Realia et naturalia

DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
256

LIINA NAGIRNAJA

Global and fine-scale genetic determinants of recurrent pregnancy loss

TARTU ÜLIKOOL
1632
UNIVERSITAS TARTUENSIS

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
256

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
256

LIINA NAGIRNAJA

Global and fine-scale genetic determinants of recurrent pregnancy loss

UNIVERSITY OF TARTU
PRESS
1632

Institute of Molecular and Cell Biology, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy (in molecular and cell biology) on March 19, 2014 by the Council of the Institute of Molecular and Cell Biology, University of Tartu.

Supervisor: Prof. Maris Laan, PhD
Institute of Molecular and Cell Biology
University of Tartu
Estonia

Opponent: Dr. Daniel Vaiman, PhD
Cochin Institute
INSERM
Paris Descartes University
France

Commencement: Room No 105, 23B Riia St, Tartu, on May 13th, 2014, at 10.00

The publication of this thesis is granted by the Institute of Molecular and Cell Biology, University of Tartu.

European Union
European Social Fund

Investing in your future

ISSN 1024–6479
ISBN 978–9949–32–526–9 (print)
ISBN 978–9949–32–527–6 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original articles referred to in the text by their Roman numerals:

I Rull K, **Nagirnaja L**, Laan M. Genetics of recurrent miscarriage: challenges, current knowledge, future directions. Front Genet. 2012 Mar 19;3:34.

II **Nagirnaja L**, Rull K, Uusküla L, Hallast P, Grigorova M, Laan M. Genomics and genetics of gonadotropin beta-subunit genes: Unique *FSHB* and duplicated *LHB*/*CGB* loci. Mol Cell Endocrinol. 2010 Nov 25;329(1–2):4–16.

III Hallast P, **Nagirnaja L**, Margus T, Laan M. Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. Genome Res. 2005 Nov;15(11):1535–46.

IV Rull K, **Nagirnaja L**, Ulander VM, Kelgo P, Margus T, Kaare M, Aittomäki K, Laan M. Chorionic gonadotropin beta-gene variants are associated with recurrent miscarriage in two European populations. J Clin Endocrinol Metab. 2008 Dec;93(12):4697–706.

V Rull K, Christiansen OB, **Nagirnaja L**, Steffensen R, Margus T, Laan M. A modest but significant effect of CGB5 gene promoter polymorphisms in modulating the risk of recurrent miscarriage. Fertil Steril. 2013 Jun; 99(7):1930–6.e6.

VI **Nagirnaja L**, Venclovas Č, Rull K, Jonas KC, Peltoketo H, Christiansen OB, Kairys V, Kivi G, Steffensen R, Huhtaniemi IT, Laan M. Structural and functional analysis of rare missense mutations in human chorionic gonadotrophin β-subunit. Mol Hum Reprod. 2012 Aug;18(8):379–90.

VII **Nagirnaja L**, Palta P, Kasak L, Rull K, Christiansen OB, Nielsen HS, Steffensen R, Esko T, Remm M, Laan M. Structural genomic variation as risk factor for idiopathic recurrent miscarriage. *Manuscript submitted*.

Author's contributions:

Ref. I, II – contributed to the preparation of the review article

Ref. III, IV – participated in experimental design, experimental data collection, analysis and interpretation and contributed to manuscript preparation

Ref. V – contributed to *in silico* data analysis, interpretation and manuscript preparation

Ref. VI – contributed to experimental design, conducted the experiments, participated in data analysis and interpretation and wrote the first draft of the manuscript

Ref. VII – contributed to experimental design, participated in data analysis, experimental data collection and interpretation and wrote the first draft of the manuscript

LIST OF ABBREVIATIONS

bp	basepair
C4A	*complement component 4A* gene
C4B	*complement component 4B* gene
cAMP	cyclic adenosine monophosphate
CGβ	chorionic gonadotropin β-subunit
CGB	*chorionic gonadotropin, beta polypeptide* gene
CHO	chinese hamster ovary cell line
CNV	copy number variant
CNVR	copy number variable region
CRE	cAMP response element
CTNNA3	*catenin (cadherin-associated protein), alpha 3* gene
DGV	Database of Genomic Variants
DKK2	*dickkopf 2 homolog* gene
DNA	deoxyribonucleic acid
EGCUT	Estonian population cohort from Estonian Biobank, Estonian Genome Center, University of Tartu
FSH	Follicle-stimulating hormone
GOLPH3	*Golgi phosphoprotein 3* gene
GWAS	genome-wide association study
hCG	human chorionic gonadotropin
hCGα	human chorionic gonadotropin α-subunit
hCGβ	human chorionic gonadotropin β-subunit
hCG-h	hyperglycosylated human chorionic gonadotropin
HEK293	human embryonic kidney 293 cell line
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
hLH/CGR	human luteinizing hormone/chorionic gonadotropin receptor
HPRT	*hypoxanthine phosphoribosyltransferase* gene
IGF2	*insulin-like growth factor 2* gene
IGKV	*Immunoglobulin kappa variable* gene cluster
IVF	*in vitro* fertilization
kb	kilobase
kDa	kilodalton
KIR	killer immunoglobulin-like receptors
KRK	korduv raseduse katkemine
LD	linkage disequilibrium
LH	luteinizing hormone (lutropin)
LHB	*luteinizing hormone beta polypeptide* gene
MAF	minor allele frequency
Mb	megabase
mRNA	messenger ribonucleic acid
mTOR	mechanistic target of rapamycin

NK cells	natural killer cells
NTF5	*neurotrophin 5* gene, currently known as *NTF4*
OR	odds ratio
PDZD2	*PDZ domain containing 2* gene
qPCR	quantitative polymerase chain reaction
RM	recurrent miscarriage
S100A8	*inflammatory marker calprotectin* gene
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	standard error of the mean
SNP	single nucleotide polymorphism
TIMP2	*TIMP metallopeptidase inhibitor 2* gene
TNF	tumor necrosis factor
TRAIL	*TNF-related apoptosis-inducing ligand* gene
WT	wild type

INTRODUCTION

Sporadic miscarriage is the most common pregnancy complication affecting up to 15% of all clinically recognized pregnancies (Stirrat 1990). Up to 3% of fertile couples have been diagnosed with recurrent miscarriage disease (RM) defined by three or more consecutive early pregnancy losses before gestational week 22. RM is a multifactorial disease accompanied by increased probability of other pregnancy complications such as preterm birth or small for gestational age newborns (Jivraj et al. 2001; van Oppenraaij et al. 2009). Although a spectrum of causes is known for RM, etiology behind approximately half of the cases remains unsolved and is defined as idiopathic (reviewed in (Christiansen et al. 2008)).

The genetic component in the etiology of idiopathic RM has been acknowledged and extensively addressed by numerous studies targeting almost 100 candidate genes to date (Kolte et al. 2011; Rull et al. 2012). However, no prevalent gene variants specific to the disease have been described sufficiently explaining the heritability of RM. Few hypothesis-free 'omics' studies, including genome-wide approach or transcriptomics of endometrial and placental samples (Lee et al. 2007; Li et al. 2010; Rull et al. 2013), have been reported identifying novel candidate genes that modify the risk of having RM and highlighting the potential of 'omics' research as the source of novel genetic biomarkers.

Human chorionic gonadotropin (hCG), also termed as the 'pregnancy hormone', is essential for pregnancy establishment and maintenance. The genes encoding the β-subunit of the heterodimeric hCG are attractive candidates in studying early pregnancy success as the aberrant expression has previously been implicated in various pregnancy disorders, including RM (Rull and Laan 2005). The four highly similar genes encoding hCGβ (*CGB*) are co-located in the *LHB/CGB* gene cluster on chromosome 19. The complex nature of the *LHB/CGB* region has hindered the fine-scale analysis of the polymorphisms within the hCGβ coding genes and their effect on early pregnancy success until now.

DNA copy number variants (CNVs) involving DNA rearrangements of >50 bp in size have been proposed to explain heritability of various complex diseases due to large coverage of the human genome (up to 30%) (Redon et al. 2006; Mills et al. 2011). Only few studies have addressed the contribution of CNVs in the manifestation of reproductive disorders, including a single CNV screening focusing on placental samples of RM cases (Rajcan-Separovic et al. 2010). Genome-wide profiling of CNVs among parents affected by the disease may prove as a powerful tool in mapping novel genetic biomarkers affecting the predisposition to RM independently but also in consort.

In this thesis, the review of literature provides up to date information on the current knowledge of risk factors and genetic association studies of RM performed at fine-scale and global level. A comprehensive overview of the

structural and functional characteristics of hCG, its β-subunit and hCGβ coding genes is provided and the contribution of CNVs in common diseases and pregnancy complications is reviewed.

The experimental part of the thesis aimed to map the fine-scale and global determinants of genetic susceptibility to RM setting the genetic diversity of hCGβ genes and global CNV profile as the examples. The research focused on firstly, the local genomic landscape of *LHB/CGB* gene cluster and evolutionary factors shaping the diversity of *CGB* genes in the context of RM. Fine-scale analysis of polymorphisms in two major hCGβ genes *CGB5* and *CGB8* as promising RM candidate genes was performed in three North European populations and the structural and functional effect of three non-synonymous mutations identified within the genes was addressed. Secondly, the impact of CNVs on the manifestation of RM was elucidated at the genome-wide scale and novel independent genetic biomarkers of RM were inferred from the global CNV profile. An improved map of genetic determinants and current status of research on the etiology of RM is discussed in the light of the findings provided in this thesis.

1. REVIEW OF LITERATURE

1.1. Dynamics and challenges of human pregnancy

Human pregnancy is a unique intricate system that involves maternal acceptance of an organism foreign by the molecular and genetic build, establishment of a transient invasive organ placenta and a complex interplay of maternal and fetal signals balancing between the maternal needs and fulfilling the fetal requirements. Well-balanced communication and timely progression of early stages of pregnancy are crucial for achieving a normal full term birth.

Hemochorial placentation defined by the direct contact of fetal chorion with maternal bloodstream is characteristic to humans, higher primates, rabbit and rodents, however the advanced invasiveness of placentation is unique to humans (Enders and Carter 2004). Following the implantation of an embryo, the extraembryonic trophoblast cells proliferate and form syncytiotrophoblast cells covering the villous trees and give rise to non-proliferative extravillous trophoblast cells that adopt an invasive phenotype and induce remodeling of maternal uterine and vascular tissues (Norwitz et al. 2001; Red-Horse et al. 2004). As the pregnancy progresses, the increasing surface of the branching placental chorionic villi is in direct contact with maternal blood providing access to nutritional, gas and waste exchange in response to the growing needs of the developing fetus. Successfully formed maternal-fetal interface is able to adapt to the changing environment and effectively resolve the challenges of pregnancy. Inadequate implantation, shallow placentation or reaction to stress factors (e.g. oxidative stress) may give rise to infertility or pregnancy complications several of which are largely human-specific, such as preeclampsia and recurrent miscarriages (Norwitz et al. 2001; Jauniaux et al. 2006). Some disorders may persist even after parturition and have an adverse impact on the long term maternal health by invoking cardiovascular and autoimmune diseases, metabolic syndrome and early mortality (Clifton et al. 2012).

1.1.1. Prerequisites of early pregnancy success

Balance and correct timing of the multitude of pathways regulating the first post-conception weeks are essential for successful establishment and subsequent course of pregnancy. Low chance of natural conception per menstrual cycle (30%) has been observed in humans with 60% of pregnancy losses occurring already prior to clinical diagnosis of pregnancy highlighting the selective nature of the implantation and placentation stages (Zinaman et al. 1996; Macklon et al. 2002).

The timely progression of blastocyst implantation (within approximately 7 – 10 days after ovulation) is strictly regulated by a bi-directional interplay between maternal and blastocyst signals (Teklenburg et al. 2010). The induction of endometrial receptivity window when the uterus supports implantation, promotion of blastocyst attachment and differentiation of endometrial cells into

decidua cells are mediated via a complex signaling network involving ovarian hormones, cytokines, growth factors, adhesion molecules and transcription factors (e.g. progesterone, estrogen, HB-EGF, KLF5, MSX1/2, LIF) (Norwitz et al. 2001; Cha et al. 2012). Failed or delayed implantation may lead to in-/subfertility or early pregnancy loss, respectively, but also predispose to late obstetric complications such as preeclampsia (**Figure 1**) (Wilcox et al. 1999; Leach et al. 2002; Cha et al. 2012). Interestingly, it has recently been noted that unexpectedly large proportion of women affected by repeated early miscarriage (41% versus predicted 8%) are not sub- but super-fertile and achieve pregnancy within only couple of cycles (Salker et al. 2010). The phenomenon has been attributed to impaired natural embryo selection, whereby uterine environment allows the implantation of unviable embryos (Teklenburg et al. 2010).

The final outcome of human pregnancy largely depends on the invasion of cytotrophoblast cells to an adequate depth in the uterus accompanied by remodeling of uterine spiral arteries and transformation into low-resistance vascular network (Jauniaux et al. 2006). Limited endovascular invasion has been related to suboptimal blood flow, subsequent hypoxia and tissue damage leading to the manifestation of preeclampsia or intrauterine growth restriction (Kaufmann et al. 2003). On the other hand, excessive placental invasion and abnormal attachment to maternal myometrium (placenta accreta/increta) that may even reach maternal organs (placenta percreta) is a serious obstetric complication with a risk of severe post-partum haemorrhage (Jauniaux and Jurkovic 2012). The advanced invasiveness characteristic to human hemo-chorial placentation has been correlated with the evolution of hyper-glycosylated human chorionic gonadotropin (hCG-h) in primate lineage expressed by invasive extravillous trophoblasts in early pregnancy (Jauniaux et al. 2006; Cole 2009; Guibourdenche et al. 2010). Concurrently, placental insufficiency due to low expression of hCG-h has been implicated in both early and late pregnancy disorders, such as early pregnancy loss and preeclampsia (Kovalevskaya et al. 2002; Keikkala et al. 2013).

It has been shown that the largest proportion of genes upregulated in the human decidua cells in response to the paracrine signals of an implanting trophoblast are related to immune response at the feto-maternal interface (Hess et al. 2007). Maternal immune tolerance to the semi-allogeneic embryo expressing paternal antigens has been most extensively studied in normal pregnancy but also in relation to pregnancy complications (Guleria and Sayegh 2007; Redman and Sargent 2010). In this feto-maternal immune system communication, critical role has been attributed to uterine natural killer (NK) cells that constitute 40% of cells in the decidua ($CD56^{bright}CD16^{-}$ NK subset) and promote trophoblast invasion and angiogenesis (Moffett-King 2002; Hanna et al. 2006). Uterine NK cells express an array of killer immunoglobulin-like receptors (KIR) that mediate the trophoblast recognition via binding MHC class I molecules (HLA-E, HLA-C and HLA-G) expressed on extravillous trophoblasts. Distinct KIR receptor-ligand combinations have been associated

with increased risk of early pregnancy loss and preeclampsia (Hiby et al. 2004; Hiby et al. 2008; Faridi and Agrawal 2011).

Upon successfully passing the selection windows at implantation and meeting all the prerequisites and challenges of a pregnancy, 30% of conceptions are estimated to reach live birth (Macklon et al. 2002).

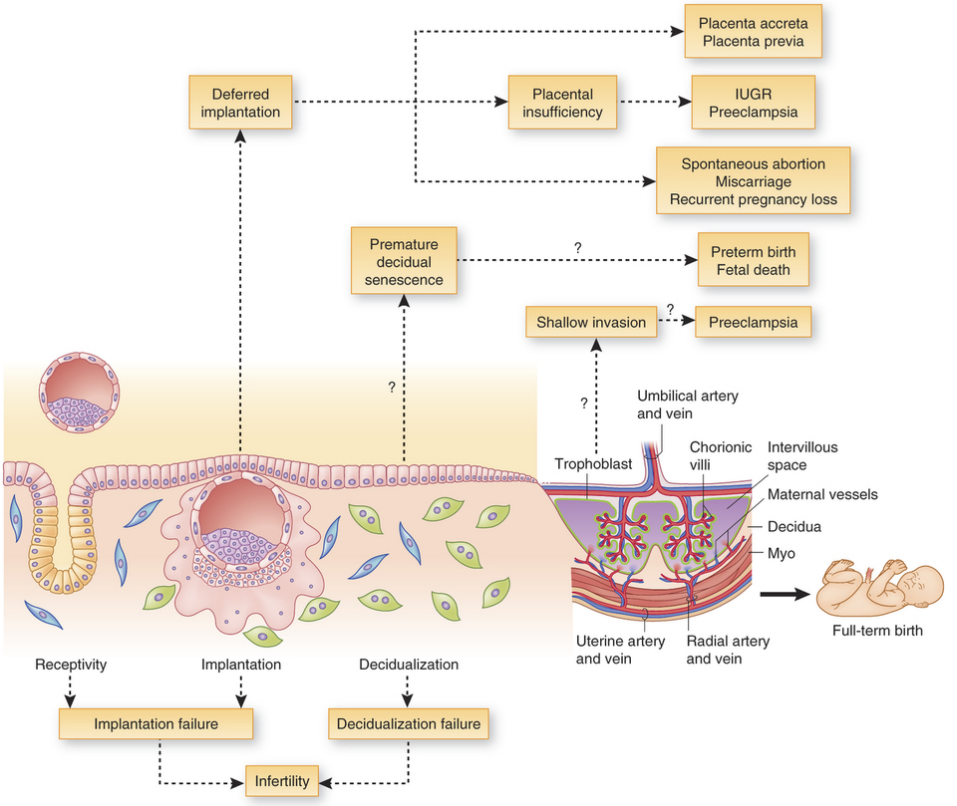

Figure 1. Cause-and-effect scenarios in human pregnancy disorders as proposed by (Cha et al. 2012). Deviations in the timing and signaling of early pregnancy stages may lead to in-/subfertility, early pregnancy loss or give rise to late pregnancy disorders due to subsequent placental insufficiency or shallow invasion. IUGR, intrauterine growth restriction.

1.1.2. Recurrent early pregnancy loss

A sporadic miscarriage is the most common pregnancy complication affecting approximately quarter of all women at least once in their lifetime and up to 15% of all clinically recognized pregnancies (Wilcox et al. 1988; Stirrat 1990). In 3% of fertile couples, three or more consecutive early pregnancy losses occur before gestational week 22, defined as recurrent miscarriage disease (RM) (Christiansen et al. 2008). In patients diagnosed with RM, each subsequent

pregnancy is accompanied with an increased risk of not only miscarriage but also other obstetric complications such as preterm birth or small for gestational age newborns (Ogasawara et al. 2000; Jivraj et al. 2001; van Oppenraaij et al. 2009). It is a distressing disease for the affected couples with approximately half of the cases remaining unsolved due to the complex etiology involving a spectrum of known but also as yet unidentified risk factors (reviewed in (Christiansen et al. 2008)).

1.1.2.1. The heterogeneity of RM etiology

The major cause of sporadic early miscarriages (in up to 75% of cases) is fetal chromosome abnormalities associated with continuously increasing age of women postponing childbearing to late 30s and early 40s in western countries (Philipp et al. 2003; Group 2010). Although high maternal age is also a risk factor in RM, other causes predominantly drive this disease as the chance of having an early pregnancy loss due to large chromosomal alterations is decreasing with an increasing number of miscarriages in a couple (**Figure 2**) (Ogasawara et al. 2000). The etiology of RM is heterogeneous and involves interplay between maternal, paternal and cumulative (placental/fetal) risk factors in pathways related to pregnancy establishment and maintenance.

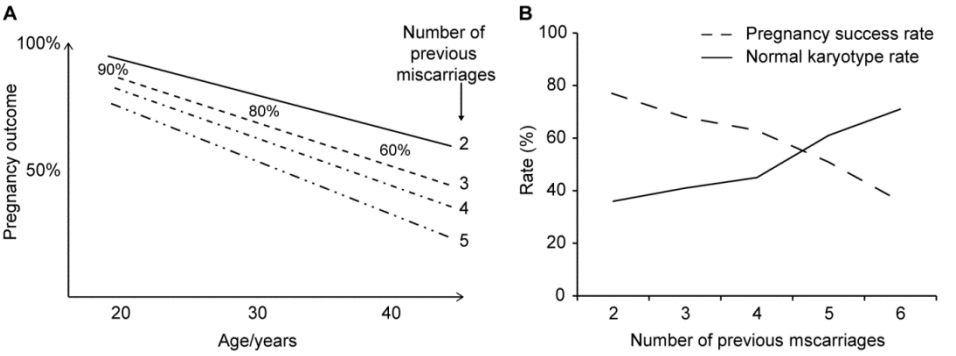

Figure 2. Rate of pregnancy success in women experiencing spontaneous miscarriages. (**A**) Pregnancy outcome depending on the maternal age in women with 2 or more miscarriages (adapted from (Matthiesen et al. 2012)). (**B**) Pregnancy success and the chance of observing fetal normal karyotype in women with 2 or more miscarriages and with average age of 31 years (based on (Ogasawara et al. 2000)).

The main factors studied in association with RM susceptibility include (reviewed in detail in (Pandey et al. 2005; Christiansen et al. 2006; Christiansen et al. 2008; Larsen et al. 2013))

- thrombophilic disturbances that contribute to thrombosis of placental vessels leading to placental infarctions
- excessive inflammatory processes causing increased apoptosis and pregnancy failure
- reproductive tract infections (e.g. HIV, malaria) triggering the inflammation of the uterine lining that disrupts embryo implantation and growth
- autoimmune (e.g. antiphospholipid syndrome) and alloimmune (couple's immunologic mismatch) factors leading to maternal rejection of the fetus
- anatomical disorders (e.g. uterine malformations)
- endocrine dysfunction (e.g. thyroid dysfunction)
- sperm DNA fragmentation introducing deleterious DNA damage to the developing embryo
- impaired embryo selection by maternal decidua
- oxidative stress damaging feto-maternal tissues

For many of the RM risk factors studied, the association with the disease is weak and/or the underlying mechanism has remained unclear (reviewed in (Christiansen et al. 2008; Larsen et al. 2013)) highlighting the need for further research to map and describe all the risk factors of this multifactorial disease

1.1.2.2. Defining the genetic component of RM (also reviewed in Ref. I)

Contribution of genetic susceptibility to the etiology of RM has been acknowledged and up to 2-fold higher occurrence of the disease has been reported within the affected families compared to control population (Christiansen et al. 1990; Kolte et al. 2011). Due to lower prevalence of large chromosomal abnormalities in miscarried embryos of RM patients compared to sporadic miscarriages (Ogasawara et al. 2000), other types of genetic variation (e.g. submicroscopic changes) likely contribute to the manifestation of the disease. Most studies addressing the genetic susceptibility to RM have used a candidate gene based approach by screening polymorphisms and mutations in genes with known functional effect on RM and/or pathways related to pregnancy establishment and maintenance. Currently, approximately 100 candidate genes have been targeted with the largest attention given to factors leading to thrombophilic, inflammatory and immunologic disturbances at the feto-maternal interface (Table 1 in Ref I).

Like other thrombophilic mutations, the most common variants G1691A in factor V (Leiden mutation) and G20210A in prothrombin, may contribute to the excessive blood coagulation and impaired placental blood circulation leading to

increased risk of early pregnancy loss in the mutation carriers (Dizon-Townson et al. 1997). Although both mutations are routinely tested in the clinical setting and have been positively associated with RM in most studies, the odds ratio (OR) ranges considerably (from 0.5 to 18) due to the heterogeneous nature of the phenotype and instead, a stronger association has been observed for late pregnancy loss (Rey et al. 2003; Kovalevsky et al. 2004; Robertson et al. 2006). Similarly, the genes regulating the balance between the action of pro-inflammatory (e.g. TNFα, IFNγ) and anti-inflammatory factors (e.g. IL6, IL10) at the feto-maternal interface have been studied extensively but have gained inconclusive results (Daher et al. 2003; Bombell and McGuire 2008). Due to weak or no impact of the studied polymorphisms, it has been proposed that instead of individual mutations, a combination and accumulation of thrombophilic and inflammation related genetic variants may shape the cumulative risk of RM (Rey et al. 2003; Jivraj et al. 2006).

As the fetal cells are expressing antigens of paternal origin, the maternal rejection of the 'foreign' fetus is potentially one of the major causes leading to miscarriage event (Wilczynski 2006). This condition of allograft rejection is bi-directional defined by maternal recognition and fetal presentation of paternal antigens and thus the genetic contribution of both counterparts has been addressed. For example, the allelic composition and presence of 14 bp insertion in the human leukocyte antigen HLA-G gene, the fetal genetic determinant expressed by extravillous cytotrophoblasts at the feto-maternal interface (Kovats et al. 1990), has been associated with the risk of RM, however the results are inconsistent (Aldrich et al. 2001; Hviid et al. 2004; Aruna et al. 2010; Wang et al. 2013). The cumulative effect of maternal recognition and presented fetal antigens has also been suggested as promising genetic factors describing susceptibility to RM, for example increased HLA-sharing between couples/ mother-fetus or specific interaction of killer immunoglobulin-like receptors (KIR) on maternal uterine NK cells and fetal HLA-C on the trophoblast cells. However, no clear consensus has also been reached for these findings (Witt et al. 2004; Beydoun and Saftlas 2005; Hiby et al. 2010; Kolte et al. 2010; Moghraby et al. 2010).

The inconclusive or controversial results of the numerous candidate gene-based SNP studies highlight the complexity of the studied phenotype and potentially low impact of single nucleotide variants in the manifestation of RM. It should be noted that many of the studies were also underpowered due to narrow study setup (only affected mothers) and insufficient number of study subjects to detect genetic variants of small effect size and/or with low population-specific prevalence (Ref. I). Thus, increasing the sample size, addressing both the maternal and paternal contribution to the disease and/or other types of variation in association with RM would prove beneficial.

1.1.2.3. Hypothesis-free 'omics' approach in screening for novel RM risk factors

Unlike candidate gene based studies, the powerful tools of genome/ transcriptome/proteome-wide analyses would enable the identification of novel genes and pathways essential in early pregnancy and modulating the risk to RM. In recent years, a number of studies have addressed various levels of variation shedding light on this heterogeneous phenotype.

Genomics

Two genome-wide studies have been reported searching for novel RM-associated genomic loci among the cases with unexplained RM. In a pilot study performed in unrelated Han Chinese patients and fertile controls, three micro-satellite markers at 6q27, 9q33.1 and Xp22.11 were significantly associated with RM (Li et al. 2010). At 9q33.1, two potential RM candidate genes were proposed, the *TNFSF8* and *TNFSF15* belonging to the tumor necrosis factor (TNF) ligand family implicated in early pregnancy loss in mice (Erlebacher et al. 2004). Four additional genomic regions (3p14.2, 6q16.3, 9p22.1 and 11q13.4) that did not overlap with the results by Li et al. (2010) were revealed in a linkage study undertaken in Danish sibling pairs affected by RM (Kolte et al. 2011). No SNP-based genome-wide association studies (GWAS) extensively used for identification of the genetic component of complex diseases have been performed for RM phenotype. The informative value of GWAS in the multifactorial RM remains unclear due to expected small effect sizes of individual genetic variants as also demonstrated for other reproductive traits (Montgomery et al. 2013).

Increasing evidence suggests that genomic imprinting mediating the expression of parent-specific alleles regulates the fetus- and placenta-specific genes and pathways (Nelissen et al. 2011; Novakovic and Saffery 2012). Aberrant gene methylation profiles have been linked to various obstetric disorders such as small for gestational age, intrauterine growth restriction, preeclampsia and gestational trophoblastic disease (Xue et al. 2004; Guo et al. 2008; Diplas et al. 2009; Yuen et al. 2009) but also recurrent miscarriage. In addition to one candidate-gene based study identifying aberrant methylation of hCG β-subunit coding *CGB5* gene in three cases with RM (Uuskula et al. 2010), a genome-scale methylation analysis of chromosomally normal miscarried chorionic villi from RM women has been reported (Hanna et al. 2013). Alternative global methylation patterns were observed among the RM cases when compared to normal chorionic villi from elective abortions and altered methylation levels were subsequently confirmed for distinct loci, such as *AXL* (*receptor tyrosine kinase*) and *DEFB1* (*defensin β 1*).

Transcriptomics

A whole transcriptome analysis of maternal decidua or placental/fetal tissues may provide novel insights on pathways and specific genes with critical functions in implantation and maintenance of pregnancy. Although the expression of pre-selected genes in chorionic villi (Baek et al. 2002; Choi et al. 2003) and endometrium (Lee et al. 2007) of RM cases has been addressed, only two studies have targeted the global transcriptome profile of these tissues. Differential gene expression analysis of the decidual tissue of women suffering from RM and undergoing a miscarriage confirmed the major contribution of pathways and genes related to immune response (23% of dysregulated genes) but also cell signaling (18%) and cell invasion (17.1%) in the reoccurrence of early pregnancy loss (Krieg et al. 2012). In the miscarried placental tissue of RM cases, significant overexpression of TNF-related apoptosis-inducing ligand (*TRAIL*) and inflammatory marker calprotectin (*S100A8*) were identified (Rull et al. 2013). The increased level of TRAIL protein has also been reported for maternal serum (Agostinis et al. 2012; Rull et al. 2013) and *S100A8* mRNA in the maternal decidua (Nair et al. 2013) of women affected by recurrent early pregnancy loss thus potentially representing novel biomarkers of RM disease.

Additional level of regulation in placental gene expression may be provided by short non-coding RNAs implicated in the post-transcriptional control of pathways related to male and female reproductive traits (Hawkins et al. 2011). Although not directly studied in recurrent miscarriages, differential microRNA (miRNA; ~22 nucleotides) profile has been reported for the cases of repeated implantation failure undergoing *in vitro* fertilization (IVF) procedures (Revel et al. 2011), indicating the potential contribution of miRNAs in the manifestation of early pregnancy loss.

Proteomics

In order to determine novel or confirm known RM risk factors at the proteome level, follicular fluid of three RM cases compared to three multiparous fertile controls was addressed with a combination of two-dimensional gel electrophoresis and mass spectrometry by (Kim et al. 2006). Aberrant expression of five proteins was observed in the follicular fluid of which the angiotensinogen, complement component C3c chain E and coagulation factors fibrinogen γ and antithrombin were subsequently also confirmed to be downregulated in the chorionic villi tissue samples thus potentially affecting embryo development or placental function. The strength of proteome analysis was further underlined when similar approach was applied for the maternal serum of RM women identifying an additional potential biomarker, the acute-phase inflammation related ITI-H4 (Kim et al. 2011).

1.2. Human chorionic gonadotropin, the 'pregnancy hormone'

One of the key factors essential in implantation and early pregnancy maintenance is the placental human chorionic gonadotropin also termed as the 'pregnancy hormone'. hCG is produced already by an 8-cell blastocyst prior to implantation (Lopata and Hay 1989) and upon reaching mother's circulation it is used as an early pregnancy biomarker in conventional pregnancy tests. The concentration of hCG doubles every two days until peaking at gestational weeks 9–11 and large inter-individual variation in the levels of hCG during pregnancy has been documented (Fig. 6B in Ref. II) (Hay 1988). Nevertheless, critically low amounts in maternal circulation have been related to adverse pregnancy outcome such as early spontaneous miscarriage or ectopic pregnancy (Korhonen et al. 1994; Rull and Laan 2005).

A variety of functions have been attributed to the pleiotropic hormone hCG since early pregnancy that include supporting progesterone production by corpus luteum until the luteo-placental shift and independent expression of progesterone by the placenta but also promoting angiogenesis, trophoblast invasiveness and decidualization of endometrium, stimulating fetal testicular testosterone production and regulating maternal immunotolerance (Huhtaniemi et al. 1977; Zygmunt et al. 2002; Kayisli et al. 2003; Guibourdenche et al. 2010; Kajihara et al. 2010; Tsampalas et al. 2010; Schumacher et al. 2013). Alternative functions and sites of expression at lower amounts have been reported for assembled hCG and/or free hCG subunits in normal non-trophoblastic tissues such as pituitary, seminal fluid or secretory endometrium (Hoermann et al. 1995; Berger et al. 2007; Zimmermann et al. 2012). Also, increased level of circulatory hCG/hCGβ production in non-pregnant organism is the marker of invasive tumour progression, including bladder and gastrointestinal cancers (reviewed in (Stenman et al. 2004)). The parallel effect of hCG in implantation/placentation and cancerogenesis is attributed to its ability to modulate pathways essential in both processes, e.g. promoting cell invasion, angiogenesis and escape of immune surveillance (reviewed in (Holtan et al. 2009)).

1.2.1. Structural characteristics of hCG

Human chorionic gonadotropin belongs to the family of heterodimeric glycosylated gonadotropins together with luteinizing hormone (LH) and follicle stimulating hormone (FSH), all formed by non-covalent association of common α-subunit and unique β-subunit that defines their functional properties and binding to specific receptor (Morgan et al. 1975; Pierce and Parsons 1981). Although LH and hCG act via the same ubiquitously expressed human luteinizing hormone/chorionic gonadotropin receptor (hLH/CGR) (reviewed in (Ascoli et al. 2002)), hCG is discerned by mainly placenta-specific expression (syncytiotrophoblast cells) and increased biopotency in induction of cAMP

signaling pathway (Casarini et al. 2012). Recently, stimulatory effects independent of hLH/CGR have also been proposed for hCG isoforms in angiogenesis and trophoblast invasion (Berndt et al. 2013; Lee et al. 2013).

The hCG α- (length 116 amino acids) and β-subunits (145 amino acids) form a similar tertiary structure determined by five cystine bonds in hCGα and six in hCGβ (Lapthorn et al. 1994). Both subunits are comprised of three hairpin loops held together by three disulfide bonds that form a characteristic cystine knot motif which is a highly conserved structural feature also found among growth factors such as transforming growth factor-β2, nerve growth factor and platelet-derived growth factor-BB (Murray-Rust et al. 1993; Lapthorn et al. 1994). Disruption of the cystine knot forming disulfide bonds in hCGβ (Cys38-Cys90, Cys34-Cys88 and Cys9-Cys57) *in vitro* has a detrimental effect on folding, assembly and subsequently function of the protein (Bedows et al. 1993; Mishra et al. 2003). Concordantly, no *in vivo* mutations affecting the production of hCG hormone have been characterized likely due to resulting compromised fertility and embryo's viability.

The stability of hCG is determined by two N-linked and four O-linked oligosaccharide chains specific to the hCGβ protein that prolong the hormone's half-life to >24 hours compared to 3–4 hours for FSH and only <1 h for LH (Table 1 in Ref. II) (Morgan et al. 1975; Lapthorn et al. 1994). A hyper-glycosylated form of hCG (hCG-h) with large complex glycan moieties has been described as the major form of hCG expressed by invasive extravillous cytotrophoblast during the first weeks of normal pregnancy and promoting early placentation process (Guibourdenche et al. 2010). Increased amounts of hCG-h have been detected in cases of testicular cancer, hydatidiform mole, choriocarcinomas and trisomy 21 (Elliott et al. 1997; Cole et al. 1999; Lempiainen et al. 2012), whereas reduced levels have been observed in early pregnancy loss and preeclampsia (Kovalevskaya et al. 2002; Keikkala et al. 2013).

1.2.2. Genomic context and evolution of hCGβ genes (also reviewed in Ref. II)

In humans, the hormone-specific β-subunit of hCG is encoded by four *CGB* genes (*CGB*, *CGB5*, *CGB7* and *CGB8*) with 97–99% DNA sequence identity that give rise to an almost identical hCGβ protein (98 – 100%) in the placenta (Bo and Boime 1992; Hollenberg et al. 1994). The *hCGβ* genes are co-located with the *LHβ* gene (*LHB*) and two putative pseudogenes (*CGB1* and *CGB2*) in tandem and inverted orientations within a common gene cluster (approximately 50 kb) at 19q13.32 (**Figure 3**) (Policastro et al., 1986).

The *CGB* genes have likely evolved by the expansion of the ancestral *LHB* gene duplicons in the primate lineage as consistently increasing number of *CGB* genes is evident among the primates today reaching five gene copies in chimpanzee and six in the human genome (**Figure 3**; Ref II) (Maston and Ruvolo 2002; Hallast et al. 2008). The duplicate *CGB* gene acquired a

frameshift mutation in the last exon elongating the protein by 24 amino acids (termed as the C-terminal extension) that introduced the attachment sites for novel 0-linked oligosaccharides and subsequently increased the half-life of CG (Maston and Ruvolo 2002). Nevertheless, the homology between human LHβ and CGβ proteins has remained at >80% (Pierce and Parsons 1981). Alterations in the promoter region transferred the expression of the novel gonadotropin from pituitary to placenta and from non-pregnant to pregnant organism (Hollenberg et al. 1994).

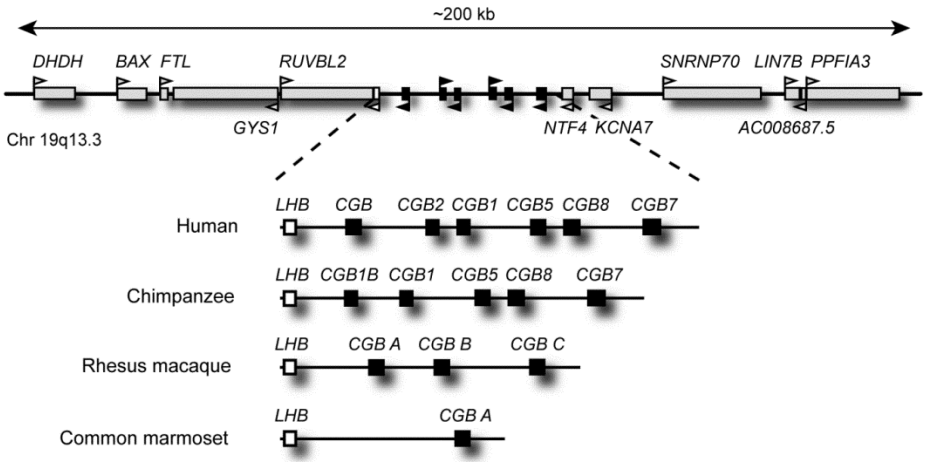

Figure 3. Genomic context of the *LHB/CGB* gene cluster in primates. The figure was drawn based on Ensembl database (http://www.ensembl.org/; Release 54). Boxes denote the genes and triangles above or below them point to the direction of transcription. The black boxes indicate *CGB*, white *LHB* and grey neighbouring genes. The *CGB* genes of rhesus macaque (*Macaca mulatta*) and common marmoset (*Callithrix jacchus*) are indicated as *CGB A-C*, since their ancestral status relative to the human *CGB* genes is unknown (reviewed in Henke and Gromoll 2008; Hallast et al. 2009) (Ref. II).

The *CGB1* and *CGB2* genes further diverged from the *CGβ* genes due to replacement of classical *CGB* 5' upstream region with a novel DNA sequence changing the open reading frame and promoting speculations on the existence of a novel protein that bears no resemblance to hCGβ (Bo and Boime 1992; Hallast et al. 2007). Although the expression of *CGB1/2* splice variants has been detected in placenta, testis and pituitary (Dirnhofer et al. 1996; Rull and Laan 2005; Parrott et al. 2011) and at increased levels in ovarian and breast cancer tissues (Giovangrandi et al. 2001; Kubiczak et al. 2013), only the classical hCGβ protein has been identified in testis possibly induced by the active transcription of *snaR-G* genes located in the novel inserted sequence upstream of *CGB1/2* (Parrott et al. 2011). The functional relevance of the

CGB1/2 genes and the existence of the hypothetical *CGB1/2* protein are currently under debate.

The appearance of *CGB* genes has been correlated with the evolution of increasingly invasive hemochorial placenta in primates as CG mediates the trophoblast invasion and placental anchoring into maternal tissues (Maston and Ruvolo 2002; Cole 2009). Concordantly, inadequate placental invasion and hCG production may lead to obstetric complications (e.g. pre-eclampsia, spontaneous miscarriage) largely unique to humans as the representatives of most advanced hemochorial placentation (Jauniaux et al. 2006).

1.2.3. hCGβ genes as candidate risk loci in pregnancy failure

The human *CGβ* genes exhibit a highly variable transcriptional activity in the placenta with *CGB5* and *CGB8* providing the largest contribution to the pool of hCGβ transcripts (up to 82%) (Bo and Boime 1992; Miller-Lindholm et al. 1997; Rull and Laan 2005). The cumulative expression level of the *hCGβ* genes during the course of pregnancy is in good correlation with the total amount of the hormone as the rate-limiting step in hCG production is the formation of the hCGβ protein (Fig. 6 in Ref. II) (Huth et al. 1992). The clinical significance of the *hCGβ* genes is indicated by the decreased placental expression among the cases of recurrent miscarriage (Rull and Laan 2005) or excessive expression in ectopic and molar pregnancy (Rull et al. 2008). The requirement of balanced biallelic transcription of maternal and paternal alleles has been observed for early pregnancy, whereas methylation allelic polymorphism patterns in the *CGB5* gene whereby paternal alleles have gained methylation has been associated with recurrent miscarriage (Uuskula et al. 2010).

Addressing the genetic variation of the *CGB* genes in relation with obstetric complications has been challenging due to high DNA sequence similarity between gene copies (up to 99%), imprecise map of gene-specific polymorphisms and likely compensatory effect of the multi-copy *hCGβ* genes. It is thus not surprising that no genetic variants with phenotypic effect have been reported in *CGB* genes until the studies reported in this thesis (also reviewed in Ref. II). Further research is needed to elucidate the genetic factors underlying the transcriptional variability of *CGB* genes seen in various pregnancy disorders but also cancer.

1.3. DNA copy number variation

Copy number variants (CNVs) (**Figure 4**) represent deletions, duplications or inversions of > 50 bp DNA sequence cumulatively covering up to 12–30% of human genome compared to 1% altered by SNPs (Redon et al. 2006; Zhang et al. 2009; Mills et al. 2011). Due to the large genomic coverage, CNVs introduce substantial dynamics and interindividual variation to the genome not only among humans but also other mammals, including chimpanzee, rhesus

macaque, dog, mouse and rat (Graubert et al. 2007; Kehrer-Sawatzki and Cooper 2007; Guryev et al. 2008; Lee et al. 2008; Conrad et al. 2010; Berglund et al. 2012), and in *Drosophila melanogaster* (Emerson et al. 2008). Recent technological advances such as next-generation sequencing have further improved fine-mapping of the CNV profiles in human genome facilitating the accurate copy number detection of even complex genomic regions (Alkan et al. 2009; Mills et al. 2011).

It has been estimated that CNVs disrupt around 13% of human RefSeq genes (McCarroll et al. 2008; Conrad et al. 2010) and approximately 18% of variability in gene expression may be attributed to these DNA variants (Stranger et al. 2007). Concordantly, a growing number of structural DNA variations have been implicated in benign phenotypic traits (e.g. human amylase gene variation) (Perry et al. 2007), rare genomic syndromes (e.g. Potocki-Lupski syndrome) (Potocki et al. 2000) and common disease susceptibility (Girirajan et al. 2011a). Both increased global burden of CNVs and distinct CNV loci have been linked to several complex disorders pointing to considerable contribution of these genetic variants in modulating individual's disease risk.

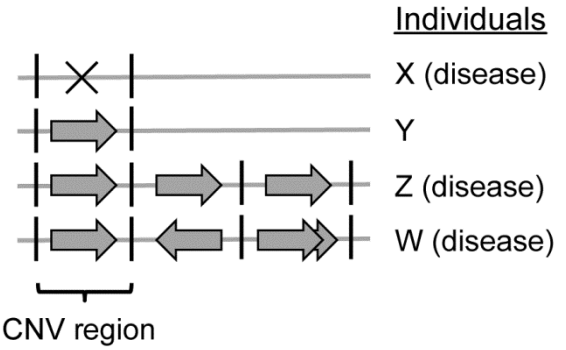

Figure 4. Schematic representation of a DNA copy number variation (CNV). In case a CNV region involves a dosage sensitive gene/regulatory region (arrow), deletion (individual X), tandem duplication (Z) or a more complex genomic rearrangement (W) may lead to a disease among the CNV carriers.

1.3.1. CNVs in common diseases

1.3.1.1. Hypothesis-free global genomic profiling of CNVs

Genome-wide screening for structural variations has been undertaken in an attempt to map the global 'risk profile' or determine particular categories of CNVs predisposing to common disease. Due to likely disease-conferring risk, rare CNVs have been preferentially addressed by numerous studies and several rare case-specific rearrangements have been implicated in the occurrence of a common disease. There is a publication bias towards the neuropsychiatric disorders likely due to the assembled consortia and access to large number of cases needed for increased statistical power of genome-wide analysis, particularly when targeting rare CNVs.

Increased burden of *de novo* and rare (<1%) CNVs among the affected individuals has been reported for several neuropsychiatric disorders, including autism (Pinto et al. 2010), schizophrenia (Consortium 2008; Walsh et al. 2008; Vrijenhoek et al. 2008; Vacic et al. 2011), bi-polar disorder (Malhotra et al. 2011) and a combined analysis of autism, intellectual disability and dyslexia (Girirajan et al. 2011a). Several novel candidate genes increasing the disease risk have been identified in these CNVs, including *VIPR2, NRXN1* and *SHANK2*. An enrichment of genes affecting neuronal signaling and development pathways has also been observed potentially suggesting that co-occurrence of individually rare CNVs may cumulatively increase the risk of a neurodevelopmental disorder (Walsh et al. 2008; Pinto et al. 2010). Concurrently, a two-hit model has been proposed, whereby at least two mutational hits of large (>500 kb) rare CNVs are needed for the manifestation of a neuropsychiatric disease (Girirajan et al. 2010; Girirajan et al. 2012).

In addition to neurodevelopmental disorders, genome-wide association studies involving both rare and common CNVs have been performed for complex diseases having a large impact on public health. In severe obesity cases, burden of long (>100 kb or >500 kb) rare CNVs was reported and significant enrichment of deletion CNVs affecting genes of G-protein coupled receptors was observed (Bochukova et al. 2010; Wheeler et al. 2013). A genome-wide analysis of all CNV classes in eight common diseases has been applied by The Wellcome Trust Case Control Consortium in altogether 16000 cases of bipolar disorder, breast cancer, coronary artery disease, Crohn's disease, hypertension, rheumatoid arthritis, type 1 diabetes and type 2 diabetes (Craddock et al. 2010). CNVs in or near altogether three loci were associated with Crohn's disease (*IRGM*), type 2 diabetes (*TSPAN8*) or with Crohn's disease, rheumatoid arthritis and type 1 diabetes (HLA region). The association of HLA region with type 1 diabetes was further confirmed by an independent study identifying 11 risk conferring CNVs, including a deletion near HLA-DQ allele (Grayson et al. 2010). Recently, two studies have reported the impact of CNVs on human longevity, whereby either burden of large (≥500 kb) common deletions or CNVs involving genes that affect RNA alternative splicing were associated with decreased lifespan (Kuningas et al. 2011; Glessner et al. 2013).

1.3.1.2. Targeted analysis of copy number variable genes

Targeted analysis of several distinct copy number variable loci has been performed in the context of common human diseases, majority of which are related to disturbances in the function of immune system and share the risk-associated loci.

The complex structural variation of duplicate genes encoding alpha- and beta-defensins (*DEFA* and *DEFB*) has been mapped in detail in human population (Hollox et al. 2003; Aldred et al. 2005). As the multi-functional defensin proteins have an important role in innate system, copy number

variation of these genes has been studied in complex disorders and increased or decreased copy numbers have been associated with psoriasis (Hollox et al. 2008; Stuart et al. 2012) and Crohn's disease (Fellermann et al. 2006), respectively.

Human Fcγ receptors are IgG-binding glycoproteins that are encoded by a duplicate gene family of *FCGR* genes and have been implicated in autoimmune and inflammatory diseases. Copy number variation of members of this gene family increases the risk of idiopathic thrombocytopenic purpura and systemic lupus erythematosus (Breunis et al. 2008;Fanciulli et al. 2007). The systemic lupus erythematosus disease has also been associated with CNVs involving complement component *C4* genes (Yang et al. 2007) and interestingly, an epistatic effect was found between *FCGR3B* and *CCL3L1* encoding for macrophage inflammatory protein (MIP)-1α influencing both systemic lupus erythematosus and rheumatoid arthritis (Mamtani et al. 2010). Susceptibility to human immunodeficiency virus (HIV) infection has been addressed with CNV studies although with contrasting results. Contribution of the gene dosage of *CCL3L1*, the human immunodeficiency virus-1 (HIV-1)-suppressive chemokine, has been debated (Gonzalez et al. 2005; Shao et al. 2007), whereas a positive association was observed for the relative amounts of activating and inhibitory KIR genes that control HIV-1 infection (Pelak et al. 2011). As the KIR gene family has been associated with a variety of human diseases and extensive copy number variability has been described in the region (Jiang et al. 2012; Kusnierczyk 2013), other traits depending on KIR gene dosage are expected to be reported.

As examples of non-immunomodulatory findings, triplication of trypsinogen gene (*PRSS1*) in patients of hereditary pancreatitis and recently, CNVs disrupting the *AUTS2* gene of unknown function among the cases of neuro-developmental disorders have been reported (Le Marechal et al. 2006; Nagamani et al. 2013) (reviewed in (Girirajan et al. 2011b)).

1.3.2. CNVs in pregnancy complications

In spite of the growing evidence of CNVs modulating the risk to various complex disorders, only few studies addressing the contribution of CNVs in pregnancy complications have been reported very recently.

A combined GWAS and CNV analysis has been performed among the cases with pre-eclampsia, the pregnancy-related hypertension, and in normotensive females (Zhao et al. 2012). Although only deletion CNV regions were considered as potentially the most deleterious type of rearrangement in the original study, three CNVs with increased prevalence among white female cases compared to controls were identified (**Table 1**). A 15 kb deletion encompassing the *PSG11* gene expressed during pregnancy by syncytiotrophoblasts was highlighted as potentially contributing to the occurrence of preeclampsia. The findings were not subsequently replicated among preeclamptic and normoten-

sive Afro-Caribbean, Hispanic and European ancestry mothers, however, a number of novel candidate rearrangements were reported (**Table 1**) (Zhao et al. 2013). Neither of the studies identified SNP associations of genome-wide significance likely due to the small number of analyzed cases.

Single reports for recurrent miscarriage and stillbirth have been published, both addressing a small number of cases (n = 27 and 29, respectively) and largely relying on the reference dataset of common CNVs in the Database of Genomic Variants (DGV) in order to infer the potential impact of identified rearrangements on the disease (**Table 1**) (Ledig et al. 2010; Rajcan-Separovic et al. 2010; Harris et al. 2011). CNVs involving the *PAPPA* and *HLA-DPA1* genes were implicated in the occurrence of stillbirths, whereas in recurrent miscarriage, two imprinted placental genes rearranged in the placental samples, *TIMP2* and *CTNNA3,* were highlighted as novel potential candidates increasing the risk of early pregnancy loss (Rajcan-Separovic et al. 2010; Harris et al. 2011).

Like candidate gene-based studies, CNV analyses are sensitive to sample size, ethnic background and adequate definition of phenotype of the study subjects and thus, most of the currently reported studies on pregnancy complications are likely to underestimate the number and effect of CNVs modulating the risk to disease. Also, none of the published studies have addressed the couple's cumulative susceptibility to disease, although both mother and father contribute to the fetal/placental genome and function (**Table 1**). Further research is needed to elaborate on the findings and assess the contribution of CNVs in manifestation of pregnancy complications.

Table 1. CNV studies in pregnancy complications.

Phenotype (Reference)	Study subjects	No of cases/ controls	Screening platform	No of CNVs	CNV size (kb)	Rearranged genes
Preeclampsia (Zhao et al. 2012)	White mothers	177/ 116	Affymetrix Genome-Wide Human SNP Array 6.0	3[a]	15–41	*PDXDC1, PSG11*
Preeclampsia (Zhao et al. 2013)	Afro-Caribbean mothers	21/ 1049	Illumina Human1M-Duo	8	4–89	*GPR39, HCG26, HCG2P7, HCG4P6, HCP5, MAGI2, MICA, UGT2B10*
	European ancestry mothers	50/ 1207	Illumina Human610-Quad	3	14–39	*AADAC*
	Hispanic mothers	62/ 661	Illumina Human1M-Duo	11	11–373	*DTX2, FRRS1, GOLPH3, HCG4P6, PDZD2, POMZP3, TXNRD2, USP37*
Stillbirth (Harris et al. 2011)	Placental samples	29/ 10; DGV[b]	Illumina CNV370-Duo	24	9–2886	*ADD2, AGO2, AUH, C9ORF47, CADM2, CHRAC1, CKS2, CLLU1, CLLU1 OS, CYP2C18, CYP2C19, DIRAS2, GADD45G, GP6, HLA-DPA1, MIR3153, MIR4290, MRPS30, NFIL3, OR4A47, OR4B1, OR4C3, OR4C45, OR4S1, OR4X1, OR4X2, OR51F2, OR51S1, OR51T1, OR52R1, PAPP-A, PTK2, PTPRJ, RDH13, S1PR3, SECISBP2, SEMA4D, SHC3, SYK, TGFA, TRAPPC9, TRRAP*
Recurrent miscarriage (Rajcan-Separovic et al. 2010)	Placental samples	27/ DGV[b]	Array-CGH	13	27–1593	*C7orf36, CSS3, CTNNA3, GPM6B, HDHD1A, LIPC, NDUFA12L, NPAS3, OFD1, PARK2, PNPLA4, POU6F, PRMT3, RAB9A, RALA, STS, TIMP2, TRPPC2, VCX, 13 ZNF family genes*

[a] Only deletions were considered in the study

[b] Database of Genomic Variants (DGV) (http://genome-euro.ucsc.edu/cgi-bin/hgGateway) was used as a reference dataset in addition to control samples if available to exclude potentially benign and common CNVs from the study

2. AIMS OF THE PRESENT STUDY

The present thesis aimed to elucidate contribution of human genome variation at global and fine-scale level in modulating the genetic etiology of recurrent miscarriage (RM). The main research aims of the thesis were as follows:

I. Fine-scale genome dynamics of duplicated *LHB/CGB* genes essential in fertility and early pregnancy maintenance (Ref. III)
   - Investigation of structural features and recombination patterns shaping the genomic landscape of the *LHB/CGB* gene cluster

II. DNA single nucleotide variants in the hCG β-subunit coding genes *CGB5* and *CGB8* in the context of RM disease
   - Screening for and association study of single nucleotide variants in the *CGB5* and *CGB8* genes in three North European populations (Ref. IV, V)
   - Testing the impact of identified missense mutations on the structural and functional features of hCGβ and hCG *in vitro* (Ref. VI)

III. DNA copy number variants (CNVs) as novel genetic determinants of recurrent miscarriage (Ref. VII)
   - Investigation of genome-wide CNV profile and its effect on functional pathways in RM cases
   - Identification of novel common CNV regions conferring risk to RM

3. RESULTS

3.1. Genome dynamics of human *LHB/CGB* gene cluster (Ref. II, III)

Rationale of the study:
The expression of the heterodimeric human chorionic gonadotropin (hCG) hormone in the placenta is of critical importance to the maintenance of early pregnancy. Due to the extensive sequence complexity and lack of detailed information on the polymorphisms of the gene family encoding the β-subunit of the hormone (*CGB* genes) the region has proven difficult to target in genomic approaches. This study was undertaken to describe the nature and origin of the diversity of the *LHB/CGB* gene cluster in order to promote further research of this region in the context of pregnancy success.

3.1.1. Genomic structure of the *LHB/CGB* gene cluster

To fine-map the genomic structure of the *LHB/CGB* gene cluster *in silico*, the sequence obtained from the NCBI GenBank database (http://www.ncbi.nlm.nih.gov; locus no NG_000019; June 26, 2002 release) was analyzed using web-based global alignment tools CLUSTALW (http://www.ebi.ac.uk/clustalw/) and BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) (Ref. III). The human *LHB/CGB* gene cluster includes the *LHB* gene (*luteinizing hormone beta polypeptide*; length 1.1 kb) encoding the LHβ protein expressed in pituitary, putative pseudogenes *CGB1* and *CGB2* (1.4 kb) and six *CGB* (c*horionic gonadotropin, beta polypeptide*) genes encoding the hCGβ subunit primarily expressed in placenta – *CGB, CGB5, CGB7* and *CGB8* (all 1.5 kb). The region is characterized by not only extensive DNA sequence identity between the genes (97–99% among the *hCGβ* genes, 92–93% when compared to *LHB*, 85% when compared to *CGB1* and *CGB2*) but also between the intergenic regions (up to 97%) (**Figure 5**). The gene cluster is closely bordered by functionally unrelated genes *RuvB-like 2* (*RUVBL2*) and *neurotrophin 5* (*NTF5;* now known as *NTF4)*.

The structural features currently seen among humans have likely arisen via several steps of duplication events involving the ancestral *LHB* giving rise to the *CGB* genes and part of the *NTF5* gene that has spread into the intergenic regions (**Figure 5**). The duplicated intergenic *Alu*-rich fragment originating from *NTF5* included the consensus *Escherichia coli* χ-sequence (GCTGGTGG) that has been associated with increased recombination and gene conversion activity (Smith 1988). The presence of χ-sites together with high content of *Alu* repetitive sequences in the intergenic regions (currently in humans 10–56%) may have cumulatively promoted the extensive rearrangement events in the primate lineage. The concept of duplication series in the evolution of this gene cluster is supported by the identification of various number of *CGB* genes in

other primates, such as chimpanzee (n = 5) (Hallast et al. 2008), rhesus macaque (n = 4) and common marmoset (n = 1) (**Figure 3**).

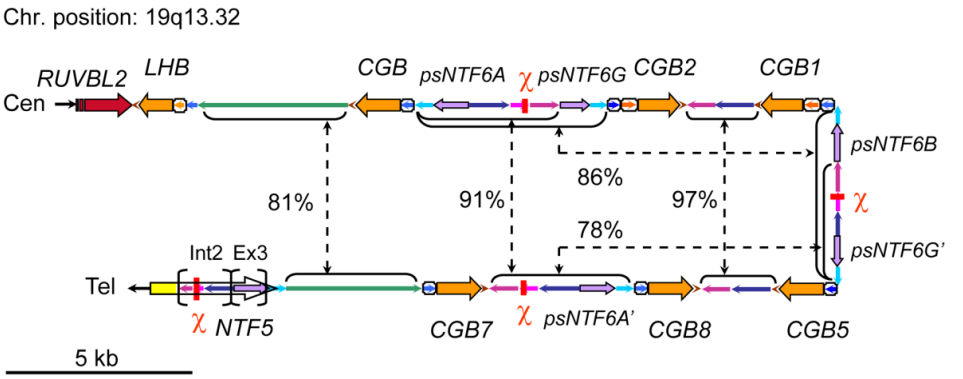

Figure 5. Structure of the human *LHB*/*CGB* gene cluster. Identical color codes refer to highly homologous DNA sequences within the cluster. Genes (*RUVBL2*, *LHB*/*CGB*s, *NTF5*) are depicted as wide arrows in the direction of transcription with promoters as boxed arrows 5' of the genes. Intergenic areas are marked as narrow lines, indicating also the direction of a segment, and broken arrows unite intergenic regions that share >75% of sequence identity. χ denotes the localization of χ-sequence. Int2 and Ex3 refer to intron 2 and exon3 in the *NTF5* gene; the former is the source of intergenic regions between *CGB* genes, and the latter has given rise to *NTF6* pseudogenes.

3.1.2. Polymorphism density and gene conversion activity

The diversity of the *LHB/CGB* genes was addressed by resequencing all genes except for *CGB8* (inaccessible due to sequence complexity) in three populations: Estonians (n = 47), African Mandenka (n = 23) and Chinese Han (n = 25). In total, 191 SNPs where identified with only a small fraction represented in the dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) (Supplementary Table S1 in Ref. III). Although the level of diversity varied among the populations analyzed, the highest polymorphism density was mutually observed for the genes at the edges of the gene cluster (*LHB*, *CGB* and *CGB7*) (Figure 2 in Ref. II; Table 1 in Ref. III). The peripheral genes *LHB*, *CGB* and *CGB7* were also characterized by strong linkage disequilibrium (LD), whereas breakdown of LD was observed toward the center of the gene cluster (*CGB1*, *CGB2*, *CGB5*) (for Estonians, **Figure 6A**; Figure 3 in Ref. III). Two major gaps in LD were noted in all populations (between the *CGB5* and *CGB8* genes and between *CGB* and *CGB2*) potentially pointing to higher recombination activity in these regions.

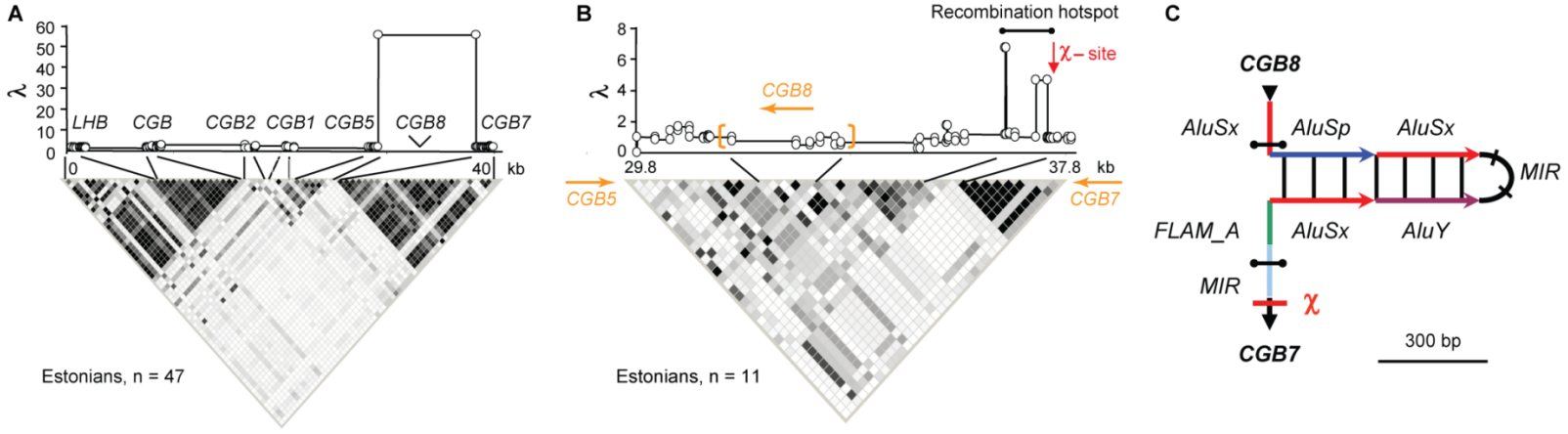

Figure 6. Linkage disequilibrium (LD) and recombination activity in the human *LHB/CGB* region and structural features of the potential recombination 'hot spot'. LD and recombination activity of the whole gene cluster were estimated in 47 Estonian population samples (**A**) and of the *CGB5-CGB7* region (**B**) in 11 Estonian samples. LD analysis was based on the r^2 statistic measuring the correlation between alleles using Haploview software (Barrett et al. 2005) and factor λ was calculated using PHASE 2.1 software (Li and Stephens 2003), both excluding SNPs with minor allele frequency <10%. (**A, B**) The regions of weak (white) and strong LD (black/dark gray) overlap with the occurrence of recombination 'warm/hot spots' and low recombination rate, respectively. (**B**) The recombination hotspot was identified between the *CGB8* and *CGB7* genes next to the recombination-associated χ-site (denoted in red). Locations of genes are marked with orange brackets (*CGB8*) or arrows denoting the direction of transcription (*CGB5*, *CGB7*). (**C**) Prediction of single-stranded DNA secondary structure for the estimated recombination 'hot spot' (bordered by black brackets) between *CGB8* and *CGB7*. The DNA structural analysis was performed with EMBOSS einverted software (http://emboss.sourceforge.net/) (Rice et al. 2000). The hotspot is flanked by double inverted *Alu*-sequences forming the stem (625 bp) of the palindrome, and its center falls within the loop (222 bp). (*MIR*) Mammalian-wide interspersed repeat; (*FLAM_A*) *Alu*-element-like repeat.

The high sequence similarity between duplicated regions in the *LHB/CGB* cluster is the rich substrate for the phenomenon of gene conversion whereby DNA sequence information is transferred between a pair of highly identical sequences via the process of homologous recombination (Figure 4A in Ref. II). A directional transfer of gene conversion tracts was observed from the central *CGB* gene copies to the distal genes with the highest number of acceptor sites identified for *CGB* (n = 8) and *CGB7* (n = 7) (Figure 2 in Ref. III) that also exhibited the highest polymorphism density and strongest LD. Gene conversion mechanism is likely the major force in shaping the diversity patters of the *LHB/CGB* region and spreading mutations that may modulate the function of the involved genes.

3.1.3. Recombination rate and structural analysis of potential recombination 'hot spot'

The recombination rate of the *LHB/CGB* gene cluster was assessed and recombination 'hot spots' were identified from unphased genotype data based on 'λ' that estimates by which factor the recombination rate between the loci exceeds the average background rate (Li and Stephens 2003). The value $\lambda = 1$ corresponds to the absence and $\lambda > 1$ to increased recombination activity. In all populations studied, potential recombination 'warm spots' and 'hot spots' were identified in the regions overlapping with the sharp LD breakdown observed in the *LHB/CGB* cluster (Figure 3 in Ref. III**)**. Among Estonians, the factor λ reached 2.36 for the potential recombination 'warm spot' between the *CGB* and *CGB2* genes and 57.1 for the 'hot spot' between *CGB5* and *CGB7* (**Figure 6A**).

In order to refine the location of the potential recombination 'hot spot' in the region between the *CGB5* and *CGB7* genes that includes *CGB8* and has remained inaccessible to *CGB8*-specific PCR and sequencing methods due to high sequence complexity, I conducted a pilot study by applying a combination of long-range (8.3 kb) and nested PCRs and subsequent resequencing in 11 Estonian populations samples. The analysis confirmed the weak LD in the region and fine-mapped the location of the recombination 'hot spot' in a <1 kb region between the *CGB8* and *CGB7* genes, embedded within an *Alu*-rich (~75% *Alu*-sequences) segment and 90–100 bp from the recombination-associated χ – sequence (**Figure 6B**). The DNA structural analysis of the 'hot spot' identified palindromic *Alu* repeats in the region that could give rise to a stem-loop secondary structure with a 625 bp stem and single-stranded loop segment (222 bp) in the middle (**Figure 6C**) potentially promoting double-strand breaks and subsequent recombination/gene conversion activity. Inverted repeats were also identified in the recombination 'warm spot' region between the *CGB* and *CGB2* genes, however with a much longer spacer (2788 bp) which might lead to decreased stability of the secondary structure and decreased recombination rate.

3.2. DNA single nucleotide variants of *CGB5* and *CGB8* in the context of recurrent miscarriage

3.2.1. SNP profile of the *CGB5* and *CGB8* genes in recurrent miscarriage (Ref. IV, V)

Rationale of the study:

Due to the irreplaceable role of hCG at the early stage of pregnancy, *CGB* genes encoding the hCG β-subunit are regarded as potential candidate genes for studying the genetic etiology of early pregnancy complications, such as recurrent miscarriage. The *CGB5* and *CGB8* genes that cumulatively provide up to 82% of hCG β-subunit transcripts in the human placenta (Miller-Lindholm et al. 1997; Rull and Laan 2005) likely have the largest effect on the quantity and quality of the hCG produced and are thus best candidates for addressing the role of this gene family in recurrent miscarriage.

3.2.1.1. Screening for variants in Northern Europe

The lack of comprehensive list of gene-specific polymorphisms for highly similar *CGB5* and *CGB8* (>92%, Ref. III) in the SNP databases prompted the full resequencing of these genes in RM case-control subjects from Northern Europe (Ref. IV). The 5' upstream (up to –435 bp from the start site of mRNA sequence) and genic region (down to +1082 bp relative to the mRNA start site) of *CGB5* and *CGB8* were resequenced in 184 female RM cases and their male partners (≥3 consecutive miscarriages before gestational week 22; Estonia n = 99, Finland n = 85) and 195 fertile women as controls (at least one live birth in Finland or three in Estonia and no previous miscarriages; Estonia n = 95, Finland n = 100). In total, 71 SNPs were identified (49 in *CGB5*, 22 in *CGB8;* Table 1 in Ref. IV) with nearly absent LD between them in both populations (Figure 2 in Ref. IV) that is largely concordant with the LD breakdown in the middle part of the *LHB/CGB* cluster reported in Ref. III.

3.2.1.2. Disease-related variants and haplotypes in *CGB8*

The SNP profile analysis of *CGB8* highlighted the potential functional relevance of the 5' upstream region of the gene. The region involved only three SNPs (Table 1 in Ref. IV) compared to the 18 SNPs in the respective DNA stretch of *CGB5* and was predicted to evolve under stronger functional constraints. Two neutrality tests were applied to explore observed versus expected distribution of SNPs and haplotypes – the Tajima's D (the difference between observed (π) and expected (θ) diversity estimates) and Ewens-Watterson homozygosity estimations (tests the observed allele frequency spectrum with the expected allele frequency spectrum under the neutral model (Hardy-Weinberg Equilibrium)). Both, the Tajima's D statistic (D^T = 2.29,

P <0.05) as well as Ewens-Watterson homozygosity test (P = 0.007) (**Table 2)** indicated a possible scenario of balancing selection driving the three apparently most efficient *CGB8* promoter variants (H1, H3, H4 in **Figure 7A**) to high frequency in both populations (Supplementary Table S2 and Supplementary Fig. S2 in Ref. IV).

Table 2. Neutrality tests of the *CGB5* and *CGB8* genes in fertile women and RM cases.

	CGB5			*CGB8*		
	Full region[a]	5'upstream region[b]	Genic region[c]	Full region[a]	5'upstream region[b]	Genic region[c]
No of SNPs	46	17	29	23	3	20
Fertile women						
Tajima D[d]	–1.12980	–1.16944	–0.88132	–0.35277	**2.29389***	–0.96141
P-value of Ewens-Watterson *F*[e]	ns	ns	ns	ns	**0.007**	ns
RM Patients						
Tajima D[d]	–1.23744	–1.23550	–0.91052	–0.56306	1.26879	–1.06361
P-value of Ewens-Watterson *F*[e]	ns	ns	ns	ns	ns	ns

[a]SNPs in 5'upstream and genic regions; [b]SNPs located in the region of –435 bp up to –1 bp from the start site of mRNA sequence; [c]SNPs located in mRNA seqence: +1 bp up to +1082 bp from the start site of mRNA sequence; [d]The basis of the Tajima's D statistics ($D^{T)}$ is the difference between observed (π) and expected (θ) diversity estimates: under neutral conditions $\pi = \theta$ and $D^T = 0$; [e]This statistic tests the observed allele frequency spectrum with the expected allele frequency spectrum under the neutral model (Hardy-Weinberg Equilibrium).
*P = 0.0169; ns – non-significant (P > 0.05)

Only one rare promoter variant H2 was identified in the 5' upstream region defined by the minor allele A of SNP c8EF-4 (position relative to transcription start site) that was solely present in RM patients, one from Finland and two from Estonia (Cochran-Armitage trend test, P = 0.071; **Figure 7A**; Table 1 in Ref. IV). This polymorphism may potentially affect the expression of the *CGB8* gene as it is located (**i**) within the AP1-like sequence overlapping the *hCGβ* initiator element critical for basal transcription and (**ii**) downstream of the Ets-2 binding site acting as a major enhancer of *hCGβ* gene expression (Ghosh et al. 2003).

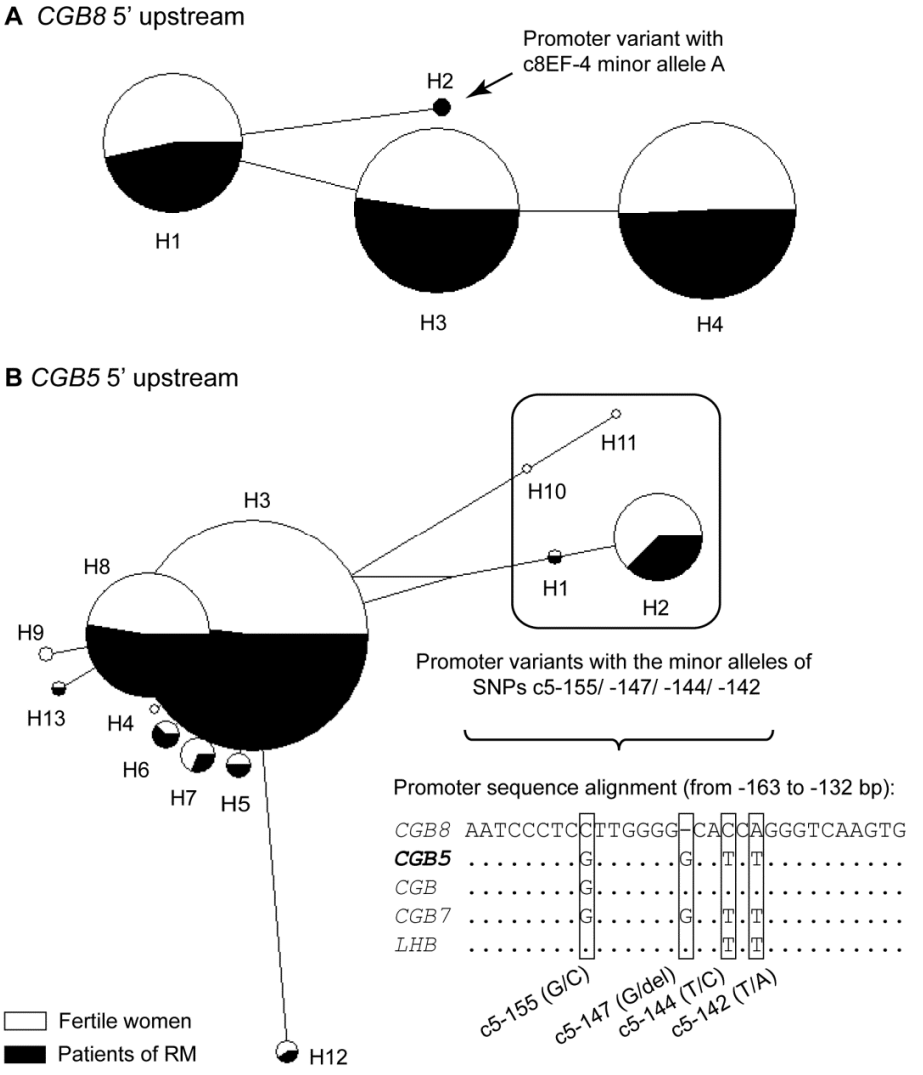

Figure 7. Haplotype networks of predicted promoter variants in the 5' upstream region of *CGB8* (**A**) and *CGB5* (**B**). Promoter variants were inferred from unphased genotype data using the Bayesian statistical method in the program PHASE 2.1.1. (Stephens et al. 2001) and networks of inferred haplotypes were drawn with NEWTWORK 4.201 software using the Median-Joining network algorithm (Bandelt et al. 1999). Singleton polymorphisms were excluded from the haplotype prediction analysis because of unreliable phasing. The size of each node is proportional to the haplotype frequency in the total dataset. The relative distribution of each haplotype among the RM cases (black) and fertile controls (white) is indicated. Haplotype nomenclature is shown in Supplementary Table S2 in Ref. IV. (**A**) Haplotype H2 defined by the minor allele of a proximal promoter mutation c8EF-4 was exclusively identified among RM patients in both Estonia and Finland. (**B**) The clade with haplotypes H1, H2, H10 and H11 carry the combination of minor alleles of four SNPs (c5-155C/c5-147del/c5-144C/c5-142A) that originates from the *CGB8* gene and has a protective effect against RM. The locations of the four SNPs are indicated in the alignment of consensus DNA sequence of the *LHβ* and *hCGβ* genes. Positions are given according to the transcription start site of *CGB5*.

Interestingly, the haplotype combining the minor alleles of c8EF-287 and c8EF-186 was absent in the current dataset in spite of the relatively high minor allele frequencies (MAF) of the SNPs (25.2% and 39.7%, respectively). The discrepancy between observed (0%) and expected (10%) haplotype frequency may be explained by the localization of these SNPs within the binding sites of Sp1/AP-2 transcription factors residing in the *hCGβ* gene region critical for the trophoblast-specific expression as well as cAMP-responsiveness of the transcription (Albanese et al. 1991). Considering also that the *CGB8* gene contributes most to the pool of hCGβ transcripts in the placenta (40%) (Rull and Laan 2005) it is thus likely that *CGB8* is harboring the most optimally functioning promoter sequence of the *hCGβ* genes.

3.2.1.3. Disease-related variants and haplotypes in *CGB5*

The SNP profile of 5' upstream region of *CGB5* (in total, 18 SNPs) involved a motif of minor alleles of four polymorphisms c5-155G/C, c5-147G/del, c5-144T/C and c5-142T/A that was completely identical to the homologous region in the *CGB8* gene exhibiting no genetic variation at these positions (**Figure 7B**). This DNA stretch (including c5-155C/c5-147del/c5-144C/c5-142A) probably originates from the *CGB8* gene via a meiotic gene conversion event between the two promoter regions. The haplotypes carrying this *CGB8*-specific DNA sequence (H1, H2, H10, H11) formed a separate clade on the haplotype network of the *CGB5* 5' upstream region (**Figure 7B**). It was speculated that in some pregnancies with impaired trophoblast growth (due to genetic, thrombophilic, immunological or other reasons), the placenta with the most efficient *CGB5* promoter haplotype (originating from and identical to the efficient *CGB8*) may have a better capacity for extra hCG production that may eventually rescue the threatened fetuses.

Concordantly, the minor allele frequency of the four *CGB5* SNPs was higher in fertile women (12.05–13.08%) compared to RM group (7.10–7.92%) in Estonian and Finnish populations combined and the association testing with the occurrence of RM identified a modest but significant protective effect for these SNPs ($P < 0.025$; OR = 0.54–0.58; Table 3 in Ref. IV). The results were successfully replicated in a meta-analysis study across the Estonian, Finnish and an independent RM case-control sample set from Denmark (RM cases with three or more consecutive miscarriages, n = 450; fertile women with at least two normal pregnancies and no miscarriages, n = 119) ($P = 0.021$; Table 2 in Ref. V). A strong protective effect against RM was also observed for a genic SNP c5EF1038 in *CGB5* intron 2 (Cochran-Armitage trend test $P < 0.007$; OR = 0.53 [95% CI 0.32–0.85]) with the frequency of 14.36% in fertile women compared to 8.15% in the RM group of Estonian and Finnish combined sample (Table 3 in Ref. IV). However, the allele frequency of this polymorphism did not differ between the RM cases and controls in the Danish replication sample

(MAF, 7.14% versus 7.42%, respectively; Cochran-Armitage trend test $P = 0.52$) and was not associated with RM (Ref. V).

3.2.1.4. Rare non-synonymous mutations in *CGB5* and *CGB8*

In addition to the spectrum of SNPs described in the *CGB5* and *CGB8* genes, four rare mutations leading to non-synonymous amino acid changes in the hCGβ protein were identified: *CGB5* p.Val56Leu in a single Finnish RM patient, p.Arg8Trp and p.Pro73Arg substitutions in the *CGB8* gene of single Estonian patients, and *CGB8* p.Val29Ile in one Finnish patient, two Estonian patients and seven Estonian fertile women (positions in the mature hCGβ protein; Table 1 in Ref. III). The three missense mutations (*CGB5* p.Val56Leu, CGB8 p.Arg8Trp and p.Pro73Arg) exclusively present among the cases were speculated to have impact on the production of hCG β-subunit protein in placenta and thus re-occurrence of miscarriages.

3.2.2. Structural and functional impact of non-synonymous mutations in *CGB5* and *CGB8* (Ref. V)

Rationale of the study:

In spite of the functional relevance of hCGβ in the establishment of pregnancy, the studies on genetic variation of the *CGB* genes and hCGβ protein has been challenging due to a high sequence similarity between the gene copies as well as the hCGβ proteins coded by these genes (98–100%) (Ref. II, III). Thus, only one naturally occurring variant of hCGβ (*CGB5* p.Val79Met) has been functionally characterized leading to inefficient hCG assembly *in vitro* (Miller-Lindholm et al. 1999). To assess the impact of the novel missense mutations identified by our screening study exclusively among the Estonian and Finnish RM cases (p.Arg8Trp, p.Val56Leu and p.Pro73Arg; Ref. IV), I conducted series of experiments addressing the structural and functional features of recombinant hCGβ proteins carrying these amino acid changes (performed in I.T. Huhtaniemi's lab, Imperial College London, UK; Ref. VI).

3.2.2.1. Prevalence and positional context of mutations p.Arg8Trp, p.Val56Leu and p.Pro73Arg

The prevalence of the rare mutations p.Arg8Trp, p.Val56Leu and p.Pro73Arg was estimated in a combined sample set of the discovery subjects from Estonia and Finland (described in Ref. IV) and replication sample from Denmark (described in Ref. V) (in total, 655 RM cases and 431 fertile female controls). Two of the studied substitutions *CGB5* p.Val56Leu (rs72556325, g.1178G>C; located in exon 3, **Figure 8A**) and *CGB8* p.Arg8Trp (rs72556341, g.806C>T; exon 2; **Figure 8A**) were each identified in a single heterozygous RM patient,

whereas the *CGB8* p.Pro73Arg mutation (rs72556345, g.1237C>G; exon 3; **Figure 8A**) was identified with a similar carrier frequency (0.46%) among RM (3/655) and control (2/431) individuals in the full screened Northern-European sample set (Estonian, Finnish, Danish, n = 1086) (Table 2 in Ref. VI).

All three hCGβ missense mutations under study are located immediately next to disulfide bond forming cysteins (Cys9, Cys57 and Cys72; Figure 1A in Ref. VI). Positions Arg8 and Val56 are involved in the central cystine knot structure essential for hCGβ folding and heterodimer assembly, whereas Pro73 is incorporated in a stable turn of the protein loop 3 that does not directly associate with the hCG α-subunit or the human luteinizing hormone/chorionic gonadotropin receptor (hLH/CGR) (**Figure 8B**). Unlike other targeted missense mutations, Val56 was found as fully conserved among hCGβ homologs from mammals to fishes (Figure 1C in Ref. VI) and was largely buried within the hCGα/β heterodimer complex with only 3% exposed to the solvent compared to 46% in case of the unassembled β-monomer based on Solvent Accessible Surface calculations (Figure 2 in Ref. VI). Therefore out of the three mutations under study, amino acid substitutions at position Val56 were predicted to have most pronounced effect on the hCGβ protein and specifically on the formation of the hCGα/β complex.

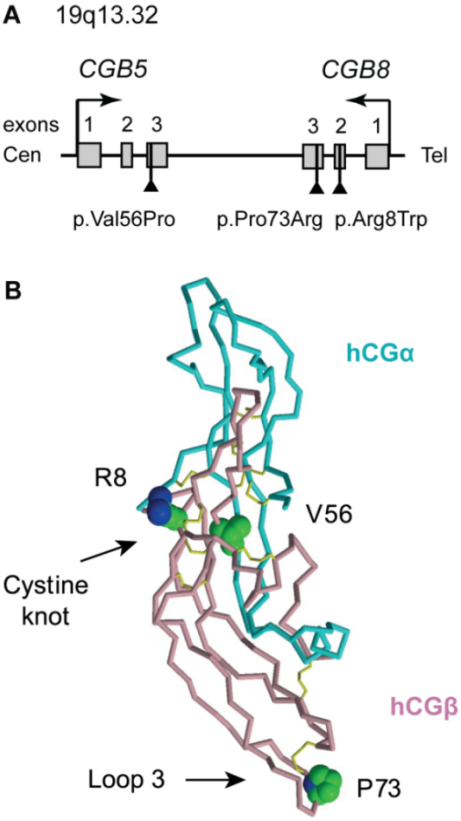

Figure 8. Localization of identified non-synonymous mutations in the *CGB5* and *CGB8* genes (**A**) and in the structure of assembled hCG molecule (**B**). (**A**) The positions of the mutations in the *CGB5* and *CGB8* gene. Gray boxes represent exons with the exon number given above. Black arrows indicate the direction of gene transcription. Cen, centromere; Tel, telomere. (**B**) Three-dimensional (3D) structure of the assembled hCG molecule based on Protein Data Bank (PDB; http://www.pdb.org) entry 1hcn. The structure of the hCG α-subunit is depicted in blue, β-subunit in pink and the disulfide bonds in yellow. The side chains of amino acids Arg8, Val56 and Pro73 in hCGβ are shown in the space-filling representation.

3.2.2.2. Structural features and assembly of recombinant hCGβ isoforms

To study the effect of the rare missense mutations in *in vitro* experimental setup, I constructed four alternative recombinant FLAG-tagged hCGβ variants – the wild-type and the hCGβ proteins carrying either p.Arg8Trp, p.Val56Leu or p.Pro73Arg mutation. Each FLAG-tagged hCGβ variant was transiently co-expressed with un-tagged α-subunit in the CHO cell line.

Forty-eight hours after transfection, the media containing the secreted hCGα and hCGβ isoforms were collected and subjected to co-immunoprecipitation using anti-FLAG antibody for specific precipitation of free FLAG-tagged hCGβ monomers and heterodimeric hCGα/β complexes. To study the structure and assembly of the expressed isoforms, the retrieved molecules were run on SDS-PAGE and detected by Western blot using either the anti-FLAG antibody which probed for both hCGβ monomers and assembled α/β heterodimers (**Figure 9A,C**), or via antiserum against hCG α-subunit which specifically visualizes assembled heterodimer hormone only (**Figure 9B**). Either non-reducing (**Figure 9A,B**) or reducing (**Figure 9C**) SDS-PAGE conditions were used based on either absence or presence of 2-mercaptoethanol and heat denaturation, respectively.

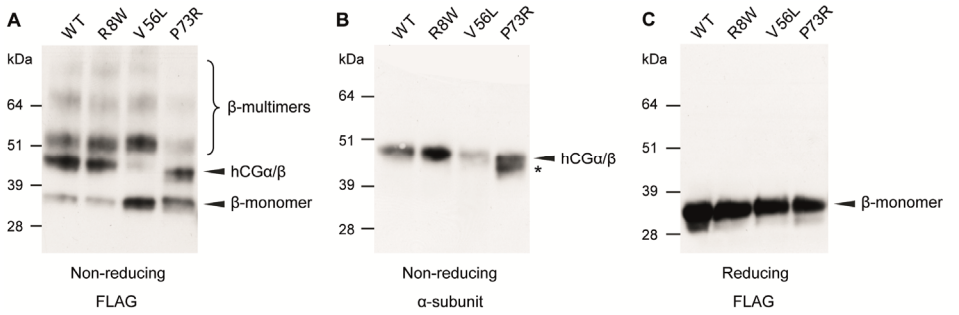

Figure 9. Co-immunoprecipitation and Western blot analysis of FLAG-tagged hCGβ variants co-expressed with hCGα in CHO cells. FLAG-tagged hCGβ monomers and associated complexes were immunoprecipitated from CHO cell culture media using anti-FLAG antibody-conjugated beads and separated by SDS-PAGE under non-reducing (**A,B**) or reducing (**C**) conditions. (**A,C**) Free and heterodimeric assembled FLAG-tagged hCGβ was detected using anti-FLAG antibody. (**B**) Heterodimeric hCG was specifically visualized using antiserum to hCG α-subunit. Bands corresponding to the heterodimeric hCG and unassembled hCGβ monomers are indicated by arrowheads; bands corresponding to β-subunit specific multimeric complexes that have been shown to be secreted from the cells, especially in the presence of mutations that affect the β-subunit folding pathway (Bedows et al. 1994; Feng et al. 1995; Feng et al. 1996), are indicated with a bracket. The data is drawn from the same experiment and it is representative of three independent co-immunoprecipitation experiments. Alternative α/β complex with the hCGβ conformational isoform caused by the p.Pro73Arg mutation is indicated by asterisk (*). WT, wild-type.

Under non-reducing conditions, the proportion of the p.Val56Leu β-subunits incorporated into α/β heterodimers (~47 kDa) relative to the freely retained β-monomers (~34 kDa) was notably reduced compared to the wild-type (**Figure 9A,B**) indicating a decreased capability of p.Val56Leu β-subunits to assemble into the heterodimer and/or destabilizing the assembled hormone. The effect was confirmed by the quantitative immunoassays detecting the assembled intact hCG (hCG Human ELISA Kit; Abcam, Cambridge, MA) or total hCGβ (hCG+β kit on Roche Elecsys 1010 system; measured in Tartu University Hospital, Tartu, Estonia) from the cell culture media. The efficiency of hCG assembly (measured as the ratio of assembled hCG to the total hCGβ amount) was decreased down to 10% compared to the wild-type hCGβ (= 100%; Student's t test, $P = 0.014$) (Figure 5 in Ref. V). The results are largely concordant with the *in silico* positional context analysis predicting the hindrance of p.Val56Leu substitution upon the hCG heterodimer formation due to its location immediately next to Cys57 in the highly conserved structural feature, the cystine knot (**Figure 8B**). The Cys57 forms the Cys9-Cys57 disulfide bond disruption of which has previously been shown to give rise to folding and assembly deficient hCGβ protein (Bedows et al. 1994; Mishra et al. 2003).

The hCGβ p.Pro73Arg mutant gave rise to approximately equal amounts of two alternative hCGβ isoforms, one corresponding to the molecular weight of the wild-type hCGβ (~34 kDa) and an additional variant with approximately 2 kDa lower molecular weight (**Figure 9A,B**). Pro73 is located next to the Cys72 that forms one of the six hCGβ disulfide bonds, Cys23-Cys72. It has been demonstrated previously that disruption of this bond affects the hCGβ folding pathway leading to secretion of an additional isoform lacking the Cys23-Cys72 bond and exhibiting a difference of 2 kDa in size on SDS-PAGE (Bedows et al. 1993; Bedows et al. 1994). Nevertheless, both isoforms detected in this study were assembled into hCG dimer with approximately equal efficiency (**Figure 9A,B**).

In the case of hCGβ p.Arg8Trp substitution, no detectable differences in the fraction of free hCGβ or in the assembly of intact hCG compared to the wild-type variant were revealed (**Figure 9**). Also, no evidence of the effect of the studied three substitutions on the glycosylation pattern of the hCGβ protein was seen based on the SDS-PAGE performed under reducing conditions that cause dissociation of hCG dimers and disruption of disulfide bonds. All tested hCGβ variants collapsed into one major (~34 kDa) and one minor (~31 kDa) isoform (**Figure 9C**), previously shown to contain either two or one N-linked oligosaccharide chains, respectively (Matzuk et al. 1987).

3.2.2.3. Bioactivity of recombinant hCGβ isoforms

Bioactivity of the hCG hormone can be estimated based on its ability to induce cAMP signaling upon binding to the human LH/CG receptor. In order to achieve sufficient signaling response, a high concentration of FLAG-tagged intact hCG (termed 'high yield' hCG) was produced using QMCF technology (by Icosagen Cell Factory OÜ; Supplemental data, Text S1 in Ref. VI) (Silla *et al.*, 2005). I used the series of diluted cell-conditioned media containing either the wild-type or mutant 'high yield' hCG adjusted for concentration of hCG heterodimer to stimulate HEK293 cell-line stably transfected with the human LH/CG receptor (HEK-hLH/CGR) and containing the cAMP-responsive (CRE) firefly luciferase reporter gene. After 5–6 h stimulation, the CRE-luciferase activity was determined and the EC_{50} values (± standard deviation, SD) [EC_{50} is defined as the concentration of the hormone required to produce 50% of maximal response] were estimated for each hCG isoform.

Interestingly, although the *CGB5* p.Val56Leu substitution leads to inefficient assembly and/or instability of the heterodimeric hCG, the *in vitro* bioactivity analysis indicated an increased potency of this isoform. The cAMP response to stimulation by the p.Val56Leu heterodimer was significantly more sensitive than to wild-type hCG (**Figure 10**), exhibiting a half-maximal response EC_{50} of 2.50 ± 0.81 pg/ml compared to the wild-type hCG EC_{50} of 11.41 ± 2.32 pg/ml (*t* test $P < 0.0013$). As a consequence, the shortage of the produced heterodimeric hormone (10% compared to wild-type) may be partly or fully compensated for by its increased bioactivity.

The hCGβ variant carrying the *CGB8* p.Pro73Arg mutation did not induce differences in the cAMP signaling when compared to the wild-type hCG ($P = 0.053$; **Figure 10**) despite the appearance of two alternative hCGβ isoforms on SDS-PAGE (**Figure 9**). Importantly, the overall functional characteristics of the assembled hCG thus remained comparable to the wild-type pointing to the functional neutrality of this mutation.

The p.Arg8Trp hCGβ isoform represented a fully neutral amino acid substitution out of the three addressed in this study as no impact on structural or functional characteristics were observed compared to the wild-type hCGβ variant.

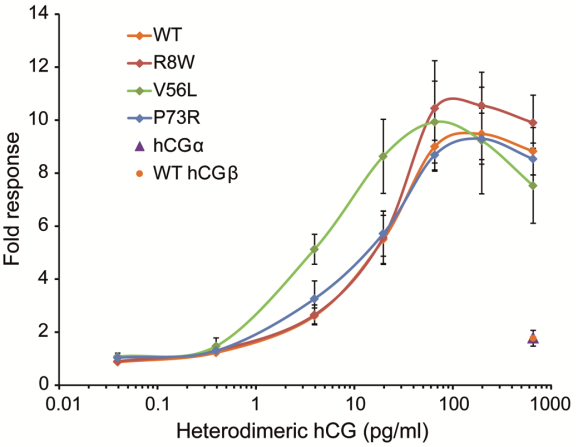

Figure 10. Bioactivity of hCGβ variants measured as hLH/CG receptor mediated cAMP signaling response to stimulation with the dosage gradient of heterodimeric wild-type (WT) and mutant hCG preparations. Stimulation with wild-type hCGβ or hCGα monomers was used as a negative control. The fold response is given as a ratio of CRE luciferase activity to unstimulated cells. The data are the mean ±SD of five independent experiments.

3.3. DNA copy number variants in recurrent miscarriage (Ref. VII)

Rationale of the study:

Previously published studies addressing the genetic etiology of RM have primarily targeted SNPs producing few confirmed disease risk variants. Although the alternative genetic variation class of copy number variants (CNVs) has been associated with various complex disorders, it has been largely understudied in pregnancy complications. My goal in this study (Ref. VII) was to elucidate the role of CNVs in modulating predisposition to RM.

3.3.1. Genome-wide profile of CNVs

In order to define genome-wide profile of CNVs specific to RM cases or fertile controls, screening with Illumina Human370CNV-Quad SNP array was undertaken in 70 Estonian subjects including 27 female and 16 male idiopathic RM cases with at least three miscarriages before gestational week 22 and 27 female controls with at least three live births and no miscarriages prior to recruitment. On average, 13.3 CNVs were determined per individual among the RM cases (in total, n = 43) and 12.6 CNVs/individual in fertile controls (n = 27). Identified CNVs clustered into 423 non-overlapping discrete Copy Number Variable Regions (CNVRs) (Supplemental Table S3 in Ref. VII) that were uniformly distributed across the genome (**Figure 11**).

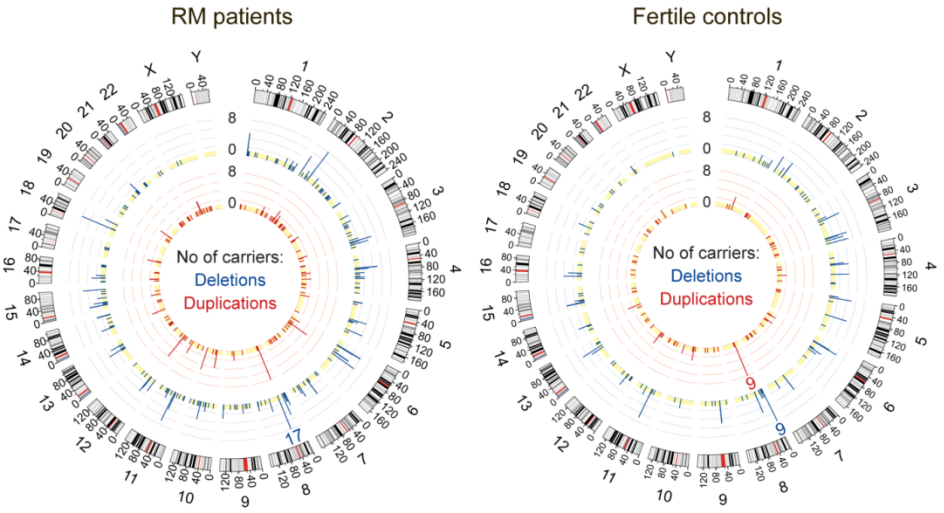

Figure 11. Circos plots with frequency and distribution of deletion and duplication CNVRs across the genome among the Estonian discovery cases (n = 43) and controls (n = 27). The length of the bars (Y-axis) in the histogram represents the number of carriers of a CNV at the genomic locus, whereas the width correlates with the size of the CNVR. Single occurrence CNVs are highlighted in the yellow background. The number of carriers for outlier CNVRs of high frequency (>8 carriers in a study group) are indicated next to respective bars.

In accordance with the known heterogeneity in the etiology of RM (Rai and Regan 2006), an individual-specific increased risk of RM was inferred for a subset of patients using case-by-case CNV profile analysis. Five outlier cases (12% of RM subjects studied) were observed with (**i**) a 5-fold excess cumulative burden of all CNVs (up to 6.1 Mb per genome) mostly comprising of long (>100 kb) deletions (**Figure 12A**) or (**ii**) accumulation of long deletions with very high gene count (up to 189 genes per individual compared to only a median of 8.3 genes among the fertile controls) (**Figure 12B**). The list of rearranged genes in the outlier cases with the largest cumulative burden of all CNVs (n = 2; **Figure 12A**) involved several known or potential candidate genes for RM and pregnancy success, such as *C4A, C4B, IGF2* or *Immunoglobulin heavy chain (IGH)* gene cluster that may contribute to the increased chances of miscarriages among these patients.

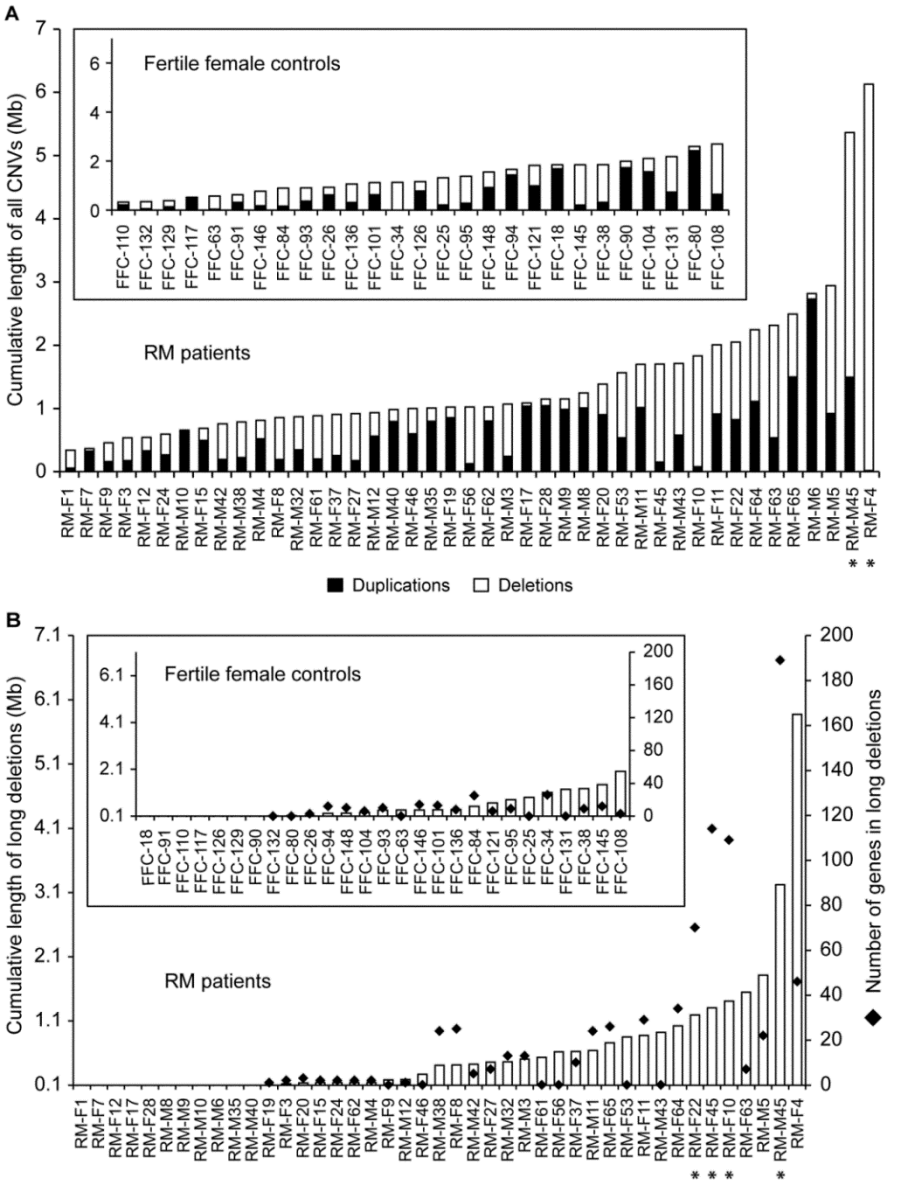

Figure 12. Genomic burden of all CNVs and the subset of long (≥100 kb) deletions in the Estonian discovery phase sample set. (**A**) Cumulative length of all deletions and duplications per individual in RM patients (n = 43) and fertile controls (n = 27). The outlier cases with increased cumulative burden of all CNVs are indicated with asterisk. (**B**) Cumulative length of long deletions and the number of disrupted genes per individual in the discovery phase cases and controls. The outlier cases with increased number of genes disrupted by long deletions are indicated with asterisk. Female and male patients with identical number-codes represent RM couples (e.g. RM-F45 and RM-M45). FFC, fertile female control; RM-F, female RM patient; RM-M, male RM patient.

3.3.2. Functional enrichment of genes disrupted by CNVs

To define the functional impact of the discovered CNVs and identify biological pathways significantly affected by the CNVs, functional enrichment analysis was performed for the list of all disrupted genes in either RM cases (1459 genes) or fertile female controls (553 genes) using g:Profiler software (Reimand et al. 2011). Among the cases, the results highlighted the specific impact of CNVs on the processes of immunomodulatory function at the feto-maternal interface related to maternal rejection of the semi-allogeneic fetus expressing paternally inherited alloantigens. Functional categories such as 'Innate immunity signaling' (REAC: 168249; 11.0% of genes disrupted by CNVs, multiple testing corrected $P = 3.57 \times 10^{-3}$), 'Fc gamma receptors interact with antigen-bound IgG' (REAC: 199161; 26.3% disrupted genes, $P = 9.97 \times 10^{-6}$) and 'Complement cascade' (REAC: 166658; 15.6% disrupted genes, $P = 1.93 \times 10^{-4}$) (Table 1 in Ref. VII) were significantly and specifically overrepresented in RM case sample. None of the immunomodulatory pathways were significantly affected by CNVs among the controls and only processes associated with general cellular function were identified (Table 1 in Ref. VII).

3.3.3. Identification of novel common CNV regions conferring risk to recurrent miscarriage

In search for novel common CNV regions that may independently modulate the predisposition to RM, I performed an experimental analysis of prioritized discrete CNV regions and conducted association study in two North European populations.

3.3.3.1. Prioritized CNVRs affecting RM in Estonia and Denmark

Nine discrete common CNV regions were selected from the whole genome SNP array genotyping data (Supplemental Table S4 in Ref. VII) for subsequent experimental testing in the Estonian discovery sample set (n = 70) using TaqMan qPCR. The selection included CNVRs that were (i) present in >1 individual, (ii) found only among RM patients or overrepresented in RM patients with OR ≥1.5 and (iii) intersected with or located in the proximity (up to approximately 200 kb) of biological candidate genes with a potential impact on the course of pregnancy based on previously published literature. Three CNVRs with precise TaqMan copy number typing assays were tested in the full Estonian sample set: CNV regions of *IGKV* (*Immunoglobulin kappa variable cluster* at 2p11.2), *DKK2* (*Dickkopf 2 homolog*, at 4q25), and *PDZD2:GOLPH3* (*PDZ domain containing 2*; *Golgi phosphoprotein 3* at 5p13.3)).

The *PDZD2:GOLPH3* duplication exhibited he strongest effect (OR = 7.28) with a higher prevalence of duplication carriers among the Estonian RM cases compared to fertile controls (9/119, 7.6% versus 1/90, 1.1%, respectively)

(Supplemental Table S7 in Ref. VII). However, analysis of the Danish replication sample (in total, 439 RM patients with ≥2 consecutive miscarriages, 115 multiparous fertile female controls) identified increased prevalence of the duplication only among female patients who exhibited carrier frequency comparable to Estonian female cases (6.6% and 7.5%, respectively) (**Figure 13**; Supplemental Table S8 in Ref. VII). Meta-analysis combining the results of the Estonian and Danish female patient-control samples (in total, cases n = 309, controls n = 205) confirmed the association of the *PDZD2:GOLPH3* CNV with an increased maternal risk of RM (OR = 4.82, $P = 0.012$). (Table 2 in Ref. VII). Concordantly, high risk of RM associated with the duplication was detected when the Estonian RM women were independently tested against female controls of Estonian population cohort from the Estonian Genome Center, University of Tartu (EGCUT) (n = 496; prevalence 1.0%; OR = 7.96, $P = 7.9 \times 10^{-4}$) (**Figure 13**).

The *IGKV* and *DKK2* CNVs with small differences in carrier frequencies in the full Estonian case-control sample (Supplemental Table S7 in Ref. VII) likely represent benign common copy number variation with no major effect on the RM phenotype in our study.

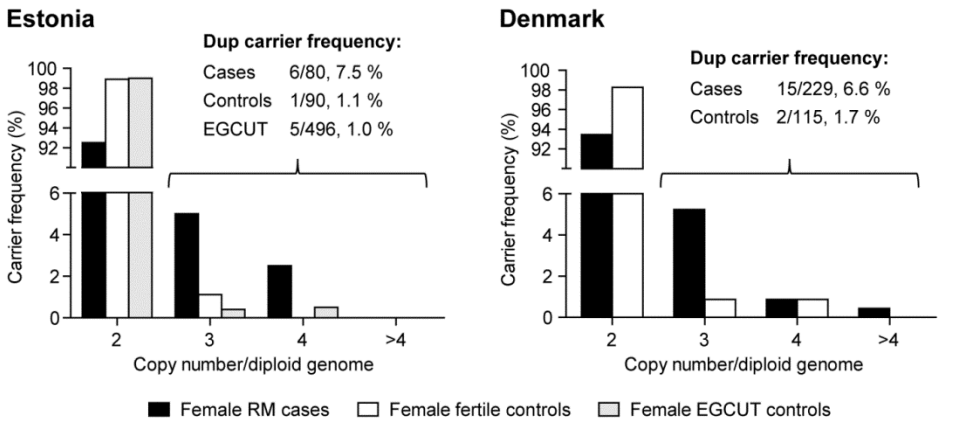

Figure 13. Copy number distribution and carrier frequency of the *PDZD2:GOLPH3* duplication at 5p13.3 among female study subjects. CNV carriers have 3 to 4 copies of the duplication per genome in Estonia and 3 to >4 copies per genome in Denmark. Dup, duplication; EGCUT, Estonian population cohort from Estonian Biobank, Estonian Genome Center, University of Tartu.

3.3.3.2. Genomic context and fine-mapping of the *PDZD2:GOLPH3* duplication

To define the range and nature of the *PDZD2:GOLPH3* duplication (occurring in up to >4 diploid copies) at 5p13.3 predisposing to RM, I set forward to fine-map the duplication breakpoints by integrating experimental (EvaGreen qPCR, PCR, sequencing) and bioinformatic approaches (including screening for

repetitive elements with RepeatMasker, http://www.repeatmasker.org/; and non-B DNA sequence motifs with Non-B DNA Motifs Search Tool, http://nonb.abcc.ncifcrf.gov/apps/nBMST/default/). The CNVR was estimated as 61.6 kb in length (positioned Chr5: 32106204 – 32167777), whereas the duplication breakpoints were located within DNA repetitive elements in the introns of *PDZD2* and *GOLPH3* genes that are transcribed in the opposite direction (**Figure 14A,B**). Application of duplication breakpoint junction-specific PCR (BP-PCR, **Figure 14B**) identified an identical recurrent tandem duplication event in all CNV carriers of this study. It was hypothesized that the *PDZD2:GOLPH3* duplication occurred via repeat-mediated rearrangement mechanisms other than non-allelic homologous recombination. Although the *PDZD2:GOLPH3* rearrangement does not involve entire coding regions of the genes, the modifications in the local genomic context may nevertheless lead to impaired function of the involved or neighboring genes as reported previously (Henrichsen et al. 2009a; Henrichsen et al. 2009b).

3.3.3.3. Expression profile of *PDZD2* and *GOLPH3*

Neither *PDZD2* nor *GOLPH3* has been associated with pregnancy success previously and have mostly gained attention due to their role in tumorigenesis (Tam et al. 2006; Scott et al. 2009). Thus little is known on the tissue expression profile of these genes.

In order to identify the major sites of expression, I performed an expression profiling analysis using human tissue cDNA panels (Human MTC panels I and II; BD Biosciences Clontech, CA) and TaqMan qPCR assays specific to the transcripts of *PDZD2* and *GOLPH3*. I detected the most prominent expression for both genes in the placenta that exceeded other sites of highest expression twofold for *GOLPH3* (average expression relative to reference *HPRT*, 2.99 ± 0.06 SEM in placenta versus 1.50 ± 0.05 in prostate) or fourfold for *PDZD2* (relative expression 1.33 ± 0.05 in the placenta versus 0.36 ± 0.01 in heart) (**Figure 14C**). Interestingly, ovary was also included in the top three sites of expression for both *GOLPH3* and *PDZD2* (1.13 ± 0.02 and 0.34 ± 0.0, respectively) and thus highlighting the functional relevance of these genes in reproductive organs and specifically in the placenta.

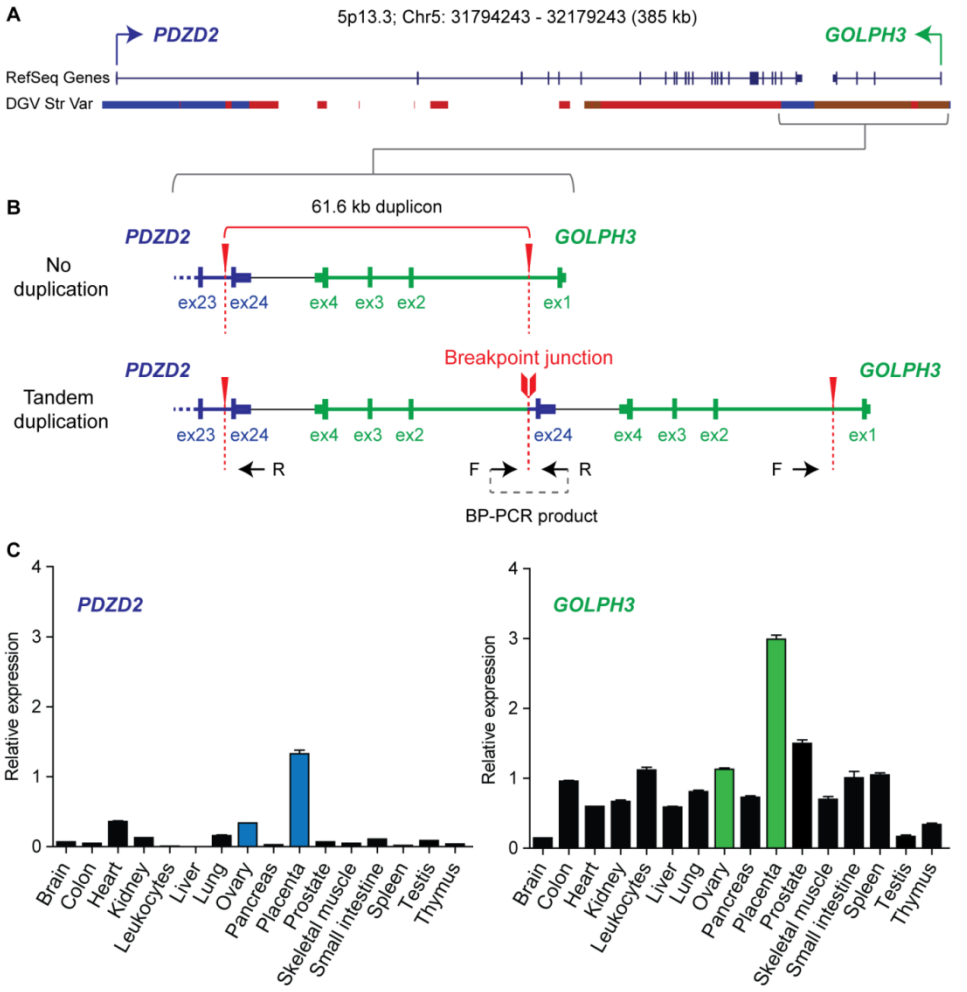

Figure 14. Fine-mapping of the *PDZD2:GOLPH3* duplication CNV at 5p13.13 and expression profiling of the *PDZD2* and *GOLPH3* genes. (**A**) Genomic context of 5p13.3 involving *PDZD2* and *GOLPH3* genes based on UCSC database (hg19; http://genome.ucsc.edu/cgi-bin/hgGateway). The opposite transcription of the *PDZD2* and *GOLPH3* genes is indicated with blue and green arrows, respectively. DGV Struc Var, structural variation data from the Database of Genomic Variants. (**B**) Schematic representation of the 5p13.3 CNV locus with or without tandem duplication. Experimentally confirmed duplication endpoints are indicated with red arrowheads and dotted lines. The breakpoint junction of the tandem duplication is marked with red arrow tail and locations of breakpoint junction-specific PCR (BP-PCR) primers are indicated with black arrows. In case of tandem duplication, a 555 bp product was amplified and subsequently confirmed by resequencing. Ex, exon. (**C**) Gene expression profile of the *PDZD2* and *GOLPH3* genes in the human cDNA tissue panels. Expression level is given relative to the reference gene *HPRT* and as average of three amplification reactions ± SEM. Gene expression levels in placenta and ovary are highlighted with colored bars.

4. DISCUSSION

4.1. Nature and impact of fine-scale genetic variation of the *LHB/CGB* region in RM

The genes encoding the hormone-specific hCGβ subunit are one of the most attractive targets in addressing fine-scale genetic etiology of early pregnancy loss as the requirement of the 'pregnancy hormone' hCG for the establishment and maintenance of pregnancy is absolute. The genomic landscape of the *LHB/CGB* gene cluster in humans described in this thesis has likely been shaped by a complex interplay between active genome dynamics and balancing act of functional constraints maintaining the effectiveness of the LHβ/CGβ genes. Extensive DNA sequence identity between DNA segments, density of *Alu* elements, formation of DNA hairpin conformation and presence of recombination warm/hot spots in the *LHB/CGB* region are rich substrates for increasing genome instability and gene conversion activity (Chen et al. 2007; Chuzhanova et al. 2009). As a result of the active genome dynamics, complex evolutionary rearrangements involving multiple duplication events have given rise to a varying number of *CGB* genes among primates with six gene copies in humans and up to even 50 copies predicted for gorilla (Dumas et al. 2007; Hallast et al. 2008). The abundant duplicated segments have been further shaped by the directional interlocus gene conversion which is a common phenomenon in the duplicated parts of the human genome and has introduced diversity at the edges of the gene cluster on one hand and on the other hand, has homogenized the duplicated genes maintaining the recombinogenic potential of the region (Dumont and Eichler 2013; Fawcett and Innan 2013).

It is noteworthy that in spite of the genomic instability and gene conversion activity spreading polymorphisms between gene copies but also introducing *de novo* mutations (Hicks et al. 2010; Dumont and Eichler 2013), no gene variants of substantial effect on pregnancy success have been identified in hCGβ coding genes so far. The data suggest an accompanying action of selection forces in order to maintain the functional efficacy of the genes and concurrently, balancing selection acting on the promoter region of *CGB8* and driving the most effective transcription among the *CGB* genes was reported in this thesis (Ref. IV). Furthermore, the polymorphism screening of the two most actively transcribed hCGβ-coding genes *CGB5* and *CGB8* and subsequent association testing noted the lack of risk alleles associated with RM disease in Northern European populations. Only a protective effect against RM was observed for the motif of four regulatory polymorphisms in the *CGB5* gene likely transferred from the respective region of *CGB8* genes via non-reciprocal gene conversion event (Ref. IV and V). Similar intricate action of gene conversion and functional constraints in modifying the local genomic context has been observed for other complex genomic regions associated with pregnancy course, including *growth hormone/ somatomammotropin* (Sedman et al. 2008), *killer*

immunoglobulin-like receptor (Yawata et al. 2006; Graef et al. 2009), *pregnancy-specific glycoprotein* (McLellan et al. 2005; Dumont and Eichler 2013) gene families and MHC class II region (von Salome et al. 2007).

Further evidence of the functional constraints acting on hCGβ genes due to their irreplaceable role in human pregnancy is suggested by the lack of non-synonymous mutations conferring large functional effect within the genes. Out of the three missense mutations identified in the major *hCGβ* genes, *CGB5* and *CGB8,* and addressed in this thesis (*CGB5* p.Val56Leu, CGB8 p.Arg8Trp and p.Pro73Arg), only the p.Val56Leu substitution was initially found to hinder the assembly of the heterodimeric hormone (Ref. VI). However, the simultaneous 5-fold increase in potency partly or fully compensated for the low proportion of the assembled heterodimer (10% compared to wild-type) restoring the overall functional effect. Although inefficient hCG assembly has also been shown for the only naturally occurring variant of hCGβ (*CGB5* p.Val79Met) characterized *in vitro* (Miller-Lindholm et al. 1999), no clinical information was available on the mutation carriers and the substitution was completely absent in a subsequent study of over 500 DNA samples from five European populations (Jiang et al. 2004). Only few other missense mutations have been identified in populations world-wide in the *CGB5* (p.Arg6Gln in Han Chinese; p.Asp117Ala in African Mandekalu; Ref. III) and *CGB8* genes (p.Val29Ile in Estonia and Finland; Ref. IV) with the carrier frequency of <10% in all cases but with no clinical information available for most of the mutation carriers. Furthermore, no individuals homozygous for any hCGβ mutations have been reported so far, which may indicate either an insufficient sample size in the conducted studies or that such genotypes result in complete pregnancy failure.

The lack of risk variants of substantial effect in the major *hCGβ* genes raises the question of missing genetic factors responsible for the large interindividual variation of *hCGβ* genes in gene expression during pregnancy and dysregulated expression in pregnancy disorders (Ref. II) (Miller-Lindholm et al. 1997; Rull and Laan 2005). It has been speculated whether the pathologically low hCG levels are indeed the primary cause of pregnancy loss or a secondary phenomenon resulting from inadequate placental maturation due to alternative risk factors (Larsen et al. 2013). Considering the genetic heterogeneity of the RM phenotype, it is feasible that both scenarios are relevant and the predisposing genetic profile of *hCGβ* genes that is likely subject to constant dynamic changes is responsible for only a subset of RM cases exhibiting decreased expression levels of hCGβ transcripts.

4.2. Genome-wide effect of CNVs in RM

4.2.1. Genomic CNV burden as risk factor for RM

The application of genome-wide CNV screening and analysis methods has enabled to compile a parental profile of CNVs predisposing to recurrent miscarriage in this thesis. A case-by-case analysis identified a subgroup of RM patients (in total, 5 out of 43; 12%) exhibiting accumulation of several long (>100 kb) deletions or long deletions with very high gene count (**Figure 12**; Ref. VII) potentially responsible for the manifestation of the disease in these individuals. An enrichment of large (cutoff 100 kb or larger) structural variants have been commonly associated with an increased risk of complex diseases previously, including neurodevelopmental disorders, longevity or severe obesity (Bochukova et al. 2010; Girirajan et al. 2011a; Kuningas et al. 2011). Furthermore, higher genome-wide prevalence of gene-rich long deletions has been reported for individuals affected by autism spectrum disorders and schizophrenia (Consortium 2008; Pinto et al. 2010; Griswold et al. 2012).

The deleterious nature of large rearrangements and specifically long deletions in gene-rich regions likely originates from increased chances of affecting loci relevant to pregnancy success. Gene content analysis of CNVs identified among RM cases with the largest overall genomic burden of rearrangements highlighted some potential dosage sensitive candidate genes, including *C4A* and *C4B* previously implicated in RM (Laitinen et al. 1991). The findings indicate that large cumulative burden of CNVs or gene-rich deletions could be regarded as an independent risk factor of RM, however the statistical significance of these findings remains to be tested further.

4.2.2. Genetic variability of immunoregulatory pathways as risk factor for RM

Adequate bi-directional communication at the feto-maternal interface and immunologic recognition of pregnancy have been proposed as of utmost importance for implantation and maintenance of early pregnancy. Fetal rejection due to either auto- or alloimmune factors is a major contributor in RM disease (reviewed in (Pandey et al. 2005; Baek et al. 2007)) and a number of candidate loci related to placental immune function have been implicated in RM independently or cumulatively, such as *KIR* and *HLA* gene families (Hviid et al. 2004; Hiby et al. 2008; Aruna et al. 2011). Similar to thrombophilic genetic factors, it has been suggested that only failure of several mechanisms or systematic deviations in pathways would lead to the occurrence of RM due to a likely redundancy in the immune function at the feto-maternal interface protecting against the fetal rejection (Jivraj et al. 2006; Larsen et al. 2013).

Multiple pieces of evidence indicate the contribution of structural variants in modifying the profile of genes related to immune function in pregnancy. In this thesis, significant enrichment of genes from immunomodulatory pathways were

observed within the rearranged loci of RM patients with half of the pathways specific to cases and not affected among controls, including complement cascade and innate immunity related processes (Table 3 in Ref. VII). The role of CNVs in promoting common genetic variability of immune system loci has previously been reported for human immunoglobulin heavy chain variable (*IGHV*) (Watson et al. 2013) and constant (*IGHC*) loci (Lefranc et al. 1991), beta-defensins (Hollox et al. 2008) and duplicate *KIR* genes (Jiang et al. 2012) indicating the potential susceptibility of these regions also to disease risk-conferring rearrangements. Furthermore, it has been shown that CNVs modulate and define the profile of presented antigens in the graft-versus-host disease, the condition of compromised immunity similar to fetal rejection in pregnancy (Guleria and Sayegh 2007; McCarroll et al. 2009). The accumulating data support the notion that CNVs may considerably affect the function of alloimmune factors and increase the risk of early miscarriages originating from feto-maternal genetic mismatch.

4.3. Global genomic analysis as the source of novel biomarkers

Hypothesis-free genome-wide approach in studying disease susceptibility is a powerful tool not only in uncovering novel pathways but also facilitating identification of novel genes and genetic variants increasing the risk of disease. Genome-wide studies targeting CNVs have been the source of numerous novel dosage sensitive candidate genes for complex diseases but also for recurrent miscarriage as reported previously by (Rajcan-Separovic et al. 2010). Aberrations of the imprinted genes *TIMP2* and *CTNNA3* in the placenta were implicated in increased risk of pregnancy loss and thus represent the risk factors specific to the fetus/offspring. Alternatively, screening of parental genomes may provide data on the genetic susceptibility of either or both parents to the occurrence of repetitive miscarriages.

A strong maternal risk of RM was identified for the 61.6 kb multicopy duplication of *PDZD2* and *GOLPH3* genes (chromosome position 5p13.3) in this thesis among women from Estonia and Denmark (meta-analysis, OR = 4.82, $P = 0.012$; Ref. VII). Interestingly, in addition to being transcribed in the female reproductive organ ovary, the highest expression of both of the involved genes was identified in the placenta (**Figure 14C**; Ref. VII). It was speculated that the increased predisposition to RM among the female duplication carriers may be attributed to the joint effect in maternal reproductive tissues and in the placental tissue carrying a maternally inherited duplication CNV.

The relevance of the *PDZD2:GOLPH3* duplication as a novel biomarker in pregnancy disorders can be estimated based on two findings. Firstly, the *PDZD2:GOLPH3* duplication CNV may represent a pleiotropic risk factor of pregnancy disorders as its risk-conferring effect is not restricted to RM. A

recent CNV association study identified the duplication as a potential risk factor in women affected by pre-eclampsia – a severe late pregnancy disorder attributed to placental dysfunction (Zhao et al. 2013). The overlapping causality of RM and pre-eclampsia (Li and Huang 2009; Baig et al. 2013) but also early pregnancy loss and fetal growth restriction (Ganguly et al. 2007) has been acknowledged previously.

Secondly, the *PDZD2* and *GOLPH3* genes affected by the 5p13.3 duplication CNV are novel candidate genes in the context of early pregnancy maintenance and the associated pathways have not been directly linked to the occurrence of early pregnancy loss previously. The function of PDZD2 is poorly defined, whereas amplification of GOLPH3, essential for Golgi trafficking and maintenance of its structure (Dippold et al. 2009; Wood et al. 2012), has been reported in various cancers and shown to induce the signaling of mechanistic target of rapamycin (mTOR) (Scott et al. 2009; Wang et al. 2012; Hu et al. 2013). The contribution of mTOR in reproductive function is however well established and alterations in mTOR signaling have been linked to multiple reproductive disorders in human and mouse, including recurrent miscarriage (Roos et al. 2007; Hirota et al. 2011; Vatin et al. 2012). The direct functional link between the GOLPH3 amplification and mTOR signaling in female reproductive tissues and placenta remains to be experimentally confirmed.

4.4. *Status quo* and future perspectives in assessing genetic determinants of RM

As pregnancy loss is the most common pregnancy complication, etiology of RM has been addressed by considerable number of studies for several decades establishing the multifactorial nature of the disease and the contribution of genetic susceptibility. Nevertheless, none of the reported variants have unequivocally proven as specific to the RM phenotype, neither prevalent among the cases leaving a void in the knowledge on disease heritability factors and also highlighting the multifactorial nature of the genetic component itself. Adjustments in study designs from the genetic and phenotypic perspective may prove beneficial in fine-mapping the genetic determinants of RM (also reviewed in Ref. I).

Genetic perspective

The common attractive expectation in studies of RM has been to identify single genetic variants of high effect efficiently applicable in clinical diagnostics. However, the accumulating research data indicate that very few genetic variants exist conferring high risk to RM likely due to the heterogeneity and redundancy of mechanisms in early pregnancy as also demonstrated for the polymorphisms in the hCGβ coding genes. It has been proposed that instead of single poly-

morphisms with independent prevalent effect, several genetic variants may cumulatively alter pathways essential for pregnancy maintenance (e.g. thrombophilic and pro-inflammatory mutations). Alternatively and feasibly, due to the multifactorial nature of RM, each case or family may carry a distinct genetic factor responsible for the increased risk of the disease observed among siblings (Christiansen et al. 2008; Kolte et al. 2011).

Although no SNP-based GWAS studies have been performed to identify the set of single nucleotide variants associated with RM, the genomic profiling of CNVs addressed in this thesis supports the hypothesis of cumulative risk of multiple rearrangements involving genes within similar pathways (in this case, processes related to immune function). The finding of the common *PDZD2:GOLPH3* duplication independently increasing the risk of RM among women with OR = 4.82 in two populations likely reflects the lack of previous studies targeting parental CNV profile in RM. Increasing the sample size may identify additional CNVs associated with RM, however the number of common causative rearrangements is likely to remain low in accordance with and as debated for other common diseases (Craddock 2010).

In order to improve the current knowledge on the whole spectrum of genetic factors leading to RM, the focus of the association studies could be further shifted from candidate gene-based analysis to genome-wide studies and incorporating data from transcriptomics, proteomics and methylomics that have become financially, methodologically and analytically more feasible and may prove beneficial in interpretation of large-scale genomic data.

Phenotypic perspective

In order to increase the power of identifying novel genetic variants of lower effect size at the genome-wide level, larger case-control samples and replication samples would be needed. However, the varying phenotyping criteria of RM at different recruitment centers such as (i) the number of miscarriages and live births at recruitment of cases and controls, respectively, (ii) exclusion criteria of cases (e.g. testing for mutations in factor V (Leiden) or factor II, prothrombin) and/or (iii) selection criteria of study subjects (only women versus couples versus placenta) may hinder the pooling of sample collections or skew replication results. Also, majority of the studies addressing the genetic component of RM have focused on only maternal or fetal/placental contribution and few have included both parents (Ref. I). As mother and father contribute equally into the offspring's genetic composition, association studies with couples may identify novel risk factors leading to early pregnancy loss. In order to uncover novel pleiotropic genetic variants not restricted to RM, parallel analysis of various pregnancy disorders may prove informative.

SUMMARY AND CONCLUSIONS

The results of this thesis can be summarized as follows:

Fine-scale genetic determinants of RM based on LHB/CGB gene cluster

1. The local genomic landscape of luteinizing hormone/chorionic gonadotropin β-subunit coding genes (*LHB/CGB*) in humans reflects the history of complex evolutionary processes involving expansion of repetitive elements and multiple rearrangements likely promoted by recombination 'warm and hot spots'. A balancing interplay between gene conversion activity spreading polymorphisms between the gene copies and functional constraints maintaining the transcription efficiency of the genes shapes the region's genomic context. It is likely that the *LHB/CGB* region is subject to further dynamic changes.
2. The most actively transcribed hCGβ genes *CGB5* and *CGB8* were characterized by the lack of genetic variants increasing the risk of RM in the case-control sample collections from Estonia, Finland and Denmark. Instead, a modest protective effect against RM was observed for a motif of four SNPs in the *CGB5* upstream region transferred from the *CGB8* gene via gene conversion event. The balancing selection acting on the upstream region of *CGB8* likely supports the highest transcription activity observed among the hCGβ genes.
3. The structural and functional analysis of three non-synonymous mutations in the *CGB5* and *CGB8* genes identified a decreased ratio of hCG heterodimer assembly for p.Val56Leu substitution. However, the low amount of formed heterodimer was compensated by its increased bioactivity upon binding to the LH/CG receptor. As only mutations of mild functional effect are tolerated and no homozygous carriers of the studied mutations or additional missense mutations with functional effect have been described in *CGB5* and *CGB8*, it was concluded that keeping the functional efficiency of the major hCGβ genes intact is essential for pregnancy maintenance.
4. Accumulating data on the fine-scale genetic variation of RM indicates that the contribution of independent variants with large effect sizes is not sufficient to explain the genetic heritability of RM and rather combinations of genetic factors or variants specific to each family lead to the manifestation of the disease.

Global genetic determinants of RM based on CNV profiling

1. Genome-wide analysis of copy number variants (CNVs) identified a high burden of all CNVs or long (>100 kb) gene-rich deletions in a subset of RM cases. It was postulated that increased genomic burden of CNVs and specifically long deletions may confer risk to RM due to increased chances of affecting genes relevant in early pregnancy establishment and maintenance.

2. Functional profiling of all genes affected by CNVs among the RM cases detected a significant and specific impact of rearrangements on the pathways related to immune function at the feto-maternal interface. The findings support the contribution of inadequate maternal immune response to semi-allogeneic fetus expressing paternal antigens in early pregnancy success.
3. Screening of distinct common CNV regions identified a novel genetic predisposition marker among women affected by RM in Estonia and Denmark. The multicopy duplication at 5p13.3 involved two genes *PDZD2* and *GOLPH3* with the highest expression in placenta but previously uncharacterized in the context of pregnancy maintenance. The *PDZD2:GOLPH3* CNV may potentially represent a novel genetic risk factor not restricted to RM.
4. Global CNV profiling of parental genomes confirmed the contribution of structural variants in shaping the genetic susceptibility to RM and may prove as an informative source of novel biomarkers in other pregnancy disorders.

Main outcome of the thesis:
This thesis established the low impact of the fine-scale genetic variation in the etiology of RM based on the *LHB/CGB* gene family and highlighted the contribution of distinct common CNVs and global burden of structural variants in shaping the genetic predisposition to recurrent early pregnancy loss.

REFERENCES

Agostinis C, Bulla R, Tisato V, De Seta F, Alberico S, Secchiero P, Zauli G. 2012. Soluble TRAIL is elevated in recurrent miscarriage and inhibits the in vitro adhesion and migration of HTR8 trophoblastic cells. *Hum Reprod* **27**(10): 2941–2947.

Albanese C, Kay TW, Troccoli NM, Jameson JL. 1991. Novel cyclic adenosine 3',5'-monophosphate response element in the human chorionic gonadotropin beta-subunit gene. *Mol Endocrinol* **5**(5): 693–702.

Aldred PM, Hollox EJ, Armour JA. 2005. Copy number polymorphism and expression level variation of the human alpha-defensin genes DEFA1 and DEFA3. *Hum Mol Genet* **14**(14): 2045–2052.

Aldrich CL, Stephenson MD, Karrison T, Odem RR, Branch DW, Scott JR, Schreiber JR, Ober C. 2001. HLA-G genotypes and pregnancy outcome in couples with unexplained recurrent miscarriage. *Mol Human Reprod* **7**(12): 1167–1172.

Alkan C, Kidd JM, Marques-Bonet T, Aksay G, Antonacci F, Hormozdiari F, Kitzman JO, Baker C, Malig M, Mutlu O et al. 2009. Personalized copy number and segmental duplication maps using next-generation sequencing. *Nat Genet* **41**(10): 1061–1067.

Aruna M, Nagaraja T, Andal Bhaskar S, Tarakeswari S, Reddy AG, Thangaraj K, Singh L, Reddy BM. 2011. Novel alleles of HLA-DQ and -DR loci show association with recurrent miscarriages among South Indian women. *Hum Reprod* **26**(4): 765–774.

Aruna M, Sudheer PS, Andal S, Tarakeswari S, Reddy AG, Thangaraj K, Singh L, Reddy BM. 2010. HLA-G polymorphism patterns show lack of detectable association with recurrent spontaneous abortion. *Tissue Antigens* **76**(3): 216–222.

Ascoli M, Fanelli F, Segaloff DL. 2002. The lutropin/choriogonadotropin receptor, a 2002 perspective. *Endocr Rev* **23**(2): 141–174.

Baek KH, Choi BC, Lee JH, Choi HK, Lee SH, Kim JW, Hill JA, Chung HM, Ko JJ, Cha KY. 2002. Comparison of gene expression at the feto-maternal interface between normal and recurrent pregnancy loss patients. *Reprod Fertil Dev* **14**(3–4): 235–240.

Baek KH, Lee EJ, Kim YS. 2007. Recurrent pregnancy loss: the key potential mechanisms. *Trends Mol Med* **13**(7): 310–317.

Baig S, Lim JY, Fernandis AZ, Wenk MR, Kale A, Su LL, Biswas A, Vasoo S, Shui G, Choolani M. 2013. Lipidomic analysis of human placental syncytiotrophoblast microvesicles in adverse pregnancy outcomes. *Placenta* **34**(5): 436–442.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**(1): 37–48.

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**(2): 263–265.

Bedows E, Huth JR, Suganuma N, Bartels CF, Boime I, Ruddon RW. 1993. Disulfide bond mutations affect the folding of the human chorionic gonadotropin-beta subunit in transfected Chinese hamster ovary cells. *J Biol Chem* **268**(16): 11655–11662.

Bedows E, Norton SE, Huth JR, Suganuma N, Boime I, Ruddon RW. 1994. Misfolded human chorionic gonadotropin beta subunits are secreted from transfected Chinese hamster ovary cells. *J Biol Chem* **269**(14): 10574–10580.

Berger P, Gruschwitz M, Spoettl G, Dirnhofer S, Madersbacher S, Gerth R, Merz WE, Plas E, Sampson N. 2007. Human chorionic gonadotropin (hCG) in the male reproductive tract. *Mol Cell Endocrinol* **260–262**: 190–196.

Berglund J, Nevalainen EM, Molin AM, Perloski M, Andre C, Zody MC, Sharpe T, Hitte C, Lindblad-Toh K, Lohi H et al. 2012. Novel origins of copy number variation in the dog genome. *Genome biology* **13**(8): R73.

Berndt S, Blacher S, Munaut C, Detilleux J, Perrier d'Hauterive S, Huhtaniemi I, Evain-Brion D, Noel A, Fournier T, Foidart JM. 2013. Hyperglycosylated human chorionic gonadotropin stimulates angiogenesis through TGF-beta receptor activation. *FASEB J* **27**(4): 1309–1321.

Beydoun H, Saftlas AF. 2005. Association of human leucocyte antigen sharing with recurrent spontaneous abortions. *Tissue Antigens* **65**(2): 123–135.

Bo M, Boime I. 1992. Identification of the transcriptionally active genes of the chorionic gonadotropin beta gene cluster in vivo. *J Biol Chem* **267**(5): 3179–3184.

Bochukova EG, Huang N, Keogh J, Henning E, Purmann C, Blaszczyk K, Saeed S, Hamilton-Shield J, Clayton-Smith J, O'Rahilly S et al. 2010. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* **463**(7281): 666–670.

Bombell S, McGuire W. 2008. Cytokine polymorphisms in women with recurrent pregnancy loss: meta-analysis. *Aust N Z J Obstet Gynaecol* **48**(2): 147–154.

Breunis WB, van Mirre E, Bruin M, Geissler J, de Boer M, Peters M, Roos D, de Haas M, Koene HR, Kuijpers TW. 2008. Copy number variation of the activating FCGR2C gene predisposes to idiopathic thrombocytopenic purpura. *Blood* **111**(3): 1029–1038.

Casarini L, Lispi M, Longobardi S, Milosa F, La Marca A, Tagliasacchi D, Pignatti E, Simoni M. 2012. LH and hCG action on the same receptor results in quantitatively and qualitatively different intracellular signalling. *PloS one* **7**(10): e46682.

Cha J, Sun X, Dey SK. 2012. Mechanisms of implantation: strategies for successful pregnancy. *Nat Med* **18**(12): 1754–1767.

Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP. 2007. Gene conversion: mechanisms, evolution and human disease. *Nature reviews Genetics* **8**(10): 762–775.

Choi HK, Choi BC, Lee SH, Kim JW, Cha KY, Baek KH. 2003. Expression of angiogenesis- and apoptosis-related genes in chorionic villi derived from recurrent pregnancy loss patients. *Mol Reprod Dev* **66**(1): 24–31.

Christiansen OB, Mathiesen O, Lauritsen JG, Grunnet N. 1990. Idiopathic recurrent spontaneous abortion. Evidence of a familial predisposition. *Acta Obstet Gynecol Scand* **69**(7–8): 597–601.

Christiansen OB, Nielsen HS, Kolte AM. 2006. Inflammation and miscarriage. *Semin Fetal Neonatal Med* **11**(5): 302–308.

Christiansen OB, Steffensen R, Nielsen HS, Varming K. 2008. Multifactorial etiology of recurrent miscarriage and its scientific and clinical implications. *Gynecol Obstet Invest* **66**(4): 257–267.

Chuzhanova N, Chen JM, Bacolla A, Patrinos GP, Ferec C, Wells RD, Cooper DN. 2009. Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair. *Hum Mutat* **30**(8): 1189–1198.

Clifton VL, Stark MJ, Osei-Kumah A, Hodyl NA. 2012. Review: The feto-placental unit, pregnancy pathology and impact on long term maternal health. *Placenta* **33 Suppl**: S37–41.

Cole LA. 2009. hCG and hyperglycosylated hCG in the establishment and evolution of hemochorial placentation. *J Reprod Immunol* **82**(2): 112–118.

Cole LA, Shahabi S, Oz UA, Bahado-Singh RO, Mahoney MJ. 1999. Hyper-glycosylated human chorionic gonadotropin (invasive trophoblast antigen) immunoassay: A new basis for gestational Down syndrome screening. *Clin Chem* **45**(12): 2109–2119.

Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* **464**(7289): 704–712.

Consortium TIS. 2008. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**(7210): 237–241.

Craddock N Hurles ME Cardin N Pearson RD Plagnol V Robson S Vukcevic D Barnes C Conrad DF Giannoulatou E et al. 2010. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* **464**(7289): 713–720.

Daher S, Shulzhenko N, Morgun A, Mattar R, Rampim GF, Camano L, DeLima MG. 2003. Associations between cytokine gene polymorphisms and recurrent pregnancy loss. *J Reprod Immunol* **58**(1): 69–77.

Diplas AI, Lambertini L, Lee MJ, Sperling R, Lee YL, Wetmur J, Chen J. 2009. Differential expression of imprinted genes in normal and IUGR human placentas. *Epigenetics: official journal of the DNA Methylation Society* **4**(4): 235–240.

Dippold HC, Ng MM, Farber-Katz SE, Lee SK, Kerr ML, Peterman MC, Sim R, Wiharto PA, Galbraith KA, Madhavarapu S et al. 2009. GOLPH3 bridges phosphatidylinositol-4- phosphate and actomyosin to stretch and shape the Golgi to promote budding. *Cell* **139**(2): 337–351.

Dirnhofer S, Hermann M, Hittmair A, Hoermann R, Kapelari K, Berger P. 1996. Expression of the human chorionic gonadotropin-beta gene cluster in human pituitaries and alternate use of exon 1. *J Clin Endocrinol Metab* **81**(12): 4212–4217.

Dizon-Townson DS, Meline L, Nelson LM, Varner M, Ward K. 1997. Fetal carriers of the factor V Leiden mutation are prone to miscarriage and placental infarction. *Am J Obstet Gynecol* **177**(2): 402–405.

Dumas L, Kim YH, Karimpour-Fard A, Cox M, Hopkins J, Pollack JR, Sikela JM. 2007. Gene copy number variation spanning 60 million years of human and primate evolution. *Genome Res* **17**(9): 1266–1277.

Dumont BL, Eichler EE. 2013. Signals of historical interlocus gene conversion in human segmental duplications. *PloS one* **8**(10): e75949.

Elliott MM, Kardana A, Lustbader JW, Cole LA. 1997. Carbohydrate and peptide structure of the alpha- and beta-subunits of human chorionic gonadotropin from normal and aberrant pregnancy and choriocarcinoma. *Endocrine* **7**(1): 15–32.

Emerson JJ, Cardoso-Moreira M, Borevitz JO, Long M. 2008. Natural selection shapes genome-wide patterns of copy-number polymorphism in Drosophila melanogaster. *Science* **320**(5883): 1629–1631.

Enders AC, Carter AM. 2004. What can comparative studies of placental structure tell us? – A review. *Placenta* **25 Suppl A**: S3–9.

Erlebacher A, Zhang D, Parlow AF, Glimcher LH. 2004. Ovarian insufficiency and early pregnancy loss induced by activation of the innate immune system. *J Clin Invest* **114**(1): 39–48.

Fanciulli M, Norsworthy PJ, Petretto E, Dong R, Harper L, Kamesh L, Heward JM, Gough SC, de Smith A, Blakemore AI et al. 2007. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat Genet* **39**(6): 721–723.

Faridi RM, Agrawal S. 2011. Killer immunoglobulin-like receptors (KIRs) and HLA-C allorecognition patterns implicative of dominant activation of natural killer cells contribute to recurrent miscarriages. *Hum Reprod* **26**(2): 491–497.

Fawcett JA, Innan H. 2013. The role of gene conversion in preserving rearrangement hotspots in the human genome. *Trends Genet* **29**(10): 561–568.

Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL, Reinisch W, Teml A, Schwab M, Lichter P et al. 2006. A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am J Hum Genet* **79**(3): 439-448.

Feng W, Bedows E, Norton SE, Ruddon RW. 1996. Novel covalent chaperone complexes associated with human chorionic gonadotropin beta subunit folding intermediates. *J Biol Chem* **271**(31): 18543–18548.

Feng W, Matzuk MM, Mountjoy K, Bedows E, Ruddon RW, Boime I. 1995. The asparagine-linked oligosaccharides of the human chorionic gonadotropin beta subunit facilitate correct disulfide bond pairing. *J Biol Chem* **270**(20): 11851–11859.

Ganguly A, McKnight RA, Raychaudhuri S, Shin BC, Ma Z, Moley K, Devaskar SU. 2007. Glucose transporter isoform-3 mutations cause early pregnancy loss and fetal growth restriction. *American journal of physiology Endocrinology and metabolism* **292**(5): E1241–1255.

Ghosh D, Ezashi T, Ostrowski MC, Roberts RM. 2003. A central role for Ets-2 in the transcriptional regulation and cyclic adenosine 5'-monophosphate responsiveness of the human chorionic gonadotropin-beta subunit gene. *Mol Endocrinol* **17**(1): 11–26.

Giovangrandi Y, Parfait B, Asheuer M, Olivi M, Lidereau R, Vidaud M, Bieche I. 2001. Analysis of the human CGB/LHB gene cluster in breast tumors by real-time quantitative RT-PCR assays. *Cancer Lett* **168**(1): 93–100.

Girirajan S, Brkanac Z, Coe BP, Baker C, Vives L, Vu TH, Shafer N, Bernier R, Ferrero GB, Silengo M et al. 2011a. Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet* **7**(11): e1002334.

Girirajan S, Campbell CD, Eichler EE. 2011b. Human copy number variation and complex genetic disease. *Annu Rev Genet* **45**: 203–226.

Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, Filipink RA, McConnell JS, Angle B, Meschino WS et al. 2012. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *New Engl J Med* **367**(14): 1321–1331.

Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C et al. 2010. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* **42**(3): 203–209.

Glessner JT, Smith AV, Panossian S, Kim CE, Takahashi N, Thomas KA, Wang F, Seidler K, Harris TB, Launer LJ et al. 2013. Copy number variations in alternative splicing gene networks impact lifespan. *PloS one* **8**(1): e53846.

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ et al. 2005. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307**(5714): 1434–1440.

Graef T, Moesta AK, Norman PJ, Abi-Rached L, Vago L, Older Aguilar AM, Gleimer M, Hammond JA, Guethlein LA, Bushnell DA et al. 2009. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. *J Exp Med* **206**(11): 2557–2572.

Graubert TA, Cahan P, Edwin D, Selzer RR, Richmond TA, Eis PS, Shannon WD, Li X, McLeod HL, Cheverud JM et al. 2007. A high-resolution map of segmental DNA copy number variation in the mouse genome. *PLoS Genet* **3**(1): e3.

Grayson BL, Smith ME, Thomas JW, Wang L, Dexheimer P, Jeffrey J, Fain PR, Nanduri P, Eisenbarth GS, Aune TM. 2010. Genome-wide analysis of copy number variation in type 1 diabetes. *PloS one* **5**(11): e15393.

Griswold AJ, Ma D, Cukier HN, Nations LD, Schmidt MA, Chung RH, Jaworski JM, Salyakina D, Konidari I, Whitehead PL et al. 2012. Evaluation of copy number variations reveals novel candidate genes in autism spectrum disorder-associated pathways. *Hum Mol Genet* **21**(15): 3513–3523.

Group TECW. 2010. Europe the continent with the lowest fertility. *Hum Reprod Update* **16**(6): 590–602.

Guibourdenche J, Handschuh K, Tsatsaris V, Gerbaud P, Leguy MC, Muller F, Brion DE, Fournier T. 2010. Hyperglycosylated hCG is a marker of early human trophoblast invasion. *J Clin Endocrinol Metab* **95**(10): E240–244.

Guleria I, Sayegh MH. 2007. Maternal acceptance of the fetus: true human tolerance. *J Immunol* **178**(6): 3345–3351.

Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C, Uxa R, Keating S, Kingdom J, Weksberg R. 2008. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. *Dev Biol* **320**(1): 79–91.

Guryev V, Saar K, Adamovic T, Verheul M, van Heesch SA, Cook S, Pravenec M, Aitman T, Jacob H, Shull JD et al. 2008. Distribution and functional impact of DNA copy number variation in the rat. *Nat Genet* **40**(5): 538–545.

Hallast P, Rull K, Laan M. 2007. The evolution and genomic landscape of CGB1 and CGB2 genes. *Mol Cell Endocrinol* **260–262**: 2–11.

Hallast P, Saarela J, Palotie A, Laan M. 2008. High divergence in primate-specific duplicated regions: human and chimpanzee chorionic gonadotropin beta genes. *BMC Evol Biol* **8**: 195.

Hanna CW, McFadden DE, Robinson WP. 2013. DNA methylation profiling of placental villi from karyotypically normal miscarriage and recurrent miscarriage. *Am J Pathol* **182**(6): 2276–2284.

Hanna J, Goldman-Wohl D, Hamani Y, Avraham I, Greenfield C, Natanson-Yaron S, Prus D, Cohen-Daniel L, Arnon TI, Manaster I et al. 2006. Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. *Nat Med* **12**(9): 1065–1074.

Harris RA, Ferrari F, Ben-Shachar S, Wang X, Saade G, Van Den Veyver I, Facchinetti F, Aagaard-Tillery K. 2011. Genome-wide array-based copy number profiling in human placentas from unexplained stillbirths. *Prenat Diagn* **31**(10): 932–944.

Hawkins SM, Buchold GM, Matzuk MM. 2011. Minireview: The roles of small RNA pathways in reproductive medicine. *Mol Endocrinol* **25**(8): 1257–1279.

Hay DL. 1988. Placental histology and the production of human choriogonadotrophin and its subunits in pregnancy. *Br J Obstet Gynaecol* **95**(12): 1268–1275.

Henrichsen CN, Chaignat E, Reymond A. 2009a. Copy number variants, diseases and gene expression. *Hum Mol Genet* **18**(R1): R1–8.

Henrichsen CN, Vinckenbosch N, Zollner S, Chaignat E, Pradervand S, Schutz F, Ruedi M, Kaessmann H, Reymond A. 2009b. Segmental copy number variation shapes tissue transcriptomes. *Nat Genet* **41**(4): 424–429.

Hess AP, Hamilton AE, Talbi S, Dosiou C, Nyegaard M, Nayak N, Genbecev-Krtolica O, Mavrogianis P, Ferrer K, Kruessel J et al. 2007. Decidual stromal cell response to paracrine signals from the trophoblast: amplification of immune and angiogenic modulators. *Biol Reprod* **76**(1): 102–117.

Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, Claas FH, Walker JJ, Redman CW, Morgan L et al. 2010. Maternal activating KIRs protect against human reproductive failure mediated by fetal HLA-C2. *J Clin Invest* **120**(11): 4102–4110.

Hiby SE, Regan L, Lo W, Farrell L, Carrington M, Moffett A. 2008. Association of maternal killer-cell immunoglobulin-like receptors and parental HLA-C genotypes with recurrent miscarriage. *Hum Reprod* **23**(4): 972–976.

Hiby SE, Walker JJ, O'Shaughnessy K M, Redman CW, Carrington M, Trowsdale J, Moffett A. 2004. Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success. *J Exp Med* **200**(8): 957–965.

Hicks WM, Kim M, Haber JE. 2010. Increased mutagenesis and unique mutation signature associated with mitotic gene conversion. *Science* **329**(5987): 82–85.

Hirota Y, Cha J, Yoshie M, Daikoku T, Dey SK. 2011. Heightened uterine mammalian target of rapamycin complex 1 (mTORC1) signaling provokes preterm birth in mice. *Proc Natl Acad Sci U S A* **108**(44): 18073–18078.

Hoermann R, Spoettl G, Berger P, Mann K. 1995. Immunoreactive human chorionic gonadotropin beta core fragment in human pituitary. *Exp Clin Endocrinol Diabetes* **103**(5): 324–331.

Hollenberg AN, Pestell RG, Albanese C, Boers ME, Jameson JL. 1994. Multiple promoter elements in the human chorionic gonadotropin beta subunit genes distinguish their expression from the luteinizing hormone beta gene. *Mol Cell Endocrinol* **106**(1–2): 111–119.

Hollox EJ, Armour JA, Barber JC. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am J Hum Genet* **73**(3): 591–600.

Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Lascorz J, Rodijk-Olthuis D, van de Kerkhof PC, Traupe H, de Jongh G, den Heijer M et al. 2008. Psoriasis is associated with increased beta-defensin genomic copy number. *Nat Genet* **40**(1): 23–25.

Holtan SG, Creedon DJ, Haluska P, Markovic SN. 2009. Cancer and pregnancy: parallels in growth, invasion, and immune modulation and implications for cancer therapeutic agents. *Mayo Clin Proc* **84**(11): 985–1000.

Hu BS, Hu H, Zhu CY, Gu YL, Li JP. 2013. Overexpression of GOLPH3 is associated with poor clinical outcome in gastric cancer. *Tumour Biol* **34**(1): 515–520.

Huhtaniemi IT, Korenbrot CC, Jaffe RB. 1977. HCG binding and stimulation of testosterone biosynthesis in the human fetal testis. *J Clin Endocrinol Metab* **44**(5): 963–967.

Huth JR, Mountjoy K, Perini F, Ruddon RW. 1992. Intracellular folding pathway of human chorionic gonadotropin beta subunit. *J Biol Chem* **267**(13): 8870–8879.

Hviid TV, Hylenius S, Lindhard A, Christiansen OB. 2004. Association between human leukocyte antigen-G genotype and success of in vitro fertilization and pregnancy outcome. *Tissue Antigens* **64**(1): 66–69.

Jauniaux E, Jurkovic D. 2012. Placenta accreta: pathogenesis of a 20th century iatrogenic uterine disease. *Placenta* **33**(4): 244–251.

Jauniaux E, Poston L, Burton GJ. 2006. Placental-related diseases of pregnancy: Involvement of oxidative stress and implications in human evolution. *Hum Reprod Update* **12**(6): 747–755.

Jiang M, Savontaus ML, Simonsen H, Williamson C, Mullenbach R, Gromoll J, Terwort N, Alevizaki M, Huhtaniemi I. 2004. Absence of the genetic variant Val79Met in human chorionic gonadotropin-beta gene 5 in five European populations. *Mol Hum Reprod* **10**(10): 763–766.

Jiang W, Johnson C, Jayaraman J, Simecek N, Noble J, Moffatt MF, Cookson WO, Trowsdale J, Traherne JA. 2012. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. *Genome Res* **22**(10): 1845–1854.

Jivraj S, Anstie B, Cheong YC, Fairlie FM, Laird SM, Li TC. 2001. Obstetric and neonatal outcome in women with a history of recurrent miscarriage: a cohort study. *Hum Reprod* **16**(1): 102–106.

Jivraj S, Rai R, Underwood J, Regan L. 2006. Genetic thrombophilic mutations among couples with recurrent miscarriage. *Hum Reprod* **21**(5): 1161–1165.

Kajihara T, Uchino S, Suzuki M, Itakura A, Brosens JJ, Ishihara O. 2010. Human chorionic gonadotropin confers resistance to oxidative stress-induced apoptosis in decidualizing human endometrial stromal cells. *Fertil Steril* **95**(4): 1302–1307.

Kaufmann P, Black S, Huppertz B. 2003. Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia. *Biol Reprod* **69**(1): 1–7.

Kayisli UA, Selam B, Guzeloglu-Kayisli O, Demir R, Arici A. 2003. Human chorionic gonadotropin contributes to maternal immunotolerance and endometrial apoptosis by regulating Fas-Fas ligand system. *J Immunol* **171**(5): 2305–2313.

Kehrer-Sawatzki H, Cooper DN. 2007. Structural divergence between the human and chimpanzee genomes. *Hum Genet* **120**(6): 759–778.

Keikkala E, Vuorela P, Laivuori H, Romppanen J, Heinonen S, Stenman UH. 2013. First trimester hyperglycosylated human chorionic gonadotrophin in serum – A marker of early-onset preeclampsia. *Placenta*.

Kim MS, Gu BH, Song S, Choi BC, Cha DH, Baek KH. 2011. ITI-H4, as a biomarker in the serum of recurrent pregnancy loss (RPL) patients. *Molecular bioSystems* **7**(5): 1430–1440.

Kim YS, Kim MS, Lee SH, Choi BC, Lim JM, Cha KY, Baek KH. 2006. Proteomic analysis of recurrent spontaneous abortion: Identification of an inadequately expressed set of proteins in human follicular fluid. *Proteomics* **6**(11): 3445–3454.

Kolte AM, Nielsen HS, Moltke I, Degn B, Pedersen B, Sunde L, Nielsen FC, Christiansen OB. 2011. A genome-wide scan in affected sibling pairs with idiopathic recurrent miscarriage suggests genetic linkage. *Mol Hum Reprod* **17**(6): 379–385.

Kolte AM, Steffensen R, Nielsen HS, Hviid TV, Christiansen OB. 2010. Study of the structure and impact of human leukocyte antigen (HLA)-G-A, HLA-G-B, and HLA-G-DRB1 haplotypes in families with recurrent miscarriage. *Hum Immunol* **71**(5): 482–488.

Korhonen J, Stenman UH, Ylostalo P. 1994. Serum human chorionic gonadotropin dynamics during spontaneous resolution of ectopic pregnancy. *Fertil Steril* **61**(4): 632–636.

Kovalevskaya G, Birken S, Kakuma T, Ozaki N, Sauer M, Lindheim S, Cohen M, Kelly A, Schlatterer J, O'Connor JF. 2002. Differential expression of human chorionic gonadotropin (hCG) glycosylation isoforms in failing and continuing pregnancies: preliminary characterization of the hyperglycosylated hCG epitope. *J Endocrinol* **172**(3): 497–506.

Kovalevsky G, Gracia CR, Berlin JA, Sammel MD, Barnhart KT. 2004. Evaluation of the association between hereditary thrombophilias and recurrent pregnancy loss: a meta-analysis. *Arch Intern Med* **164**(5): 558–563.

Kovats S, Main EK, Librach C, Stubblebine M, Fisher SJ, DeMars R. 1990. A class I antigen, HLA-G, expressed in human trophoblasts. *Science* **248**(4952): 220–223.

Krieg SA, Fan X, Hong Y, Sang QX, Giaccia A, Westphal LM, Lathi RB, Krieg AJ, Nayak NR. 2012. Global alteration in gene expression profiles of deciduas from women with idiopathic recurrent pregnancy loss. *Mol Human Reprod* **18**(9): 442–450.

Kubiczak M, Walkowiak GP, Nowak-Markwitz E, Jankowska A. 2013. Human Chorionic Gonadotropin Beta Subunit Genes CGB1 and CGB2 are Transcriptionally Active in Ovarian Cancer. *International journal of molecular sciences* **14**(6): 12650–12660.

Kuningas M, Estrada K, Hsu YH, Nandakumar K, Uitterlinden AG, Lunetta KL, van Duijn CM, Karasik D, Hofman A, Murabito J et al. 2011. Large common deletions associate with mortality at old age. *Hum Mol Genet* **20**(21): 4290–4296.

Kusnierczyk P. 2013. Killer cell immunoglobulin-like receptor gene associations with autoimmune and allergic diseases, recurrent spontaneous abortion, and neoplasms. *Front Immunol* **4**: 8.

Laitinen T, Lokki ML, Tulppala M, Ylikorkala O, Koskimies S. 1991. Increased frequency of complement C4 'null' alleles in recurrent spontaneous abortions. *Hum Reprod* **6**(10): 1384–1387.

Lapthorn AJ, Harris DC, Littlejohn A, Lustbader JW, Canfield RE, Machin KJ, Morgan FJ, Isaacs NW. 1994. Crystal structure of human chorionic gonadotropin. *Nature* **369**(6480): 455–461.

Larsen EC, Christiansen OB, Kolte AM, Macklon N. 2013. New insights into mechanisms behind miscarriage. *BMC medicine* **11**: 154.

Le Marechal C, Masson E, Chen JM, Morel F, Ruszniewski P, Levy P, Ferec C. 2006. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nat Genet* **38**(12): 1372–1374.

Leach RE, Romero R, Kim YM, Chaiworapongsa T, Kilburn B, Das SK, Dey SK, Johnson A, Qureshi F, Jacques S et al. 2002. Pre-eclampsia and expression of heparin-binding EGF-like growth factor. *Lancet* **360**(9341): 1215–1219.

Ledig S, Ropke A, Wieacker P. 2010. Copy number variants in premature ovarian failure and ovarian dysgenesis. *Sex Dev* **4**(4–5): 225–232.

Lee AS, Gutierrez-Arcelus M, Perry GH, Vallender EJ, Johnson WE, Miller GM, Korbel JO, Lee C. 2008. Analysis of copy number variation in the rhesus macaque genome identifies candidate loci for evolutionary and human disease studies. *Hum Mol Genet* **17**(8): 1127–1136.

Lee CL, Chiu PC, Hautala L, Salo T, Yeung WS, Stenman UH, Koistinen H. 2013. Human chorionic gonadotropin and its free beta-subunit stimulate trophoblast invasion independent of LH/hCG receptor. *Mol Cell Endocrinol* **375**(1–2): 43–52.

Lee J, Oh J, Choi E, Park I, Han C, Kim do H, Choi BC, Kim JW, Cho C. 2007. Differentially expressed genes implicated in unexplained recurrent spontaneous abortion. *Int J Biochem Cell Biol* **39**(12): 2265–2277.

Lefranc MP, Hammarstrom L, Smith CI, Lefranc G. 1991. Gene deletions in the human immunoglobulin heavy chain constant region locus: molecular and immunological analysis. *Immunodefic Rev* **2**(4): 265–281.

Lempiainen A, Hotakainen K, Blomqvist C, Alfthan H, Stenman UH. 2012. Hyperglycosylated human chorionic gonadotropin in serum of testicular cancer patients. *Clin Chem* **58**(7): 1123–1129.
Li M, Huang SJ. 2009. Innate immunity, coagulation and placenta-related adverse pregnancy outcomes. *Thromb Res* **124**(6): 656–662.
Li N, Stephens M. 2003. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**(4): 2213–2233.
Li W, Zeng Chan W, Cui X, Xiao Feng L, Mao Sheng Y. 2010. Genome-wide screening for risk loci of idiopathic recurrent miscarriage in a Han Chinese population: a pilot study. *Reprod Sci* **17**(6): 578–584.
Lopata A, Hay DL. 1989. The potential of early human embryos to form blastocysts, hatch from their zona and secrete HCG in culture. *Hum Reprod* **4**(8 Suppl): 87–94.
Macklon NS, Geraedts JP, Fauser BC. 2002. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. *Hum Reprod Update* **8**(4): 333–343.
Malhotra D, McCarthy S, Michaelson JJ, Vacic V, Burdick KE, Yoon S, Cichon S, Corvin A, Gary S, Gershon ES et al. 2011. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron* **72**(6): 951–963.
Mamtani M, Anaya JM, He W, Ahuja SK. 2010. Association of copy number variation in the FCGR3B gene with risk of autoimmune diseases. *Genes Immun* **11**(2): 155–160.
Maston GA, Ruvolo M. 2002. Chorionic gonadotropin has a recent origin within primates and an evolutionary history of selection. *Mol Biol Evol* **19**(3): 320–335.
Matzuk MM, Krieger M, Corless CL, Boime I. 1987. Effects of preventing O-glycosylation on the secretion of human chorionic gonadotropin in Chinese hamster ovary cells. *Proc Natl Acad Sci U S A* **84**(18): 6354–6358.
Matthiesen L, Kalkunte S, Sharma S. 2012. Multiple pregnancy failures: an immunological paradigm. *Am J Reprod Immunol* **67**(4): 334–340.
McCarroll SA, Bradner JE, Turpeinen H, Volin L, Martin PJ, Chilewski SD, Antin JH, Lee SJ, Ruutu T, Storer B et al. 2009. Donor-recipient mismatch for common gene deletion polymorphisms in graft-versus-host disease. *Nat Genet* **41**(12): 1341–1344.
McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PI, Maller JB, Kirby A et al. 2008. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**(10): 1166–1174.
McLellan AS, Zimmermann W, Moore T. 2005. Conservation of pregnancy-specific glycoprotein (PSG) N domains following independent expansions of the gene families in rodents and primates. *BMC Evol Biol* **5**: 39.
Miller-Lindholm AK, Bedows E, Bartels CF, Ramey J, Maclin V, Ruddon RW. 1999. A naturally occurring genetic variant in the human chorionic gonadotropin-beta gene 5 is assembly inefficient. *Endocrinology* **140**(8): 3496–3506.
Miller-Lindholm AK, LaBenz CJ, Ramey J, Bedows E, Ruddon RW. 1997. Human chorionic gonadotropin-beta gene expression in first trimester placenta. *Endocrinology* **138**(12): 5459–5465.
Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK et al. 2011. Mapping copy number variation by population-scale genome sequencing. *Nature* **470**(7332): 59–65.
Mishra AK, Mahale SD, Iyer KS. 2003. Disulfide bonds Cys(9)-Cys(57), Cys(34)-Cys(88) and Cys(38)–Cys(90) of the beta-subunit of human chorionic gonadotropin are crucial for heterodimer formation with the alpha-subunit: experimental evidence

for the conclusions from the crystal structure of hCG. *Biochim Biophys Acta* **1645**(1): 49–55.
Moffett-King A. 2002. Natural killer cells and pregnancy. *Nature reviews Immunology* **2**(9): 656–663.
Moghraby JS, Tamim H, Anacan V, Al Khalaf H, Moghraby SA. 2010. HLA sharing among couples appears unrelated to idiopathic recurrent fetal loss in Saudi Arabia. *Hum Reprod* **25**(8): 1900–1905.
Montgomery GW, Zondervan KT, Nyholt DR. 2013. The future for genetic studies in reproduction. *Mol Human Reprod*.
Morgan FJ, Birken S, Canfield RE. 1975. The amino acid sequence of human chorionic gonadotropin. The alpha subunit and beta subunit. *J Biol Chem* **250**(13): 5247–5258.
Murray-Rust J, McDonald NQ, Blundell TL, Hosang M, Oefner C, Winkler F, Bradshaw RA. 1993. Topological similarities in TGF-beta 2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* **1**(2): 153–159.
Nagamani SC, Erez A, Ben-Zeev B, Frydman M, Winter S, Zeller R, El-Khechen D, Escobar L, Stankiewicz P, Patel A et al. 2013. Detection of copy-number variation in AUTS2 gene by targeted exonic array CGH in patients with developmental delay and autistic spectrum disorders. *Europ J Hum Genet* **21**(3): 343–346.
Nair RR, Khanna A, Singh K. 2013. Role of inflammatory proteins S100A8 and S100A9 in pathophysiology of recurrent early pregnancy loss. *Placenta* **34**(9): 824–827.
Nelissen EC, van Montfoort AP, Dumoulin JC, Evers JL. 2011. Epigenetics and the placenta. *Hum Reprod Update* **17**(3): 397–417.
Norwitz ER, Schust DJ, Fisher SJ. 2001. Implantation and the survival of early pregnancy. *New Engl J Med* **345**(19): 1400–1408.
Novakovic B, Saffery R. 2012. The ever growing complexity of placental epigenetics – role in adverse pregnancy outcomes and fetal programming. *Placenta* **33**(12): 959–970.
Ogasawara M, Aoki K, Okada S, Suzumori K. 2000. Embryonic karyotype of abortuses in relation to the number of previous miscarriages. *Fertil Steril* **73**(2): 300–304.
Pandey MK, Rani R, Agrawal S. 2005. An update in recurrent spontaneous abortion. *Arch Gynecol Obstet* **272**(2): 95–108.
Parrott AM, Sriram G, Liu Y, Mathews MB. 2011. Expression of type II chorionic gonadotropin genes supports a role in the male reproductive system. *Mol Cell Biol* **31**(2): 287–299.
Pelak K, Need AC, Fellay J, Shianna KV, Feng S, Urban TJ, Ge D, De Luca A, Martinez-Picado J, Wolinsky SM et al. 2011. Copy number variation of KIR genes influences HIV-1 control. *PLoS Biol* **9**(11): e1001208.
Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R et al. 2007. Diet and the evolution of human amylase gene copy number variation. *Nat Genet* **39**(10): 1256–1260.
Philipp T, Philipp K, Reiner A, Beer F, Kalousek DK. 2003. Embryoscopic and cytogenetic analysis of 233 missed abortions: factors involved in the pathogenesis of developmental defects of early failed pregnancies. *Hum Reprod* **18**(8): 1724–1732.
Pierce JG, Parsons TF. 1981. Glycoprotein hormones: structure and function. *Annu Rev Biochem* **50**: 465–495.
Pinto D Pagnamenta AT Klei L Anney R Merico D Regan R Conroy J Magalhaes TR Correia C Abrahams BS et al. 2010. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**(7304): 368–372.

Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD et al. 2000. Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* **24**(1): 84–87.

Rai R, Regan L. 2006. Recurrent miscarriage. *Lancet* **368**(9535): 601–611.

Rajcan-Separovic E, Diego-Alvarez D, Robinson WP, Tyson C, Qiao Y, Harvard C, Fawcett C, Kalousek D, Philipp T, Somerville MJ et al. 2010. Identification of copy number variants in miscarriages from couples with idiopathic recurrent pregnancy loss. *Hum Reprod* **25**(11): 2913–2922.

Red-Horse K, Zhou Y, Genbacev O, Prakobphol A, Foulk R, McMaster M, Fisher SJ. 2004. Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface. *J Clin Invest* **114**(6): 744–754.

Redman CW, Sargent IL. 2010. Immunology of pre-eclampsia. *Am J Reprod Immunol* **63**(6): 534–543.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W et al. 2006. Global variation in copy number in the human genome. *Nature* **444**(7118): 444–454.

Reimand J, Arak T, Vilo J. 2011. g:Profiler – a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res* **39**(Web Server issue): W307–315.

Revel A, Achache H, Stevens J, Smith Y, Reich R. 2011. MicroRNAs are associated with human embryo implantation defects. *Hum Reprod* **26**(10): 2830–2840.

Rey E, Kahn SR, David M, Shrier I. 2003. Thrombophilic disorders and fetal loss: a meta-analysis. *Lancet* **361**(9361): 901–908.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**(6): 276–277.

Robertson L, Wu O, Langhorne P, Twaddle S, Clark P, Lowe GD, Walker ID, Greaves M, Brenkel I, Regan L et al. 2006. Thrombophilia in pregnancy: a systematic review. *Br J Haematol* **132**(2): 171–196.

Roos S, Jansson N, Palmberg I, Saljo K, Powell TL, Jansson T. 2007. Mammalian target of rapamycin in the human placenta regulates leucine transport and is down-regulated in restricted fetal growth. *J Physiol* **582**(Pt 1): 449–459.

Rull K, Hallast P, Uuskula L, Jackson J, Punab M, Salumets A, Campbell RK, Laan M. 2008. Fine-scale quantification of HCG beta gene transcription in human trophoblastic and non-malignant non-trophoblastic tissues. *Mol Hum Reprod* **14**(1): 23–31.

Rull K, Laan M. 2005. Expression of beta-subunit of HCG genes during normal and failed pregnancy. *Hum Reprod* **20**(12): 3360–3368.

Rull K, Nagirnaja L, Laan M. 2012. Genetics of recurrent miscarriage: challenges, current knowledge, future directions. *Front Genet* **3**: 34.

Rull K, Tomberg K, Koks S, Mannik J, Mols M, Sirotkina M, Varv S, Laan M. 2013. Increased placental expression and maternal serum levels of apoptosis-inducing TRAIL in recurrent miscarriage. *Placenta* **34**(2): 141–148.

Salker M, Teklenburg G, Molokhia M, Lavery S, Trew G, Aojanepong T, Mardon HJ, Lokugamage AU, Rai R, Landles C et al. 2010. Natural selection of human embryos: impaired decidualization of endometrium disables embryo-maternal interactions and causes recurrent pregnancy loss. *PloS one* **5**(4): e10287.

Schumacher A, Heinze K, Witte J, Poloski E, Linzke N, Woidacki K, Zenclussen AC. 2013. Human chorionic gonadotropin as a central regulator of pregnancy immune tolerance. *J Immunol* **190**(6): 2650–2658.

Scott KL, Kabbarah O, Liang MC, Ivanova E, Anagnostou V, Wu J, Dhakal S, Wu M, Chen S, Feinberg T et al. 2009. GOLPH3 modulates mTOR signalling and rapamycin sensitivity in cancer. *Nature* **459**(7250): 1085–1090.

Sedman L, Padhukasahasram B, Kelgo P, Laan M. 2008. Complex signatures of locus-specific selective pressures and gene conversion on Human Growth Hormone/ Chorionic Somatomammotropin genes. *Hum Mutat* **29**(10): 1181–1193.

Shao W, Tang J, Song W, Wang C, Li Y, Wilson CM, Kaslow RA. 2007. CCL3L1 and CCL4L1: variable gene copy number in adolescents with and without human immunodeficiency virus type 1 (HIV-1) infection. *Genes Immun* **8**(3): 224–231.

Smith GR. 1988. Homologous recombination in procaryotes. *Microbiol Rev* **52**(1): 1–28.

Stenman UH, Alfthan H, Hotakainen K. 2004. Human chorionic gonadotropin in cancer. *Clin Biochem* **37**(7): 549–561.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**(4): 978–989.

Stirrat GM. 1990. Recurrent miscarriage. *Lancet* **336**(8716): 673–675.

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C et al. 2007. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**(5813): 848–853.

Stuart PE, Huffmeier U, Nair RP, Palla R, Tejasvi T, Schalkwijk J, Elder JT, Reis A, Armour JA. 2012. Association of beta-defensin copy number and psoriasis in three cohorts of European origin. *J Invest Dermatol* **132**(10): 2407–2413.

Zhang F, Gu W, Hurles ME, Lupski JR. 2009. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* **10**: 451–481.

Zhao L, Bracken MB, Dewan AT. 2013. Genome-Wide Association Study of Pre-Eclampsia Detects Novel Maternal Single Nucleotide Polymorphisms and Copy-Number Variants in Subsets of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study Cohort. *Ann Hum Genet*.

Zhao L, Triche EW, Walsh KM, Bracken MB, Saftlas AF, Hoh J, Dewan AT. 2012. Genome-wide association study identifies a maternal copy-number deletion in PSG11 enriched among preeclampsia patients. *BMC Pregnancy Childbirth* **12**: 61.

Zimmermann G, Ackermann W, Alexander H. 2012. Expression and production of human chorionic gonadotropin (hCG) in the normal secretory endometrium: evidence of CGB7 and/or CGB6 beta hCG subunit gene expression. *Biol Reprod* **86**(3): 87.

Zinaman MJ, Clegg ED, Brown CC, O'Connor J, Selevan SG. 1996. Estimates of human fertility and pregnancy loss. *Fertil Steril* **65**(3): 503–509.

Zygmunt M, Herr F, Keller-Schoenwetter S, Kunzi-Rapp K, Munstedt K, Rao CV, Lang U, Preissner KT. 2002. Characterization of human chorionic gonadotropin as a novel angiogenic factor. *J Clin Endocrinol Metab* **87**(11): 5290–5296.

Tam CW, Cheng AS, Ma RY, Yao KM, Shiu SY. 2006. Inhibition of prostate cancer cell growth by human secreted PDZ domain-containing protein 2, a potential autocrine prostate tumor suppressor. *Endocrinology* **147**(11): 5023–5033.

Teklenburg G, Salker M, Heijnen C, Macklon NS, Brosens JJ. 2010. The molecular basis of recurrent pregnancy loss: impaired natural embryo selection. *Mol Human Reprod* **16**(12): 886–895.

Tsampalas M, Gridelet V, Berndt S, Foidart JM, Geenen V, Perrier d'Hauterive S. 2010. Human chorionic gonadotropin: a hormone with immunological and angiogenic properties. *J Reprod Immunol* **85**(1): 93–98.

Uuskula L, Rull K, Nagirnaja L, Laan M. 2010. Methylation allelic polymorphism (MAP) in chorionic gonadotropin beta5 (CGB5) and its association with pregnancy success. *J Clin Endocrinol Metab* **96**(1): E199–207.

Vacic V, McCarthy S, Malhotra D, Murray F, Chou HH, Peoples A, Makarov V, Yoon S, Bhandari A, Corominas R et al. 2011. Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature* **471**(7339): 499–503.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, Nord AS, Kusenda M, Malhotra D, Bhandari A et al. 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**(5875): 539–543.

van Oppenraaij RH, Jauniaux E, Christiansen OB, Horcajadas JA, Farquharson RG, Exalto N. 2009. Predicting adverse obstetric outcome after early pregnancy events and complications: a review. *Hum Reprod Update* **15**(4): 409–421.

Wang JH, Chen XT, Wen ZS, Zheng M, Deng JM, Wang MZ, Lin HX, Chen K, Li J, Yun JP et al. 2012. High expression of GOLPH3 in esophageal squamous cell carcinoma correlates with poor prognosis. *PloS one* **7**(10): e45622.

Wang X, Jiang W, Zhang D. 2013. Association of 14-bp insertion/deletion polymorphism of HLA-G gene with unexplained recurrent spontaneous abortion: a meta-analysis. *Tissue Antigens* **81**(2): 108–115.

Vatin M, Burgio G, Renault G, Laissue P, Firlej V, Mondon F, Montagutelli X, Vaiman D, Serres C, Ziyyat A. 2012. Refined mapping of a quantitative trait locus on chromosome 1 responsible for mouse embryonic death. *PloS one* **7**(8): e43356.

Watson CT, Steinberg KM, Huddleston J, Warren RL, Malig M, Schein J, Willsey AJ, Joy JB, Scott JK, Graves TA et al. 2013. Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Am J Hum Genet* **92**(4): 530–546.

Wheeler E, Huang N, Bochukova EG, Keogh JM, Lindsay S, Garg S, Henning E, Blackburn H, Loos RJ, Wareham NJ et al. 2013. Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat Genet* **45**(5): 513–517.

Wilcox AJ, Baird DD, Weinberg CR. 1999. Time of implantation of the conceptus and loss of pregnancy. *New Engl J Med* **340**(23): 1796–1799.

Wilcox AJ, Weinberg CR, O'Connor JF, Baird DD, Schlatterer JP, Canfield RE, Armstrong EG, Nisula BC. 1988. Incidence of early loss of pregnancy. *New Engl J Med* **319**(4): 189–194.

Wilczynski JR. 2006. Immunological analogy between allograft rejection, recurrent abortion and pre-eclampsia – the same basic mechanism? *Hum Immunol* **67**(7): 492–511.

Witt CS, Goodridge J, Gerbase-Delima MG, Daher S, Christiansen FT. 2004. Maternal KIR repertoire is not associated with recurrent spontaneous abortion. *Hum Reprod* **19**(11): 2653–2657.

von Salome J, Gyllensten U, Bergstrom TF. 2007. Full-length sequence analysis of the HLA-DRB1 locus suggests a recent origin of alleles. *Immunogenetics* **59**(4): 261–271.

Wood CS, Hung CS, Huoh YS, Mousley CJ, Stefan CJ, Bankaitis V, Ferguson KM, Burd CG. 2012. Local control of phosphatidylinositol 4-phosphate signaling in the Golgi apparatus by Vps74 and Sac1 phosphoinositide phosphatase. *Mol Biol Cell* **23**(13): 2527–2536.

Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Sabatti C, Geurts van Kessel A, Brunner HG, Ophoff RA, Veltman JA. 2008. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. *Am J Hum Genet* **83**(4): 504–510.

Xue WC, Chan KY, Feng HC, Chiu PM, Ngan HY, Tsao SW, Cheung AN. 2004. Promoter hypermethylation of multiple genes in hydatidiform mole and choriocarcinoma. *The Journal of molecular diagnostics: JMD* **6**(4): 326–334.

Yang Y, Chung EK, Wu YL, Savelli SL, Nagaraja HN, Zhou B, Hebert M, Jones KN, Shu Y, Kitzmiller K et al. 2007. Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. *Am J Hum Genet* **80**(6): 1037–1054.

Yawata M, Yawata N, Draghi M, Little AM, Partheniou F, Parham P. 2006. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. *J Exp Med* **203**(3): 633–645.

Yuen RK, Avila L, Penaherrera MS, von Dadelszen P, Lefebvre L, Kobor MS, Robinson WP. 2009. Human placental-specific epipolymorphism and its association with adverse pregnancy outcomes. *PloS one* **4**(10): e7389.

SUMMARY IN ESTONIAN

Korduva raseduse katkemise genoomsed ja geneetilised riskitegurid

Raseduse katkemine ehk spontaanabort on kõige sagedasem rasedustüsistus ning kuni 15% kliinilisest diagnoositud rasedustest katkevad iseeneslikult. Ligi 3%-l viljakas eas paaridel leiab spontaanabort aset järjestikku kolm või enam korda, mille puhul diagnoositakse korduv raseduse katkemine (KRK). Järgnevate raseduste korral on KRK patsientidel suurenenud risk nii uue raseduse katkemiseks kui ka enneaegse sünnituse esinemiseks. KRK näol on tegemist heterogeense haigusega, mille teadaolevate riskitegurite spekter on lai, kuid umbes pooltel haigusjuhtudel ei ole tekkepõhjus endiselt määratletav. Selge põhjuseta ehk idiopaatilist KRK esinemise risk on patisentide lähi-sugulastel kaks korda kõrgem kui mitte-suguluses olevatel indiviididel viidates pärilikkuse kompnendile KRK patogeneesis. Idiopaatilise KRK geneetiliste riskide kaardistamiseks on tänaseks läbi viidud hulgaliselt uuringuid rohkem kui saja kandidaat-geeni põhjal, mille funktsioon varajase raseduses on teada. Sellele vaatamata ei ole siiani tuvastatud ühtegi sagedast geneetilist markerit, mis oleks spetsiifline KRK-le ning mis oleks rakendatav kliinilises diagnostikas ja ravis. Kaasaegsed kogu genoomi, transkriptoomi ja/või proteoomi-põhised metoodikad võimaldavad aga kompleksset lähenemist KRK pärilikkuskomponendi analüüsil ning tuvastada uusi geene ja mehhanisme, mis potentisaalselt mõjutavad varajase raseduse kulgu.

Inimese koorioni gonadotropiini (*human chorionic gonadotropin*, hCG), mida kutsutakse ka ‘raseduse hormooniks’, toodab blastotsüst juba 8-rakulises staadiumis ning selle olemasolu on embrüo pesastumiseks ja raseduse säilimiseks kriitilise tähtsusega. Heterodimeerse hCG spetsiifilisuse määrab β-alaühik, mille normist kõgem või madalam avaldumine on tuvastatud nii emakavälise raseduse, moolraseduse kui ka idiopaatilise KRK puhul. Seega on β-alaühikut kodeerivad geenid ja neis esinev geneetiline varieeruvus huvipakkuv kui potentsiaalne KRK riskitegur. hCGβ alaühikut kodeerivad neli duplikaat-geeni (*CGB*, *CGB7* ning kõige kõrgema avaldumisega *CGB5* ja *CGB8*), mis asuvad ühises geeniklastris 19. kromosoomil koos lutropiini β-alaühikut tootva (*LHB*) ja kahe hCGβ alaühikut mitte-kodeeriva (*CGB1* ja *CGB2*) geeniga. Siiani ei ole teostatud põhjalikke hCGβ geenide varieeruvuse uuringuid KRK kontekstis valdavalt metoodiliste takistuste tõttu, mis lähtuvad *LHB/CGB* geeniklastri keerulisest struktuurist, suurest DNA kordusjärjestuste arvust ning *CGB* geenide sarnasusest nii DNA (97–99%) kui valgu (98–100%) tasemel.

DNA koopiaarvu varieeruvus (*copy number variants*, CNV), mis hõlmab DNA struktuurseid ümberkorraldusi suurusega üle 50 aluspaari, on laiaulatuslik fenomen mõjutades kuni 30% inimese kogu genoomi järjestusest. Üksikuid CNV piirkondasid kui ka CNV-de summarset rikastatust inimese genoomis on

seostatud mitmete haiguste patogeneesiga nagu näiteks autism, skisofreenia, ülekaalulisus ja pikaealisus. Inimese reproduktiivhaiguste kontekstis on CNV-sid vähe uuritud ning idiopaatilise KRK puhul on CNV-de profiili analüüsi varasemalt teostatud ainult 24 aborteerunud platsenta proovis. Suuremahulise platsenta-põhise CNV-de analüüsi limiteerivaks faktoriks KRK puhul, mis leiab aset varajase raseduse käiugus, on aga kvaliteetse bioloogilise materjali raskendatud kättesaadavus. Kuna loote genoomi panustavad ema ja isa võrdselt, siis on alternatiivina võimalik hinnata vanemate geneetilist eelsoodumust KRK esinemiseks verest eraldatud DNA põhjal.

Käesoleva doktoritöö kirjanduse ülevaates on koondatud info tänaseks teadaolevatest idiopaatilise KRK riskiteguritest ning geneetilistest assotsiatsiooniuuringutest, mis on teostatud nii kogu genoomi kui ka üksikute markerite tasemel. Lisaks on kirjeldatud hCG hormooni, hCGβ alaühikut, *CGB* geenide funktsionaalseid ja struktuurseid omadusi ning ühtlasi CNV-de rolli nii kompleks-haiguste kui ka rasedustüsistuste kujunemisel.

Doktoritöö eksperimentaalses osas on püstitatud järgmised eesmärgid:

I. *LHB/CGB* geeniklastri kaardistamine ning detailne DNA järjestuse analüüs eesmärgiga tuvastada tegureid, mis mõjutavad piirkonna struktuurset ja evolutsioonilist dünaamikat
II. DNA ühenukleotiidsed variandid kõige rohkem hCG β-alaühikut tootvates geenides *CGB5* ja *CGB8* ning nende seos KRK esinemisega:
    - DNA ühenukleotiidsete variantide tuvastamine *CGB5* ja *CGB8* geenides ning juht-kontrolluuringu teostamine Eesti, Soome ja Taani valimites
    - *CGB5* ja *CGB8* geenides esinevate mitte-sünonüümsete mutatsioonide struktuurne ja funktsionaalne analüüs *in vitro*
III. DNA koopiaarvu varieeruvus (CNV) kui uus KRK geneetiline marker:
    - CNV-de genoomse profiili analüüs ning CNV-de poolt häiritud bioloogiliste radade tuvastamine
    - Uute CNV piirkondade identifitseerimine, mis suurendavad KRK kujunemise riski

Uurimistöö peamised tulemused on järgmised:

Ühenukleotiidete geneetilise markerite roll KRK esinemisel LHB/CGB geeniklastri näitel

1. *LHB/CGB* geeniklaster, mis koosneb ühest *LHB* ja kuuest *CGB* geenikoopiast inimesel, on kujunenud korduvate DNA ümberkorralduste tagajärjel, mida on tõenäoliselt soodustanud aktiivsed rekombinatsiooni alad piir-

konnas. *LHB/CGB* geenide varieeruvust on mõjutanud ühelt poolt geenikonversioon, mis levitab polümorfisme geenikoopiate vahel, ning teisalt funktsionaalsed piirangud, mis garanteerivad hCGβ geenide aktiivse avaldumise. Arvestades genoomse piirkonna dünaamikat võib eeldada, et *LHB/CGB* geeniklaster on ka käesolevalt struktuuri ja varieeruvuse poolest ebastabiilne.

2. Kõige rohkem hCGβ-t tootvatele geenidele *CGB5* ja *CGB8* on iseloomulik riskivariantide puudumine Eesti, Soome ja Taani KRK-ga patsientidel. *CGB5* promootor-alas esineb neljast polümorfismist koosnev motiiv, mis omab protektiivset efekti KRK esinemise suhtes ning mis on üle kandunud vastavast *CGB8* promootori piirkonnast geenikonversiooni teel. *CGB8* on omakorda kõige efektiivsem hCG β-alaühikut produtseeriv geenikoopia ning selle promootor-ala on balansseeriva valiku all.
3. Kolme mitte-sünonüümse *CGB5* ja *CGB8* geenis asuva mutatsiooni struktuurne ja funktsionaalne uuring tuvastas häirunud hCG hormooni assambleerumise p.Val56Leu aminohappe asenduse puhul (tekib 10% metsik-tüüpi hCG-st). Samas põhjustab p.Val56Leu asendus hormooni suuremat bioaktiivsust LH/CG retseptoriga seondumisel seega täielikult või osaliselt kompenseerides vähenenud hCG kogust. Kuna analüüsitud mutatsioonide puhul ei ole uuritud indiviidide seas tuvastatud homosügootseid kandjaid ning tõenäoliselt tolereeritakse *CGB5* ja *CGB8* geenides ainult „kerge" funktsionaalse mõjuga aminohappe muutusi, siis võib eeldada, et antud geenide hoidmine funktsionaalselt aktiivsetena on oluline edukaks raseduse kulgemiseks.

Globaalsete geneetiliste markerite roll KRK esinemisel CNV-de genoomse profiili näitel

1. CNV-de genoomne kaardistamine tuvastas, et KRK patisentidel on CNV-de summaarne ulatus ja pikkade (>100 kb) geenirikaste deletsioonide osakaal genoomi kohta suurem kui viljakatel kontrollidel. Ulatuslik CNV-de esinemine genoomis suurendab tõenäosust, et ümberkorraldused hõlmavad geene, mis on olulised varajases raseduse etapis.
2. KRK patsientidel esinevate CNV-de poolt ümberkorraldatud geenide funktsionaalse rikastatuse analüüsil ilmnes, et struktuursed ümberkorraldused mõjutavad kõige enam bioloogilisi radasid, mis on seotud immuunvastuse väljakujunemisega platsentas. Tulemused kinnitavad, et emapoolse balansseeritud immuntolerantsuse tekkel loote poolt esitletud isa antigeenidele on määrava tähstusega roll edukaks raseduse kulgemiseks.
3. Enamlevinud CNV piirkondade-põhine uuring Eesti ja Taani valimites tuvastas uue geneetilise markeri, mille esinemisega kaasneb suurem KRK risk naispartneritel. Mitme koopiana esinev duplikatsioon 5p13.3 kromosomaalses piirkonnas hõlmab kahte geeni *PDZD2* ja *GOLPH3*, mis on kõige

aktiivsemalt avaldunud platsentas ning mida pole varasemalt seostatud varajase rasedusega.

Vaatamata laialdastele uuringutele KRK geneetiliste põhjuste selgitamiseks, ei ole siiani tuvastatud haiguse-spetsiifilisi markereid. Raseduse katkemise riski oluliselt mõjutavate geenide puhul, nagu seda on hCGβ duplikaat-geenid, on ilmselt põhjuseks loote elujõulisust oluliselt häirivate geneetiliste variandite evolutsioon negatiivse valiku surve all. Üha kasvava andmestiku põhjal on pakutud, et üksikute suure efektiga mutatsioonide ja polümorfsete markerite asemel võivad KRK kujunemist mõjutada väikese efektiga geneetiliste variantide kombinatsioonid või esineb igal KRK perekonnal oma unikaalne riskitegur.

DNA koopiaarvu varieeruvuse analüüs kogu genoomi tasemel on võimaldanud hüpoteesi-vaba lähenemist ja uute KRK-seoseliste lookuste tuvastamist ning on kinnitanud DNA struktuursete ümberkorralduste rolli varajase raseduse katkemisel. Ühtlasi rõhutavad käesoleva doktoritöö raames esitatud tulemused KRK geneetilise etioloogia heterogeensust.

ACKNOWLEDGEMENTS

Out of clutter, find simplicity,
From discord, find harmony,
In the middle of difficulty,
lies opporunity.

Albert Einstein

On the path towards this PhD thesis, occasionally paved with clutter and discord, I owe many thanks to my supervisor, Professor Maris Laan, for shedding light on opportunities, inspiring with countless ideas and for providing a multitude of interesting research objectives.

I am grateful to all my wonderful colleagues throughout the years for always providing help and advice whenever needed. My special thanks go to the girls from the 'Girls' room', Marina Grigorova and Laura Kasak, but also to former members of our research group and my dear friends Liis Uusküla-Reimand and Kärt Tomberg for sharing the PhD experience and for being there despite the distance.

The research of my thesis is predominantly focusing on human genetics with a touch of medical sciences that I cannot claim to be expert of. Thus I acknowledge the most capable multi-tasker Dr. Kristiina Rull for all the discussions and insights on the medical aspects of scientific research. And for being true to her profession by persistently encouraging us not to waste time with scientific degrees and become pregnant.

Within the framework of my PhD studies, I was presented with an opportunity to work in Professor Ilpo T. Huhtaniemi's research group in the Institute of Reproductive and Developmental Biology, Imperial College London, in the vibrant city of London. I am grateful to Professor Huhtaniemi for this valuable experience and appreciate Hellevi Peltoketo, Kim C. Jonas and other group members for all the help and know-how.

For the last but not least, I owe everything to the pillars of my life – my family and closest friends who have always supported me in everything and anything.

This work was financially supported by personal stipends from Kristjan Jaak Stipend Program, Graduate School in Biomedicine and Biotechnology, Ernst Jaakson memorial fund, Seitsmenda Samba Fond of University of Tartu, Margot M. ja Herbert R. Linn stipend fund and Estonian Students Fund, USA.

PUBLICATIONS

CURRICULUM VITAE

Name: Liina Nagirnaja
Date of birth: 24.12.1980
Address: Institute of Molecular and Cell Biology
Riia 23, Tartu 51010
E-mail: nagir@ut.ee
Current position: Researcher, Institute of Molecular and Cell Biology

Education:

2005–2012	PhD student in gene technology, Human Molecular Genetics research group, Institute of Molecular and Cell Biology, University of Tartu
2003–2005	MSc studies in gene technology, Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu
1999–2003	BSc studies in biology (*cum laude*), University of Tartu
1988–1999	Miina Härma Gymnasium, Tartu

Working Experience:

Since 2011	Researcher, Institute of Molecular and Cell Biology
2009–2010	Visiting researcher, Institute of Reproductive and Developmental Biology, Imperial College London, United Kingdom
2005–2010	Instructor of course "Genetics and evolution", The Gifted and Talented Development Centre, University of Tartu
2008	Specialist, Estonian Biocentre
2007–2008	Instructor in the Estonian delegation, International Junior Science Olympiad (2007, Taipei, Taiwan; 2008, Changwon, South-Korea)
2006–2009	Committee member, Estonian Junior Science Olympiad
2006–2007	Member of the Faculty of Science and Technology Council, University of Tartu
2005–2008	Committee member, Estonian Biology Olympiad
2002	Administration officer in the Reception Committee, University of Tartu

Main Fields of Research:
The role of DNA single nucleotide variants in the human luteinizing hormone and chorionic gonadotropin β-subunit (*LHB/CGB*) gene family and the contribution of DNA copy number variants (CNVs) in modifying the genetic predisposition to recurrent miscarriages.

Publications:

Nagirnaja L, Palta P, Kasak L, Rull K, Christiansen OB, Nielsen HS, Steffensen R, Esko T, Remm M, Laan M. (2014). Structural genomic variation as risk factor for idiopathic recurrent miscarriage. (*Manuscript submitted*)

Rull K, Christiansen, OB, Nagirnaja L, Steffensen R, Margus T, Laan M. (2013). A modest but significant effect of *CGB5* gene promoter polymorphisms in modulating the risk of recurrent miscarriage. Fertility and Sterility, 99(7), 1930–1936

Nagirnaja L, Venclovas Č, Rull K, Jonas KC, Peltoketo H, Christiansen OB, Kairys V, Kivi G, Steffensen R, Huhtaniemi IT, Laan M. (2012). Structural and functional analysis of rare missense mutations in human chorionic gonadotrophin β-subunit. Molecular Human Reproduction. Aug;18(8):379–90

Rull K, Nagirnaja L, Laan M. (2012). Genetics of recurrent miscarriage: challenges, current knowledge, future directions. Frontiers in Genetics, 3:34

Uusküla L, Rull K, Nagirnaja L, Laan M. (2011). Methylation Allelic Polymorphism (MAP) in *Chorionic Gonadotropin {beta}5* (*CGB5*) and Its Association with Pregnancy Success. Journal of Clinical Endocrinology and Metabolism, Jan;96(1):E199–207

Nagirnaja L, Rull K, Uusküla L, Hallast P, Grigorova M, Laan M. (2010). Genomics and genetics of gonadotropin beta-subunit genes: Unique *FSHB* and duplicated *LHB/CGB* loci. Molecular and Cellular Endocrinology, 329(1–2), 4–16

Rull K, Nagirnaja L, Ulander V.-M, Kelgo P, Margus T, Kaare M, Aittomäki K, Laan M (2008). Chorionic Gonadotropin Beta gene variants are associated with recurrent miscarriage in two European populations. Journal of Clinical Endocrinology and Metabolism, 93(12), 4697–4706

Hallast P, Nagirnaja L, Margus T, Laan M. (2005). Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. Genome Research, 15(11), 1535–46

Scholarships and Awards:

2012 Kristjan Jaak stipend, Archimedes Foundation
2011 Runner-up for The Elsiever *Trophoblast Research* Award 2011 at "IFPA (International Federation of Placenta Associations) Meeting 2011", Geilo, Norway
2011 Kristjan Jaak stipend, Archimedes Foundation
2010 Kristjan Jaak stipend, Archimedes Foundation
2009 Estonian Students Fund, USA
2009 Ernst Jaakson memorial fund scholarship, Tartu University Foundation
2007 Kristjan Jaak stipend, Archimedes Foundation
2006 Stipend of Seitsmenda Samba Fond, University of Tartu
2006 Stipend of Margot M. ja Herbert R. Linn, Estonian World Council

2005 II Prize for MA thesis, Estonian National Contest for Young Scientists at university level
2005 Stipend of Tartu Raefond, University of Tartu
2005 Kristjan Jaak stipend, Archimedes Foundation
2004 Harald Raudsepp's Stipend of Estonian-Revelia Academic Fund
2003 Diploma for B.Sc. thesis, National Contest of Students' Scientific Research

Other Scientific Activities:
Since 2001 Member of the Estonian Society of Human Genetics

Supervised dissertations:
2010–2012 Laura Kasak, MSc studies, Institute of Molecular and Cell Biology
Since 2012 Diana Nõmmemees, BSc studies, Institute of Molecular and Cell Biology

Interests:
Since 2005 Member of Mountaineering Club "Firn"

ELULOOKIRJELDUS

Nimi: Liina Nagirnaja
Sünniaeg: 24.12.1980
Aadress: TÜ Molekulaar- ja Rakubioloogia Instituut
Riia 23, Tartu 51010
E-post: nagir@ut.ee
Praegune töökoht: Teadur, TÜ Molekulaar- ja Rakubioloogia Instituut

Haridus:

2005–2012	PhD õpingud, geenitehnoloogia eriala, Inimese Molekulaargeneetika töögrupp, Molekulaar- ja Rakubioloogia Instituut, Tartu Ülikool
2003–2005	MSc (teadusmagister), geenitehnoloogia eriala, Biotehnoloogia õppetool, Molekulaar- ja Rakubioloogia Instituut, Tartu Ülikool
1999–2003	BSc bioloogia eriala (*cum laude*), Tartu Ülikool
1988–1999	Miina Härma Gümnaasium, Tartu

Töökogemus:

Alates 2011	Teadur, Molekulaar- ja Rakubioloogia Instituut
2009–2010	Külalisteadlane, Institute of Reproductive and Developmental Biology, Imperial College London, London, Suurbritannia
2005–2010	TÜ Teaduskooli kursuse „Geneetika ja evolutsioon" koordinaator
2008	Spetsialist, Eesti Biokeskus
2007–2008	Loodusteaduste Olümpiaadi rahvusvahelise vooru Eesti delegatsiooni liige (2007 a., Taipei, Taiwan; 2008 a., Changwon, Lõuna-Korea)
2006–2009	Loodusteaduste Olümpiaadi žürii liige
2006–2007	Tartu Ülikooli Bioloogia-geograafiateaduskonna Nõukogu liige
2005–2008	Eesti Bioloogiaolümpiaadi žürii liige
2002	Üliõpilaste vastuvõtu korraldaja, Tartu Ülikool

Peamised uurimisvaldkonnad:
Inimese luteiniseeriva hormooni ja koorion gonadotropiini beeta-subühiku geeniperekonna ning DNA struktruusete ümberkorralduste potentsiaalne seotus korduvate spontaanabortidega.

Publikatsioonide loetelu:

Nagirnaja L, Palta P, Kasak L, Rull K, Christiansen OB, Nielsen HS, Steffensen R, Esko T, Remm M, Laan M. (2014). Structural genomic variation as risk factor for idiopathic recurrent miscarriage. (*esitatud avaldamiseks*)

Rull K, Christiansen, OB, Nagirnaja L, Steffensen R, Margus T, Laan M. (2013). A modest but significant effect of *CGB5* gene promoter polymorphisms in modulating the risk of recurrent miscarriage. Fertility and Sterility, 99(7), 1930–1936

Nagirnaja L, Venclovas Č, Rull K, Jonas KC, Peltoketo H, Christiansen OB, Kairys V, Kivi G, Steffensen R, Huhtaniemi IT, Laan M. (2012). Structural and functional analysis of rare missense mutations in human chorionic gonadotrophin β-subunit. Molecular Human Reproduction. Aug;18(8):379–90

Rull K, Nagirnaja L, Laan M. (2012). Genetics of recurrent miscarriage: challenges, current knowledge, future directions. Frontiers in Genetics, 3:34

Uusküla L, Rull K, Nagirnaja, L, Laan M. (2011). Methylation Allelic Polymorphism (MAP) in *Chorionic Gonadotropin {beta}5* (*CGB5*) and Its Association with Pregnancy Success. Journal of Clinical Endocrinology and Metabolism, Jan; 96(1): E199–207

Nagirnaja L, Rull K, Uusküla L, Hallast P, Grigorova M, Laan M. (2010). Genomics and genetics of gonadotropin beta-subunit genes: Unique *FSHB* and duplicated *LHB/CGB* loci. Molecular and Cellular Endocrinology, 329(1–2), 4–16

Rull K, Nagirnaja L, Ulander V-M, Kelgo P, Margus T, Kaare M, Aittomäki K, Laan M (2008). Chorionic Gonadotropin Beta gene variants are associated with recurrent miscarriage in two European populations. Journal of Clinical Endocrinology and Metabolism, 93(12), 4697–4706

Hallast P, Nagirnaja L, Margus T, Laan M (2005). Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. Genome Research, 15(11), 1535–46

Saadud stipendiumid ja auhinnad:

2012 Kristjan Jaagu nimeline stipendium, Sihtasutus Archimedes

2011 II koht postri konkursil The Elsevier *Trophoblast Research* Award 2011, konverents "IFPA (International Federation of Placenta Associations)Meeting2011", Geilo, Norra

2011 Kristjan Jaagu nimeline stipendium, Sihtasutus Archimedes

2010 Kristjan Jaagu nimeline stipendium, Sihtasutus Archimedes

2009 Eesti Üliõpilaste Toetusfond, USA

2009 Ernst Jaaksoni mälestusfondi stipendium, TÜ Sihtasutus

2007 Kristjan Jaagu nimeline stipendium, Sihtasutus Archimedes

2006 TÜ Seitsmenda Samba Fondi stipendium

2006 Margot M. ja Herbert R. Linna stipendium, Ülemaailmne Eesti Kesknõukogu

2005 II preemia, „Üliõpilaste teadustööde riiklik konkurss", Haridus- ja Teadusministeerium/Sihtasutus Archimedes
2005 Tartu Raefondi stipendium, TÜ Sihtasutus
2005 Kristjan Jaagu nimeline stipendium, Sihtasutus Archimedes
2004 Harald Raudsepa nimeline stipendium, Estonian-Revelia Academic Fund
2003 Diplom, „Üliõpilaste teadustööde riiklik konkurss", Haridus- ja Teadusministeerium

Muu teaduslik organisatsiooniline ja erialane tegevus:
Alates 2001 Eesti Inimesegeneetika Ühingu liige

Juhendatud väitekirjad:
2010–2012 Laura Kasak, MSc väitekiri, TÜ Molekulaar- ja Rakubioloogia Instituut
Alates 2012 Diana Nõmmemees, BSc õpingud, TÜ Molekulaar- ja Rakubioloogia Instituut

Huvialad
Alates 2005 Alpiklubi „Firn" liige.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro.* Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous grassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.
102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.

103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.
106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.
122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.

123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.
126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007. 112 p.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.
142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.

143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.
147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2009, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.

162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.
166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.

182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp.** Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.
187. **Virve Sõber.** The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro.** The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold.** Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert.** Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu.** Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik.** ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber.** Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper.** Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak.** Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo.** Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel.** Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus.** Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius.** Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv.** Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk.** Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.

202. **Arno Põllumäe.** Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht.** Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus.** Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.
206. **Kessy Abarenkov.** PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.
207. **Marina Grigorova.** Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar.** The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild.** Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag.** Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.
212. **Aleksei Lulla.** Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.
213. **Mari Järve.** Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers. Tartu, 2012, 119 p.
214. **Ott Scheler**. The application of tmRNA as a marker molecule in bacterial diagnostics using microarray and biosensor technology. Tartu, 2012, 93 p.
215. **Anna Balikova**. Studies on the functions of tumor-associated mucin-like leukosialin (CD43) in human cancer cells. Tartu, 2012, 129 p.
216. **Triinu Kõressaar.** Improvement of PCR primer design for detection of prokaryotic species. Tartu, 2012, 83 p.
217. **Tuul Sepp.** Hematological health state indices of greenfinches: sources of individual variation and responses to immune system manipulation. Tartu, 2012, 117 p.
218. **Rya Ero.** Modifier view of the bacterial ribosome. Tartu, 2012, 146 p.
219. **Mohammad Bahram.** Biogeography of ectomycorrhizal fungi across different spatial scales. Tartu, 2012, 165 p.
220. **Annely Lorents.** Overcoming the plasma membrane barrier: uptake of amphipathic cell-penetrating peptides induces influx of calcium ions and downstream responses. Tartu, 2012, 113 p.

221. **Katrin Männik.** Exploring the genomics of cognitive impairment: whole-genome SNP genotyping experience in Estonian patients and general population. Tartu, 2012, 171 p.
222. **Marko Prous.** Taxonomy and phylogeny of the sawfly genus *Empria* (Hymenoptera, Tenthredinidae). Tartu, 2012, 192 p.
223. **Triinu Visnapuu.** Levansucrases encoded in the genome of *Pseudomonas syringae* pv. tomato DC3000: heterologous expression, biochemical characterization, mutational analysis and spectrum of polymerization products. Tartu, 2012, 160 p.
224. **Nele Tamberg.** Studies on Semliki Forest virus replication and pathogenesis. Tartu, 2012, 109 p.
225. **Tõnu Esko.** Novel applications of SNP array data in the analysis of the genetic structure of Europeans and in genetic association studies. Tartu, 2012, 149 p.
226. **Timo Arula.** Ecology of early life-history stages of herring *Clupea harengus membras* in the northeastern Baltic Sea. Tartu, 2012, 143 p.
227. **Inga Hiiesalu.** Belowground plant diversity and coexistence patterns in grassland ecosystems. Tartu, 2012, 130 p.
228. **Kadri Koorem.** The influence of abiotic and biotic factors on small-scale plant community patterns and regeneration in boreonemoral forest. Tartu, 2012, 114 p.
229. **Liis Andresen.** Regulation of virulence in plant-pathogenic pectobacteria. Tartu, 2012, 122 p.
230. **Kaupo Kohv.** The direct and indirect effects of management on boreal forest structure and field layer vegetation. Tartu, 2012, 124 p.
231. **Mart Jüssi.** Living on an edge: landlocked seals in changing climate. Tartu, 2012, 114 p.
232. **Riina Klais.** Phytoplankton trends in the Baltic Sea. Tartu, 2012, 136 p.
233. **Rauno Veeroja.** Effects of winter weather, population density and timing of reproduction on life-history traits and population dynamics of moose (*Alces alces*) in Estonia. Tartu, 2012, 92 p.
234. **Marju Keis.** Brown bear (*Ursus arctos*) phylogeography in northern Eurasia. Tartu, 2013, 142 p.
235. **Sergei Põlme.** Biogeography and ecology of *alnus*- associated ectomycorrhizal fungi – from regional to global scale. Tartu, 2013, 90 p.
236. **Liis Uusküla.** Placental gene expression in normal and complicated pregnancy. Tartu, 2013, 173 p.
237. **Marko Lõoke.** Studies on DNA replication initiation in *Saccharomyces cerevisiae.* Tartu, 2013, 112 p.
238. **Anne Aan.** Light- and nitrogen-use and biomass allocation along productivity gradients in multilayer plant communities. Tartu, 2013, 127 p.
239. **Heidi Tamm.** Comprehending phylogenetic diversity – case studies in three groups of ascomycetes. Tartu, 2013, 136 p.

240. **Liina Kangur.** High-Pressure Spectroscopy Study of Chromophore-Binding Hydrogen Bonds in Light-Harvesting Complexes of Photosynthetic Bacteria. Tartu, 2013, 150 p.
241. **Margus Leppik.** Substrate specificity of the multisite specific pseudouridine synthase RluD. Tartu, 2013, 111 p.
242. **Lauris Kaplinski.** The application of oligonucleotide hybridization model for PCR and microarray optimization. Tartu, 2013, 103 p.
243. **Merli Pärnoja.** Patterns of macrophyte distribution and productivity in coastal ecosystems: effect of abiotic and biotic forcing. Tartu, 2013, 155 p.
244. **Tõnu Margus.** Distribution and phylogeny of the bacterial translational GTPases and the Mqsr/YgiT regulatory system. Tartu, 2013, 126 p.
245. **Pille Mänd**. Light use capacity and carbon and nitrogen budget of plants: remote assessment and physiological determinants. Tartu, 2013, 128 p.
246. **Mario Plaas.** Animal model of Wolfram Syndrome in mice: behavioural, biochemical and psychopharmacological characterization. Tartu, 2013, 144 p.
247. **Georgi Hudjašov.** Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations. Tartu, 2013, 115 p.
248. **Mari Lepik**. Plasticity to light in herbaceous plants and its importance for community structure and diversity. Tartu, 2013, 102 p.
249. **Ede Leppik**. Diversity of lichens in semi-natural habitats of Estonia. Tartu, 2013, 151 p.
250. **Ülle Saks.** Arbuscular mycorrhizal fungal diversity patterns in boreonemoral forest ecosystems. Tartu, 2013, 151 p.
251. **Eneli Oitmaa**. Development of arrayed primer extension microarray assays for molecular diagnostic applications. Tartu, 2013, 147 p.
252. **Jekaterina Jutkina.** The horizontal gene pool for aromatics degradation: bacterial catabolic plasmids of the Baltic Sea aquatic system. Tartu, 2013, 121 p.
253. **Helen Vellau.** Reaction norms for size and age at maturity in insects: rules and exceptions. Tartu, 2014, 132 p.
254. **Randel Kreitsberg.** Using biomarkers in assessment of environmental contamination in fish – new perspectives. Tartu, 2014, 107 p.
255. **Krista Takkis.** Changes in plant species richness and population performance in response to habitat loss and fragmentation.Tartu, 2014, 135 p.