2020, 242 p.
308. **Liisi Raam**. Molecular Alterations in the Pathogenesis of Two Chronic Dermatoses – Vitiligo and Psoriasis. Tartu, 2020, 164 p.
309. **Artur Vetkas**. Long-term quality of life, emotional health, and associated factors in patients after aneurysmal subarachnoid haemorrhage. Tartu, 2020, 127 p.
310. **Teele Kasepalu**. Effects of remote ischaemic preconditioning on organ damage and acylcarnitines' metabolism in vascular surgery. Tartu, 2020, 130 p.
311. **Prakash Lingasamy**. Development of multitargeted tumor penetrating peptides. Tartu, 2020, 246 p.
312. **Lille Kurvits.** Parkinson's disease as a multisystem disorder: whole transcriptome study in Parkinson's disease patients' skin and blood. Tartu, 2021, 142 p.
313. **Mariliis Põld.** Smoking, attitudes towards smoking behaviour, and nicotine dependence among physicians in Estonia: cross-sectional surveys 1982–2014. Tartu, 2021, 172 p.
314. **Triin Kikas**. Single nucleotide variants affecting placental gene expression and pregnancy outcome. Tartu, 2021, 160 p.
315. **Hedda Lippus-Metsaots**. Interpersonal violence in Estonia: prevalence, impact on health and health behaviour. Tartu, 2021, 172 p.

316. **Georgi Dzaparidze.** Quantification and evaluation of the diagnostic significance of adenocarcinoma-associated microenvironmental changes in the prostate using modern digital pathology solutions. Tartu, 2021, 132 p.
317. **Tuuli Sedman.** New avenues for GLP1 receptor agonists in the treatment of diabetes. Tartu, 2021, 118 p.
318. **Martin Padar.** Enteral nutrition, gastrointestinal dysfunction and intestinal biomarkers in critically ill patients. Tartu, 2021, 189 p.
319. **Siim Schneider.** Risk factors, etiology and long-term outcome in young ischemic stroke patients in Estonia. Tartu, 2021, 131 p.
320. **Konstantin Ridnõi.** Implementation and effectiveness of new prenatal diagnostic strategies in Estonia. Tartu, 2021, 191 p.
321. **Risto Vaikjärv.** Etiopathogenetic and clinical aspects of peritonsillar abscess. Tartu, 2021, 115 p.
322. **Liis Preem.** Design and characterization of antibacterial electrospun drug delivery systems for wound infections. Tartu, 2022, 220 p.
323. **Keerthie Dissanayake.** Preimplantation embryo-derived extracellular vesicles: potential as an embryo quality marker and their role during the embryo-maternal communication. Tartu, 2022, 203 p.
324. **Laura Viidik.** 3D printing in pharmaceutics: a new avenue for fabricating therapeutic drug delivery systems. Tartu, 2022, 139 p.
325. **Kasun Godakumara.** Extracellular vesicle mediated embryo-maternal communication – A tool for evaluating functional competency of pre-implantation embryos. Tartu, 2022, 176 p.