DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
338

NATALIA PERVJAKOVA

Genomic imprinting in complex traits

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
338

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
338

NATALIA PERVJAKOVA

Genomic imprinting in complex traits

UNIVERSITY OF TARTU
Press

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

This dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Molecular Biomedicine on March 29, 2018 by the Council of the Institute of Molecular Cell Biology, University of Tartu.

Supervisors: Reedik Mägi, PhD
Estonian Genome center, Institute of Genomics, University of Tartu

Prof Andres Metspalu, MD, PhD,
Chair of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
Estonian Genome Center, Institute of Genomics, University of Tartu

Reviewer: Prof. Ain Heinaru, PhD
Chair of Genetics, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

Opponent: Prof. Stephan Beck, PhD
Chair of Medical Genomics, University College London, UCL Cancer Institute, London (UK)

Commencement: Room No. 105, 23B Riia St., Tartu, on June 22, 2018, at 10:15 pm.

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

This research was funded by EU H2020 grants 692145, 633589, Estonian Research Council Grant IUT20-60, IUT24-6, and European Union through the European Regional Development Fund Project No. 2014-2020.4.01.15-0012 GENTRANSMED, Tartu Ülikooli ASTRA projekt PER ASPERA, NIASC, Estonian Center of Genomics/Roadmap II (project No. 2014-2020.4.01.16-0125). Data analyses were carried out in part in the High Performance Computing Center of University of Tartu.

ISSN 1024-6479
ISBN 978-9949-77-734-1 (print)
ISBN 978-9949-77-735-8 (pdf)

Copyright: Natalia Pervjakova, 2018

University of Tartu Press
www.tyk.ee

“Somewhere, something incredible
is waiting to be known.”

Carl Sagan

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS 9
ABBREVIATIONS 12
INTRODUCTION 13
1. REVIEW OF THE LITERATURE 15
  1.1. Parent-of-origin effects 15
  1.2. Features of genomic imprinting 17
  1.3. Three major phases in life cycle of genomic imprinting 18
    1.3.1. Establishment of imprinting.. 18
    1.3.2. Maintenance of imprinting 20
    1.3.3. Erasure of imprinting 21
  1.4. Dynamics of genomic imprints 23
    1.4.1. Imprinted genes with both active copies 24
    1.4.2. Genes that have imprinted and bi-allelic expression within the same organism 26
  1.5. Germline DMRs as an indicator of imprinted expression 28
    1.5.1. Placenta-specific germline DMRs 30
    1.5.2. Transiently methylated germline DMRs 33
  1.6. Post-natal loss of imprinted expression 34
2. AIMS OF THE STUDY 37
3. RESULTS AND DISCUSSION 38
  3.1. Imprinted genes and imprinting control regions are stable and show predominant intermediate methylation (Ref. I) 38
    3.1.1. Description of cohort and materials 38
    3.1.2. Detection of methylation patterns for imprinted and non-imprinted genes 39
    3.1.3. Equality of variances 40
    3.1.4. The location of intermediately methylated probes 41
    3.1.5. Germline DMRs are not ubiquitously methylated across somatic tissues 41
  3.2. Parent-of-origin effects in *PTPRD* gene affect the level of triglycerides in medium very low density lipoprotein NMR metabolite (Ref. II) 44
    3.2.1. Description of cohorts and methods 44
    3.2.2. Meta-analysis revealed an effect of *PTPRD* gene to VLDL lipoprotein 45
    3.2.3. Allele-specific expression analysis 46
  3.3. Parent-of-origin allelic association affecting BMI and Age at Menarche complex traits (Ref. III and Ref. IV) 47
    3.3.1. Description of cohort and materials 48
    3.3.2. Association analysis revealed enrichment in imprinted genes 48

CONCLUSIONS 50
SUMMARY IN ESTONIAN 52
REFERENCES 54
ACKNOWLEDGMENTS 65
PUBLICATIONS 67
CURRICULUM VITAE 152
ELULOOKIRJELDUS 156

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, referred to in the text by Roman numerals (Ref. I to Ref. IV):

I **Pervjakova N,** Kasela S, Morris AP, Kals M, Metspalu A, Lindgren CM, Salumets A, Mägi R. 2016. **Imprinted genes and imprinting control regions show predominant intermediate methylation in adult somatic tissues.** *Epigenomics. 2016 Jun;8(6):789-99. doi: 10.2217/epi.16.8.*

II **Pervjakova N,** Kukushkina V, Haller T, Kasela S, Joensuu A, Kristiansson K, Annilo T, Perola M, Salomaa V, Jousilahti P, Metspalu A, Mägi R. 2018. **Genome-wide analysis of NMR metabolites revealed parent-of-origin effect on triglycerides in medium VLDL in *PTPRD* gene.** *Biomark Med. 2018 Mar 14. doi: 10.2217/bmm-2018-0020.*

III Clive J. Hoggart, Giulia Venturini, Massimo Mangino, Felicia Gomez, Giulia Ascari, Jing Hua Zhao, Alexander Teumer, Thomas W. Winkler, **Natalia Tšernikova**, Jian'an Luan, Evelin Mihailov, Georg B. Ehret, Weihua Zhang, David Lamparter, Tõnu Esko, Aurelien Macé, Sina Rüeger, Pierre-Yves Bochud, Matteo Barcella, Yves Dauvilliers, Beben Benyamin, David M. Evans, Caroline Hayward, Mary F. Lopez, Lude Franke, Alessia Russo, Iris M. Heid, Erika Salvi, Sailaja Vendantam, Dan E. Arking, Eric Boerwinkle, John C. Chambers, Giovanni Fiorito, Harald Grallert, Simonetta Guarrera, Georg Homuth, Jennifer E. Huffman, David Porteous, Generation Scotland Consortium, The LifeLines Cohort study, The GIANT Consortium, Darius Moradpour, Alex Iranzo, Johannes Hebebrand, John P. Kemp, Gert J. Lammers, Vincent Aubert, Markus H. Heim, Nicholas G. Martin, Grant W. Montgomery, Rosa Peraita-Adrados, Joan Santamaria, Francesco Negro, Carsten O. Schmidt, Robert A. Scott, Tim D. Spector, Konstantin Strauch, Henry Völzke, Nicholas J. Wareham, Wei Yuan, Jordana T. Bell, Aravinda Chakravarti, Jaspal S. Kooner, Annette Peters, Giuseppe Matullo, Henri Wallaschofski, John B. Whitfield, Fred Paccaud, Peter Vollenweider, Sven Bergmann, Jacques S. Beckmann, Mehdi Tafti, Nicholas D. Hastie, Daniele Cusi, Murielle Bochud, Timothy M. Frayling, Andres Metspalu, Marjo-Riitta Jarvelin, André Scherag, George Davey Smith, Ingrid B. Borecki, Valentin Rousson, Joel N. Hirschhorn, Carlo Rivolta, Ruth J. F. Loos, Zoltán Kutalik. 2014. **Novel approach identifies SNPs in *SLC2A10* and *KCNK9* with evidence for parent-of-origin effect on body mass index.** *PLoS Genet. 2014 Jul 31;10(7):e1004508. doi: 10.1371/journal.pgen.1004508.*

IV Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E, Ang WQ, Corre T, Cousminer DL, Feenstra B, Franceschini N, Ganna A, Johnson AD, Kjellqvist S, Lunetta KL, McMahon G, Nolte IM, Paternoster L, Porcu E, Smith AV, Stolk L, Teumer A, **Tšernikova N**, Tikkanen E, Ulivi S, Wagner EK, Amin N, Bierut LJ, Byrne EM, Hottenga JJ, Koller DL, Mangino M, Pers TH, Yerges-Armstrong LM, Zhao JH, Andrulis IL, Anton-Culver H, Atsma F, Bandinelli S, Beckmann MW, Benitez J, Blomqvist C, Bojesen SE, Bolla MK, Bonanni B, Brauch H, Brenner H, Buring JE, Chang-Claude J, Chanock S, Chen J, Chenevix-Trench G, Collée JM, Couch FJ, Couper D, Coveillo AD, Cox A, Czene K, D'adamo AP, Smith GD, De Vivo, Demerath EW, Dennis J, Devilee P, Dieffenbach AK, Dunning AM, Eiriksdottir G, Eriksson JG, Fasching PA, Ferrucci L, Flesch-Janys D, Flyger H, Foroud T, Franke L, Garcia ME, García-Closas M, Geller F, de Geus EE, Giles GG, Gudbjartsson DF, Gudnason V, Guénel P, Guo S, Hall P, Hamann U, Haring R, Hartman CA, Heath AC, Hofman A, Hooning MJ, Hopper JL, Hu FB, Hunter DJ, Karasik D, Kiel DP, Knight JA, Kosma VM, Kutalik Z, Lai S, Lambrechts D, Lindblom A, Mägi R, Magnusson PK, Mannermaa A, Martin NG , Masson G, McArdle PF, McArdle WL, Melbye M, Michailidou K, Mihailov E, Milani L, Milne RL, Nevanlinna H, Neven P, Nohr EA, Oldehinkel AJ, Oostra BA, Palotie A, Peacock M, Pedersen NL, Peterlongo P, Peto J, Pharoah PD, Postma DS, Pouta A, Pylkäs K, Radice P, Ring S, Rivadeneira F, Robino A, Rose LM, Rudolph A, Salomaa V, Sanna S, Schlessinger D, Schmidt MK, Southey MC, Sovio U, Stampfer MJ, Stöckl D, Storniolo AM, Timpson NJ, Tyrer J, Visser JA, Vollenweider P, Völzke H, Waeber G, Waldenberger M, Wallaschofski H, Wang Q, Willemsen G, Winqvist R, Wolffenbuttel BH, Wright MJ, Australian Ovarian Cancer Study, GENICA Network, kConFab, LifeLines Cohort Study, InterAct Consortium, Early Growth Genetics (EGG) Consortium, Boomsma DI, Econs MJ, Khaw KT, Loos RJ, McCarthy MI, Montgomery GW, Rice JP, Streeten EA, Thorsteinsdottir U, van Duijn CM, Alizadeh BZ, Bergmann S, Boerwinkle E, Boyd HA, Crisponi L, Gasparini P, Gieger C, Harris TB, Ingelsson E, Järvelin MR, Kraft P, Lawlor D, Metspalu A, Pennell CE, Ridker PM, Snieder H, Sørensen TI, Spector TD, Strachan DP, Uitterlinden AG, Wareham NJ, Widen E, Zygmunt M, Murray A, Easton DF, Stefansson K, Murabito JM, Ong KK. **Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche.** *Nature. 2014 Oct 2;514(7520):92-97. doi: 10.1038/nature13545*

The publications listed above have been reprinted with the permission of the copyright owners.

My contributions to the listed publications were as follows:

Ref. I Participated in study design, conducting the experiments, analysed the data, and wrote the manuscript.

Ref. II Participated in study design, conducting the experiments, analysed the data, and wrote the manuscript.

Ref. III Analysed the Estonian Biobank data in both discovery and replication phase, participated in the critical review of the paper.

Ref. IV Analysed the Estonian Biobank data, participated in the critical review of the paper.

ABBREVIATIONS

AAEEs	Antagonistic allele expression effects
ASE	Allele-specific expression
BMI	Body mass index
CoEEs	Allele co-expression effects
DAEEs	Differential allele expression effects
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DRN	Dorsal raphe nucleus
EGCUT	Estonian Genome Center, University of Tartu
FINRISK	Finnish population survey on risk factors on chronic, noncommunicable diseases
gDMR	Germline DMR
GWAS	Genome-wide association study
hPSCs	Human pluripotent stem cells
ICM	Inner Cell Mass
ICR	Imprinting Control Region
iPS	Induces pluripotent stem cells
LRT	Likelihood ratio test
MAF	Minor allele frequency
mRNA	Messenger RNA
NMR	Nuclear magnetic resonance
PGCs	Primordial germ cells
POEs	Parent-of-Origin Effects
sDMR	Somatic DMR
SNP	Single nucleotide polymorphism
SWAN	Subset-quantile Within Array Normalization
tDMR	Transiently methylated germline DMR
TE	Trophectoderm
TSS	Transcription start site
UTR	Untranslated region
WAMIDEX	Web atlas of murine genomic imprinting and differential expression
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine

INTRODUCTION

The human genome contains ca. 20,000 protein coding genes and up to 150 of them are validated or are suspected to be imprinted. The process of imprinting itself was discovered nearly 30 years ago and for a long time it was suggested as a unique mechanism that causes the differences in expression depending on allelic origin. Being established in the germline, genomic imprinting was thought to be stable and constant within an organism. Later, single evidences from different animal and human studies have demonstrated that the expression of imprinted genes may depend on the specific tissue and the developmental stage of an organism. Moreover, specific epigenomic effects, such as maternal genetic effects, were found to mimic genomic imprinting. Also, recent studies have demonstrated the overall tendency that expression dependent on parental alleles is much more widespread than previously thought. It all lead to combining genomic imprinting and other allele-specific effects into the one group of effects called parent-of-origin effects becoming a good practise for the researchers nowadays.

In this thesis, specifically in the literature overview, I will focus mainly on the process of genomic imprinting. Specifically, I will provide an overview of the topics that are most relevant to the occurrence of tissue-specificity and the non-ubiquitous distribution of imprinted expression across an organism. Also, in the literature review, I will present the evidence for the plasticity and dynamics of genomic imprinting. It is important to note that most of the examples come from mouse studies, although my preference is for examples that are from genes for which imprinted expression is conserved between mouse and human genomes.

The experimental part of this thesis consists of two sections that address the tissue-specific imprinting question and effect of imprinted genes on complex traits. To test the hypotheses raised in this study we used different datasets, including the genome-wide association analysis data from international consortia and a panel of post-mortem tissues from donors. Interestingly, at the time of publishing the first paper related to the occurrence of tissue-specificity in imprinted genes, we didn't have enough data to confirm our hypotheses that would explain imprinted expression based on the methylation pattern of the germline DMR. One year later, by the time of writing, several studies performed by independent groups confirmed our earlier hypothesis, making a tremendous impact in this area. With the help of these studies I obtained a vision of this process, which I present in the literature overview.

Exploring the second hypothesis of how imprinted genes modulate the complex traits, you may find that I am only in the beginning of this path. My thesis is predominated by experiments that resulted in discoveries of associations between imprinted genes and complex traits, such as age at menarche, BMI and blood metabolites. These analyses are very important in the light of the further research, where imprinted genes should be examined for the actual expression in tissues related to the specific phenotype and their gene dosage

effect should be studied more closely. I personally feel that the question of how tissue-specificity arises in the human genome is still a *terra incognita*, and future research is required to provide a deeper understanding of the establishment and regulation of imprinted genes throughout the entirety of the human life.

1. REVIEW OF THE LITERATURE

1.1. Parent-of-origin effects

Parent-of-origin effects (POEs) comprise a group of epigenetic phenomena modulating or affecting different complex traits such as pre- and postnatal development, survival and growth of an organism, cognitive abilities, metabolic processes, and several human diseases and other complex traits. Specifically, POEs refer to a phenomenon in which the gene expression pattern depends on the parental allelic origin, caused by a difference in the methylation state of the two alleles (Lawson et al., 2013; Reik and Walter, 2001). POEs are caused by several different mechanisms that, in the end, have a similar phenotypic outcome, which is reflected in altered gene expression, thereby making it challenging to uncover the process behind. Nevertheless such effects may be considered as an evolutionary driving mechanism of human genome organisation allowing regulation of gene dosage at specific moments of development, growing and aging (Mohammad et al., 2010; Peters, 2014; Prickett and Oakey, 2012).

The process of genomic imprinting can be viewed as a main cause for parent-of-origin effects. It is mainly associated with achieving the non-equal influence of parental genomes to the offspring (Reik and Walter, 2001). Genomic imprinting is first established in the germline, and results in mainly monoallelic gene expression according to the parental origin of the alleles (Li and Sasaki, 2011). It is the process that has been widely observed in analyses of complex traits, when the effect of epigenetic variation on the genetic architecture of particular phenotypes is explored. The process of genomic imprinting in mammals was discovered nearly 30 years ago in a series of nuclear transplantation experiments (Surani et al., 1984). During these experiments, mouse uniparental embryos that had only one out of the two sets of parental chromosomes were created. In parallel, an additional mouse model was designed to have uniparental disomies, which means that they inherited specific chromosomes from one of the parents. Both experiments clearly demonstrated that the function of certain genes depends on parental origin, and molecular errors in such inheritance are associated with variety of processes (Surani et al., 1984). Uniparental inheritance of these regions, later referred to as imprinted, revealed the importance of this process in pre- and postnatal development, including growth and ability to survive in the first hours after birth, behaviour, maternal care, metabolic processes and many others (Peters, 2014).

Parent-of-origin effects are often considered as being equivalent to the process of genomic imprinting. However, there are other phenomena that can cause the appearance of such effects. Even if we only concentrate attention on these two processes, and admit their synonymy, continuing active study of genomic imprinting brings more novel evidence that goes against the initial viewpoint of uniformity of this process, and points, instead, to a more complex structure of POEs (Peters, 2014). Recently discovered diverse allelic effects

found in in neonatal brain that cause mosaics of mutant and wild-type cell for heterozygotes may easily lead to the appearance of POE-like phenotypic outcomes (Huang et al., 2017). The other type of POEs may arise due to differential rates of mutational transmission, which lead to a higher likelihood of disease-associated mutation transmission from one parent to the offspring, as demonstrated in myotonic dystrophy type-1 disorder (Pearson, 2003; Tomé et al., 2011). Additional examples include exclusively transmitted genetic factors, such as mitochondrial DNA or the Y chromosome, and biased transmissions associated with meiosis onset and *de novo* germline genomic rearrangement (Kelsey and Feil, 2013a).

Maternal genetic effects are another category of POEs that often disguise as genomic imprinting. Maternal genetic effects are the mother's genetically mediated phenotype influences on the phenotype of the offspring via maternally provided environmental factors, and such effects may extend beyond the direct inheritance of alleles (Hager et al., 2008; Svensson et al., 2009). Non-human studies clearly demonstrate mimicry of the maternal genetic effects that produce the same patterns of phenotypic variation expected from an imprinting (Svensson et al., 2009). Mice studies demonstrated that the appearance of maternal effects in a heterozygous knock-out mouse model of anxiety, where heterozygous serotonin A1 receptor, *Htr1a*, mothers gave birth to offspring that did not inherit the mutation themselves, but presented an anxiety-like phenotype (Gleason et al., 2010). Similar paternal genetic effects are also expected to exist, but are assumed to be much less influential to the offspring, as the father does not provide such a strong environmental influence to the offspring.

Taken together, parent-of-origin effects are much more widespread than predicted by early studies. Considering parent-of-origin effects as a set of several processes that may arise due to different evolutionary processes, but have a similar phenotypic outcome at the expression level, would be the best strategy for a researcher who aims to catalogue and categorise the parent-of-origin effects.

In my thesis, I will focus only on one part of this huge phenomenon – on the process of genomic imprinting. Specifically, I will provide an overview of the topics that are most relevant to the occurrence of tissue-specificity and the non-ubiquitous distribution of imprinted expression across an organism. Also, in the literature review, I will present the evidence for the plasticity and dynamics of genomic imprinting. It is important to note that most of the examples come from mouse studies, although my preference is for examples that are from genes for which imprinted expression is conserved between mouse and human genomes. I personally feel that the question of how tissue-specificity arises in the human genome is still a *terra incognita,* and future research is required to provide a deeper understanding of the establishment and regulation of imprinted genes throughout the entirety of the human lifespan.

1.2. Features of genomic imprinting

Imprinted genes have several features that distinguish them from the genes with biallelic expression. One of the unique features of imprinted genes is the ability of the transcriptional machinery of a cell to enable the expression that is clearly denoted as paternal or maternal. Moreover, the mechanism of exclusive expression from one of the homologs is constant, and repeats after each division of a cell. Therefore, the process regulating imprinted expression has at least four properties: i) imprinted genes and their regulative sequences have a clear signal affecting the transcriptomic pathway by activating or repressing it; ii) an epigenetic signal is heritable in somatic lineage, allowing daughter cells to transmit the information regarding the parental origin of chromosomes; iii) an epigenetic signal is established in a period when maternal and paternal copies of DNA are clearly divided, i.e., during gametogenesis and shortly after post-fertilization; and iv) an epigenetic signal is erasable, allowing a new signal to be established, i.e., the maternally inherited signal is lost during spermatogenesis to mark contributing maternal DNA as paternal (Bartolomei and Ferguson-Smith, 2011). Taken together, genomic imprinting maintains the complex mechanism of establishment and cleavage of epigenetic marks that are tightly regulated and are necessary for this biological phenomenon.

Studies in mice demonstrated that nearly 80% of imprinted genes are located in clusters that contain between 2 and 15 genes, and vary in size from >100 kb to several megabases. To date, most of the clusters are described in the mouse genome, while orthologue clusters share the similar regulation arrangement between mice and humans (MouseBook Imprinting Catalogue). This clustered organization allows *cis*-regulation of monoallelic expression of an entire cluster via imprinting control regions (ICRs) (Barlow, 2011). These regions have been validated in a series of mouse knock-out experiments, and have been shown to occur in human epigenetic diseases, revealing the unique contribution of the respective genes (Chotalia et al., 2009; Sanz et al., 2008). In some studies, the equivalence of an ICR to a differentially methylated region (DMR) – the CpG rich region that exhibits parent-specific DNA methylation (Kelsey and Feil, 2013a) – can be seen, while other studies identify ICR as a combination of several DMRs (Prickett and Oakey, 2012). I would also favour the idea that ICR is the primary DMR established during germline development (gDMR), while an ICR itself may also harbour more secondary differentially methylated regions, also called somatic DMRs (sDMRs). Somatic DMRs are under the hierarchical influence of gDMR (Coombes et al., 2003; Kagami et al., 2010; Lopes et al., 2003; Sanchez-Delgado et al., 2016) for which methylation acquired after fertilisation (Dent and Isles, 2014). It is always important to note the origin of DMRs located within a cluster as their functional consequence can differ drastically. For example, some DMRs are established at early stage of parental germ cells and maintain imprinted status in all developmental stages and tissues of an embryo (Court et al., 2014a). Other DMRs, in contrast,

demonstrate multiple changes in their methylation pattern, which is associated with appearance of tissue-specific imprinted expression (Dent and Isles, 2014).

1.3. Three major phases in life cycle of genomic imprinting

The cycle consists of three major phases: i) establishment that takes place during gametogenesis; ii) maintenance after the fertilization process; and iii) erasure that occurs again in the germline, but already of a next generation organism (Figure 1). The last step is followed by establishment, thereby completing the cycle (Li and Sasaki, 2011).

1.3.1. Establishment of imprinting

It has been long hypothesized that imprints are established solely in the germline, as this is the time when two genomes are separated and allocated in the cell-specific environment (Reik and Dean, 2001). However it is also possible that some imprints are established later, during *de novo* demethylation, which takes place immediately after fertilization (Messerschmidt et al., 2014; Robertson et al., 1999). Overall, male and female establishment in germline occurs in completely different ways, resulting in distinct methylation profiles of a mature oocyte and sperm (Figure 1). Indeed, recent genome-wide methylation studies revealed a substantial difference in the pattern of two parental DNA molecules. In the male germline, the process of establishment starts in the prenatal mitotically arrested (G1/G0) prospermatogonia cells, after mouse embryonic day 14.5 (E14.5), with progressive *de novo* methylation (Davis et al., 2000; Kato et al., 2007; Kelsey and Feil, 2013a; Ueda et al., 2000). Later, the imprints will be fully established during the neonatal stage (Kelsey and Feil, 2013a). In contrast, female germline *de novo* DNA methylation occurs asynchronously during the postnatal oocyte growth phase of meiotic prophase I. The maternal imprints will already be fully established by the mature oocyte stage. Established imprints in the germline are next transmitted to the fertilized cell and maintained through the pre- and postnatal development (Davis et al., 2000; Kelsey and Feil, 2013a; Li and Sasaki, 2011; Ueda et al., 2000).

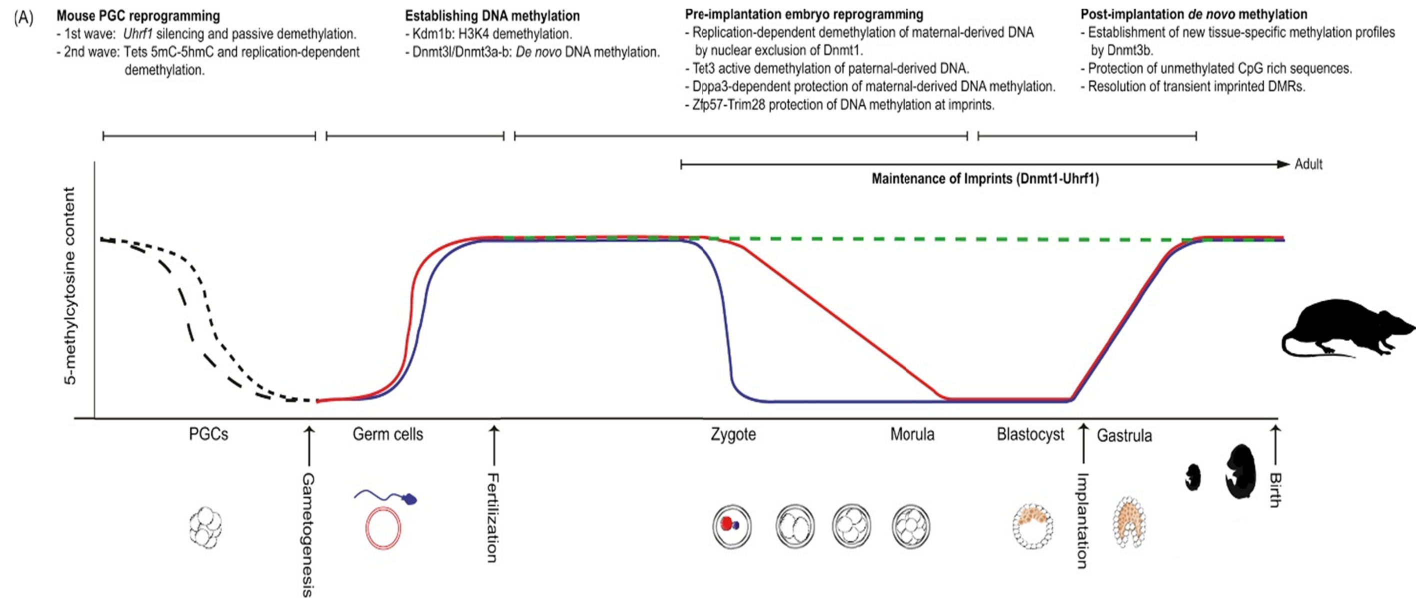


Figure 1. Erasure, establishment and maintenance of genomic imprints in mammals. Erasure starts with the migration of primordial germ cells (PGCs) into the genital ridge. Next, both maternal and paternal gDMRs become methylated in oocyte and sperm, respectively. Established imprints are maintained, despite global methylation reprogramming after fertilization. In a fertilized egg, paternal and maternal genomes undergo active and passive demethylation, respectively. *De novo* methylation occurs in the latest stage of pre-implantation, which results in the maintenance of imprints in somatic cells throughout the lifetime of the organism. The cycle ends with the re-establishment of the imprints, during gametogenesis, which will be transmitted to the next generation. Adapted from Monk (2015).

The most important question in this stage is how parental alleles are marked with their parental origin. The mechanism of recognition of parent-specific regions is also important to understand how tissue-specificity may arise. The main research focus has shifted to ICRs – key regulatory elements in controlling an entire cluster of imprinted genes, although other strategies have also been proposed. As mentioned above, ICR encompasses at least one DMR established in the germline (Court et al., 2014a; Prickett and Oakey, 2012). DMRs that are maintained throughout development are referred to as germline or primary DMRs (gDMR), whereas others that activate in later stages of cell development are called somatic or secondary DMRs (sDMR) (Court et al., 2014a; Prickett and Oakey, 2012). Imprinting control regions were proposed as the first target of recognition by methylation machinery, as knock-out methods have demonstrated that loss of such regions causes further inability to establish an imprinted pattern for multiple genes (Bartolomei and Ferguson-Smith, 2011). Several studies of primordial germ cells (PGCs) have demonstrated that the biased timing of establishing genomic imprints begins with the ICRs (Cantão et al., 2017; Kelsey and Feil, 2013b; von Meyenn and Reik, 2015). Most ICRs are methylated on the maternal allele (Bourc'his et al., 2001; Kaneda et al., 2004; Li and Sasaki, 2011; Prickett and Oakey, 2012). To date, it is known that maternally methylated regions use the tetrameric complex that consists of the *de novo* DNA methyltransferase DNMT3A and its stimulatory protein DNMT3L (Aapola et al., 2000; Hata et al., 2002; Kaneda et al., 2004; Kato et al., 2007). This complex methylates preferentially CpG sequences, which are 8–10 base pairs apart, are specific for maternally, but not paternally methylated loci (Ferguson-Smith and Greally, 2007). Another proposed strategy comprises of a cascade of oocyte-specific transcription across ICRs, followed by establishment or maintenance of open chromatin domains that are required for further establishment of DNA methylation. The study performed by Chotalia et al., demonstrated that the transcriptional machinery is directing the DNA methylation and not *vise versa* (Chotalia et al., 2009). Establishment of the imprinting of genes is a complex process that is yet to be described, although it is clear that this process includes many factors such as CpG spacing, variable timing of establishment, post-translational histone modifications and oocyte-specific timed transcription.

1.3.2. Maintenance of imprinting

After imprints are set, they begin to function via appropriate monoallelic expression that is maintained throughout the development of an organism. The first challenge that DMRs and imprints meet, at this stage, is genome-wide epigenetic reprogramming, where DNA methylation and chromatin modifications are erased and then reset (Figure 1). The demethylation process, itself, is also genome-specific and occurs both by active and by passive mechanisms. Passive demethylation is a result of DNMT1 inhibition during the cell proliferation process (Hassan et al., 2017; Wu and Zhang, 2010). In contrast, active demethylation is a replication independent process where TET proteins transform

5-hydrozymethylcytosine (5hmC) into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Ito et al., 2011; Wu and Zhang, 2017). TET proteins are also known to be part of TET-TDG pathway, where thymine DNA glycosylase (TDG) in the process of base excision repair mediates conversion of 5caC back to the C nucleotide (Wu and Zhang, 2017). The paternal DNA copy is actively demethylated for a few hours after birth, whereas the maternal copy is largely demethylated passively (Shen et al., 2014). At this step, both genomes are allocated in one cell, so that all DMRs on the maternal strand should be protected from any unexpected wave of active demethylation triggered by chromatin remodelling of paternal genome (Kelsey and Feil, 2013a; Li and Sasaki, 2011).

Interestingly, many CpG-rich sequences become DNA-methylated in oocyte and sperm, but only a fraction of them survive at early stages of development (Bartolomei and Ferguson-Smith, 2011; Kelsey and Feil, 2013a; Wu and Zhang, 2017). Therefore, the hypothesis of a selection process in pre-implantation embryos that specifically protects gDMRs from the demethylation is currently an active area of research. Such models include a combination of genetic factors that are not imprinted themselves, but provide the appropriate environment. One example is ZFP57 – a KRAB zinc finger protein, which is critical for DNA methylation maintenance (Abrink et al., 2001; Friedman et al., 1996). Li and his colleagues, in the setting of maternal and zygotic knock-out *Zfp57* experiments in mice, observed embryonic lethality and complete loss of methylation in numerous imprinted loci (Li et al., 2008). The follow-up of this study confirmed the loss of methylation in murine *Snrpn*, *Peg3*, *Zac1*, *Nespas*, and *H19* DMRs (Takahashi et al., 2016). It is also important to note that other studies have demonstrated conservation in the role of *ZFP57* in maintenance of DNA methylation between mice and humans (Amarasekera et al., 2014; Riso et al., 2016). Therefore, maintenance should be viewed as a hugely complex process associated with the cell biology of the maturing oocyte. One of the interesting findings is that hormone-induced superovulation also affects proper DNA maintenance at both maternal and paternal gDMRs, whereas chromatin organisation was also suggested to play a role in aberration of proteins in an oocyte (Fauque et al., 2007; Market-Velker et al., 2009). Taken together, the process of the maintenance is still not fully understood, but future research of these issues could explain the connection of the evolution of maintenance factors and their involvement in genomic imprinting.

1.3.3. Erasure of imprinting

The start of the erasure process is associated with specification of primordial germ cells from epiblast cells. Erasure of genomic imprints occurs together with global DNA demethylation and chromatin reorganisation, which ultimately leads to the generation of totipotency (Breindel et al., 2017; Hajkova et al., 2010; von Meyenn and Reik, 2015). Three parallel independent studies published in 2015 revealed a remarkable conservation of global methylation erasure

between mice and human genomes (Gkountela et al., 2015; Guo et al., 2015; Tang et al., 2015). Despite the suggested fundamental conservation of the erasure process, some methylation measurements and timelines varied between studies. On one hand, such differences can indicate batch effects caused by difficulties in isolating human cells and different protocols used in the study. On the other hand, it may point to the existence of a biological background of the process (von Meyenn and Reik, 2015). The most important differences were noted for the erasure time point of ICRs. Compared to mice, human ICRs appeared to have an earlier erasure that occurs before genital ridge colonisation (Tang et al., 2015), which may contribute to subsequent differences in imprinted patterns among species.

The process of erasure is also complex, although it is thought to be similar for male and female genomes (Figure 1). One of the mechanisms detected in PGCs, which is associated with loss of methylation, is activation-induced cytidine deaminase (AID) – a molecule that can catalyse modification of 5mC, leading indirectly to removal of methylation (Morgan et al., 2004; Popp et al., 2010). The idea behind the function of this enzyme, coded by the human *AICDA* gene, is in deamination of 5-methylcytosine (5mC) base followed by conversion into thymidine (T). As a result, T-G mismatches are recognized by the DNA repair mechanism, whereas thymidine is replaced by an adenine nucleotide and thus methylation cannot be restored at this position (Morgan et al., 2004). The AID enzyme has been demonstrated to act in both global demethylation and imprinted locus-specific demethylation of ICRs, i.e. mouse *H19* and *Kcnq1ot* loci (Popp et al., 2010). It is worth noting that the AID system is widely used in a variety of processes, such as B cell differentiation, mRNA processing, and negative regulation of methylation-dependent chromatin silencing – therefore future research might uncover new mechanisms of how AID enzymes mediate the demethylation in PGCs. Additionally, the evidence that AID-deficient PGCs are still able to continue the process of demethylation points to a more complex structure of this process.

Another strategy of demethylation has also been suggested through the conversion of 5mC to 5-hydroxymethylcytosine (5hmC) mediated by the translocation family proteins (TET), for which significantly increased expression was detected at mouse E11.5 and E12.5, when the imprinted DMRs undergo demethylation. Specifically, *Tet1* plays a critical role in the process of erasure, where lack of *Tet1* product results in placental, fetal and post-natal defects provided by dysregulation of imprinted *Peg3* and *Peg10* (Yamaguchi et al., 2013). Another Ten-Eleven Translocator protein, Tet2, was detected as part of the erasure complex in the mouse hybrid model (4n). The gene *Tet1* is a key factor in inducing 5-methylcytosine oxidation at ICRs, whereas *Tet2* is further required for the reprogramming of embryonic germ cells (Piccolo et al., 2013). The process of erasure of imprints is an area of active research, as better understanding of each particular step will significantly contribute to a variety of processes and techniques, including cell reprogramming technology for animal cloning and iPS cell generation.

1.4. Dynamics of genomic imprints

After the wave of genome-wide *de-novo* methylation, genomic imprints and their ICRs are maintained in somatic cells throughout the lifetime of an organism. More and more evidence demonstrates that the process of genomic imprinting might be dynamic, with varying degrees of monoallelic expression between tissues and developmental stages within the one organism. It is also important to note that genomic imprinting is not necessarily conserved between species, which makes studying tissue-specific imprinting even more challenging (Monk et al., 2006). For example, mouse, macaque and humans share the *IGF2R* gene, which is consistently imprinted in animal species, but not in humans (Cheong et al., 2015). In humans, imprinted expression of *IGF2R* is polymorphic and occurs in about 50% of individuals (Vu et al., 2004, 2006).

One of the first attempts to cluster expression patterns of imprinted genes, depending on tissue sample, was performed using the Web Atlas of Murine genomic Imprinting and Differential Expression (WAMIDEX) (Prickett and Oakey, 2012; Schulz et al., 2008). When comparing this map to human imprinted genes, it was noted that both human and mouse genomes maintain the relative preservation of imprinted patterns across tissues (Babak et al., 2015). The preservation of imprinted status for conserved genes across species does not guarantee the presence of the same mechanism of establishment. ICR controlling *PEG3* and *PLAGL1* imprinted expression has been found to be established in the post-fertilisation period in non-human primates and humans. However, in mice, imprinted expression of these genes is controlled by germline ICRs that appear at the one cell stage (Cheong et al., 2015). The major switches from monoallelic to biallelic expression are expected between embryonic, extraembryonic and adult organism tissues (Babak et al., 2015; Baran et al., 2015; Prickett and Oakey, 2012). Likewise, changes in imprinted expression can be roughly classified into three categories: i) genes that perform switches among the different tissues from one allele to another, while both copies remain functional; ii) genes that switch from imprinted expression in various types of tissues to biallelic expression in other types of tissues; and iii) genes that are known to have mainly biallelic expression, but were detected as imprinted in a certain tissue. Understanding the nature of tissue-specific changes in imprinted expression is very important in the context of diseases or traits affected by these genes, as the function of a gene may also depend on the allele expressed. The proposed classification of switches in imprinted expression was also based on the idea that they might correspond to the stages of development, and thus will add the value to the clarification of a diversity and the overall evolution pattern of imprinting.

1.4.1. Imprinted genes with both active copies

Mouse *Grb10* (Growth Factor Receptor Bound Protein 10) was one of the first genes for which tissue-specific imprinting was discovered. Although its expression pattern is well measured, the mechanisms regulating expression remain mostly unknown. The *Grb10* gene is unique in exhibiting imprinted expression from both parental alleles in a tissue-specific manner (Garfield et al., 2011). Moreover, each allele contributes to the distinct function of the gene product. Activation of the maternal copy in the placenta controls fetal growth and insulin signalling (Monk et al., 2009; Plasschaert and Bartolomei, 2015), while expression from the paternal copy in the central nervous system is transmitted through to adulthood and affects social behaviour (Garfield et al., 2011). Additional function of the *Grb10* gene in adult stem cells is associated with hematopoietic stem cell self-renovation and regeneration (Yan et al., 2016).

The tissue-specific expression of *Grb10* occurs during embryogenesis at ICRs, which regulates activation or repression of parental alleles through DNA methylation and histone modification (Dent and Isles, 2014). Maternal expression occurs in placenta and many of the adult tissues, while paternal expression occurs primary in neuronal tissues (Plasschaert and Bartolomei, 2015). Maternal transcripts arise from the major promoter, whereas paternal expression is detected from the three alternative promoters located next to the ICR, which contains brain-specific DMR and germline DMR (Figure 2) (Monk et al., 2006; Plasschaert and Bartolomei, 2015; Sanz et al., 2008).

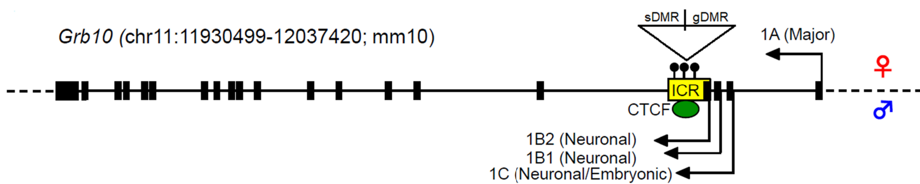


Figure 2. Schematic representation of the *Grb10* locus. Allele-specific expression designated as black arrows on both sides of the line (maternal expression above the line, paternal expression below the line). The yellow box indicates the ICR that consists of two DMRs, regulating the expression of the gene. The black dots located on the ICR indicate methylation marks allowing maternal expression in the majority of tissues of an adult organism. Adapted from Plasschaert and Bartolomei (2015).

ICR-regulation is among the major factors proposed to regulate tissue-specificity, although the interplay between DNA methylation, CTCF-binding protein, and histone-modifying enzymes has important supportive role in this mechanism (Hikichi et al., 2003; Kabir and Kazi, 2014; Lin et al., 2011). The absence of highly methylated epigenetic marks on the repressive H3K27me3 histone, associated with brain-specific somatic DMR, allows paternal expression in the developing neural lineage (Plasschaert and Bartolomei, 2015; Sanz et al., 2008).

For the same reason, the maternal allele is repressed by histone modification H3K9me3 and H4K20me3 (Sanz et al., 2008). It is also worth speculating that only one specific paternally expressed transcript of *Grb10* was detected in neurons, although there are more truncated isoforms that are expected to exist because of the presence of major and cell-specific promoters of *Grb10* (Mukhopadhyay et al., 2015; Plasschaert and Bartolomei, 2015; Stringer et al., 2012). These transcripts need to be validated, as the complete mechanism of tissue-specific regulation of this particular locus would not be fully understood without information regarding the final products of *Grb10*.

Another example of the ability to keep both alleles as a functional unit is the *IGF2* (Insulin-like Growth Factor 2) gene, for which expression from the maternal allele was demonstrated in the human brain, while other somatic tissues canonically express the paternal allele (Baran et al., 2015). It is interesting that, in the case of brain tissue, *H19* also remained expressed exclusively from the maternal copy (Baran et al., 2015; Renfree et al., 2013), which may point to an independent mechanism of regulation. Imprinted clusters are generally known to contain at least one non-coding RNA (ncRNA) (Edwards and Ferguson-Smith, 2007), playing a direct role in silencing mRNA genes within the cluster (Pauler and Barlow, 2006). *H19* is also a long non-coding RNA gene that controls imprinting of mouse *Igf2* and should be expressed from the allele opposite to *Igf2*, which means the maternal allele for the majority of tissues (Martinet et al., 2016; Monnier et al., 2013). On the other hand, the *Igf2r* cluster also contains the *Air* ncRNA, which is necessary to control mRNAs within the cluster (Pauler and Barlow, 2006) and its role has not been investigated in the human brain. Such activation of human *IGF2* alleles was explained by the presence of different promoters, and there was, therefore, also an attempt to identify more transcription start sites that are associated with maternal expression in the brain. The Database of Transcriptional Start Sites did not provide any suggestion for brain versus other tissues (Yamashita et al., 2012). The effect found for the human brain serves as an interesting example of non-canonical regulation within the specific tissue of a brain.

Further investigation of *IGF2* and *GRB10* is also important in the light of spontaneous abortion and assisted reproductive technology-conceived offspring. Higher methylation levels were demonstrated for these genes in case of assisted reproductive technology-conceived offspring, although more detailed analysis has suggested that aberrant methylation patterns, and thus imprinting defects of these genes, are associated with spontaneous abortion, which might not be due to assisted reproductive technology treatments (Zheng et al., 2013). Both *GRB10* and *IGF2* tend to be distinct from mainstream imprinted behaviour having unique features with unclear etiology of tissue-specificity, providing important evidence to further profiling the cell-type specificity of imprinted genes.

1.4.2. Genes that have imprinted and bi-allelic expression within the same organism

Imprinted genes may also exhibit monoallelic expression in embryonic or extra-embryonic tissues and biallelic expression in adult tissues. This type of genes was found in both human and mouse genomes. One example is the human *PLAGL1* (*ZAC*, *LOT1*, pleomorphic adenoma of the salivary gland gene like 1) gene that was found to be expressed exclusively from the paternal allele in most tissues during fetal development (Kamiya et al., 2000), including placenta (Arima et al., 2001), and throughout life (Kamiya et al., 2000). In contrast, human peripheral blood leukocytes and pancreas have demonstrated biallelic expression of this gene (Mackay et al., 2002). The tissue-specific phenomenon of the *PLAGL1* gene was explained by the presence of an alternative promoter (P2) located in within a second and unmethylated CpG island (Valleley et al., 2007). Indeed, the P2 promoter provides a biallelic expression, whereas paternal expression is always produced by the P1 promoter, but in lower concentrations, thus allowing simultaneous transcription of both products within a specific tissue (Valleley et al., 2007). In the mouse genome, the orthologue gene *Zac1* is conserved and was also denoted as imprinted with paternal expression in the majority of tissues of an adult organism (Piras et al., 2000; Smith et al., 2002). Biallelic expression in mice is more widespread and was observed in liver, kidney and skeletal muscle (Piras et al., 2000). According to the overall expression level from both P1 and P2 promoters in human liver (Valleley et al., 2007), it is possible that *PLAGL1* also has both transcripts in this tissue.

From an evolutionary perspective, the combination of monoallelic and biallelic expression or the situation where only the temporal switch to the biallelic state exists could be important in regulation of gene dosage. The original idea of gene dosage came from the example of the mouse *Kcnq1* gene that reverts from maternal to biallelic transcript in cardiac lineages between embryonic days E13.5 and E14.5 (Korostowski et al., 2011). It has been demonstrated that activation of both copies is concordant with changes to the chromatin structure and thus results in an increased amount of product from *Kncq1* gene (Korostowski et al., 2011). The *Kcnq1* gene is almost unique as an example of temporal change to biallelic expression patterns, although the majority of genes that maintain the pattern of switches from imprinted to bi-allelic state support the model in which imprinted expression manifests during embryogenesis and then lost during development (Babak et al., 2015).

As with the *Kcnq1* example, different transcripts and their specific expression states of *PLAGL1* might be associated with distinct functions. Interestingly, mouse *Plagl1*/*Zac1* is known to regulate function of the *Igf2-H19* locus (Varrault et al., 2006, 2017). As in humans, mouse *H19* and *Igf2* genes are neighbours and, in addition to the ICR, share two enhancers located ~8 kb downstream of *H19* and ~80 kb downstream of *Igf2*. Direct binding to these enhancers of *Plagl1*/*Zac1* result in transactivation of *Igf2* and *H19* promoters. Interestingly, *Plagl1/Zac1* was also shown to regulate a gene network that has at

least 246 genes (Varrault et al., 2017). Within this network *Plagl1* was also able to alter other imprinted genes that do not belong to the *Igf2-H19* cluster (Figure 3) (Varrault et al., 2006).

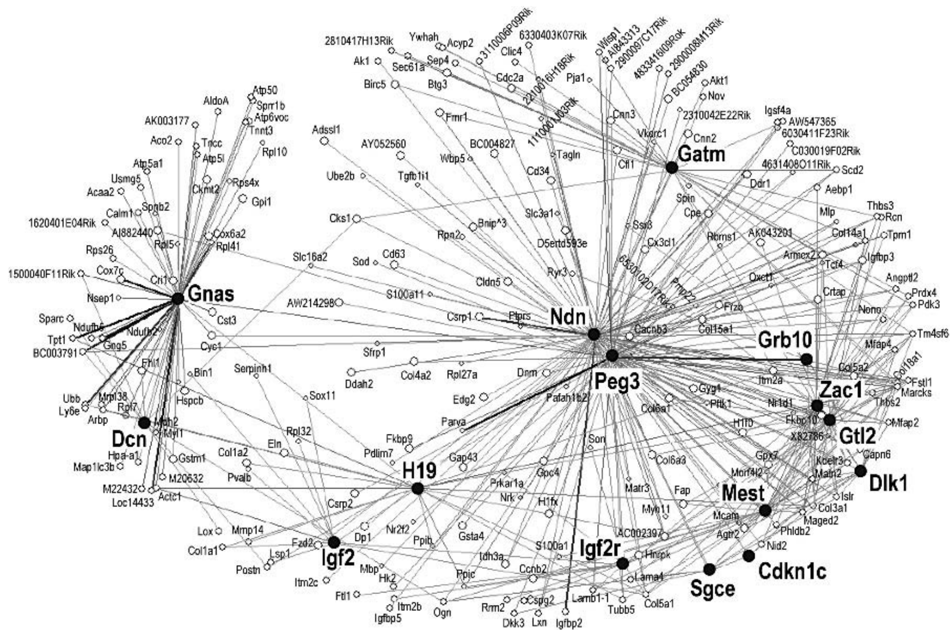


Figure 3. The imprinted gene network that controls mouse embryonic growth and differentiation. *Peg3* gene was identified as a gene altered by *Plagl1/Zac1* mutation. Other imprinted genes highlighted in black. Adapted from Varrault et al. (2006, 2017).

Among the imprinted genes represented in the network that undergoes regulation of the mouse *Plagl1/Zac1* gene, the *Peg3* gene is another example, which in turn impacts on three large gene families to bring about changes in brain and placenta tissue (Keverne, 2015). These three large families (prolactins, ceacams, and pregnancy-specific glycoproteins) are downstream of *Peg3* and together provide a robust framework for fetal development (Blois et al., 2014). Thus, identification of an exact mechanism of *Peg3* alteration by *Plagl1*/*Zac1* is very important in the context of embryo development and mother-fetus coadaptive processes. Additionally, in this paper, the authors concentrated their attention on the paternally expressed transcripts from the main promoter P1 only. However, as the P1 promoter seven-finger isoform product was reported to have functional differences from the P2 five-finger isoform product (Bilanges et al., 2001), it is highly likely that the splicing ratio of these two transcripts has an additional function. As the function of the *Plagl1*/*Zac1* gene may depend on the specific transcript, further investigation in this direction would be necessary to establish all interconnections between imprinted genes, which might give a better understanding of their evolutional mission.

1.5. Germline DMRs as an indicator of imprinted expression

The *Igf2* and *Plagl1*/*Zac1* examples have demonstrated that tissue-specific expression may depend on the promoters. Imprinting status of a gene can be also detected by studying the methylation status of its germline DMRs. The human *GNAS* cluster serves as an excellent example that illustrates how DMRs can provide such regulation. The *GNAS* cluster is a highly complex structure, as it gives rise to maternally, paternally and biallelically expressed transcripts determined by differentially imprinted promoters (Figure 4) (Peters et al., 1999). The upstream first exon of the *NESP* gene has a maternally expressed *NESP* transcript (Williamson et al., 2006). Exon XL is located downstream of the *NESP* gene and gives rise to a paternally expressed *GNASXL* transcript (Hayward et al., 1998; Peters and Williamson, 2008). In addition, there is a *GNAS* gene that contains 13 exons (Peters et al., 1999), although only the transcript produced from exon 1 is called *GNAS*. To avoid further confusion in the following text I will use combination of words such as *GNAS* locus and *GNAS* transcript to clarify the regulation mechanism. The *GNAS* transcript is biallelically expressed in many tissues (Mantovani et al., 2004), but was found to have maternal expression in anterior pituitary (Hayward et al., 2001), thyroid and ovary (Mantovani et al., 2002). Also, the 1A exon located in very close proximity to the *GNAS* exon 1 is a separate unit and produces the paternally expressed *EXON 1A* transcript (Kelsey, 2010). Lastly, the cluster contains the *NESPAS* exon 1 from which paternally expressed noncoding transcripts exist as spliced and unspliced forms (Williamson et al., 2004).

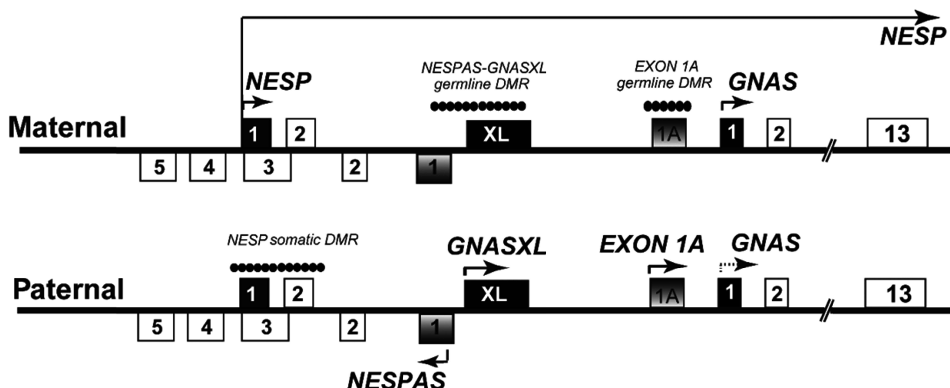


Figure 4. Human *GNAS* cluster. Solid black boxes denote first exons of the *NESP* (1), *GNASXL* (XL) and *GNAS* (1) genes. White boxes with numbers indicate other exons of a gene, for simplicity exons of *GNAS* transcript are omitted. Narrows denote direction of transcription. Black dots over the boxes represent three DMRs within the cluster. Adapted from Kelsey et al. (2010).

Despite the fact that mouse and human *GNAS* clusters have almost identical structures, except that the *GNAS* gene contains 12 exons (Peters et al., 1999), the mouse *Gnas* gene has demonstrated maternal expression among tissues that do not overlap with human tissues. For example, maternally expressed transcripts in mice were found in renal proximal tubules, and in brown and white adipose tissues (Kelsey, 2010; Yu et al., 1998). A different distribution of maternally expressed transcripts between mice and humans may point to different mechanisms of establishment, for which two models were proposed. The *GNAS/Gnas* cluster contains three DMRs, two germline and one somatic, which control allelic expression of the entire cluster (Kelsey, 2010; Peters and Williamson, 2008). The *GNAS* transcript that provides biallelic and monoallelic expression is solely regulated by *EXON 1A* gDMR (Kelsey, 2010; Peters and Williamson, 2008). The first model, also called the repressor model, suggests that a silencer protein binds to the unmethylated *Exon 1A* DMR on the paternal allele, thus causing prevention of *GNAS* transcription by establishing a repressive chromatin domain (Peters et al., 2006; Sakamoto et al., 2004; Weinstein et al., 2001). This model has a limitation, because in this case a silencer protein must be restricted to the tissues in which *GNAS* shows maternal expression. The second model is, in contrast, an insulator model, under which *EXON 1A* is able to bind CTCF protein on the paternal allele. The presence of CTCF protein leads us to expecting a block of the *GNAS* promoter access for enhancers and thus the paternal allele would be silenced. The same Exon 1A gDMR located on the maternal allele is methylated and therefore cannot act as a CTCF binding domain, resulting in expression of the maternal allele (Peters et al., 2006; Weinstein et al., 2001). The second model is more probable as the *GNAS* promoter itself is located within a constitutively unmethylated CpG island in very close proximity to *Exon 1A* DMR with methylation on the maternal allele (Kelsey, 2010).

The exact mechanism of tissue-specific DMR methylation is yet to be clarified, although there is more and more evidence that DMRs are a valuable source for detection of a precise methylation pattern over the regulated genes. Likewise, intermediate methylation status of a DMR, which is shown to be associated with imprinting, arises from the situation in which one parental allele is methylated whereas the other is unmethylated, and thus measuring the methylation status of both alleles gives an intermediate value (Bar et al., 2017; Court et al., 2014a). The main question regarding the change from imprinted status into the biallelic state at gDMRs is whether so called "loss of imprinting" itself correlates with changes in DNA methylation at control regions. Correlation analysis of methylation levels of gDMRs and biallelic scores calculated genome-wide in human pluripotent stem cells (hPSCs) has clearly demonstrated a strong correlation between low methylation of maternal gDMRs and biallelic expression of genes (Bar et al., 2017). Also, the clear concordance of intermediate DMR methylation was noted for monoallelic expression of a gene. For example, paternal expression of the *PSIMCT-1* gene was connected to the intermediate methylation status of the related gDMR (Bar et al., 2017; Court et

al., 2014a). In the case of clusters that undergo both paternally and maternally methylated gDMRs, such as the *GNAS* locus, opposing methylation changes were noted (Bar et al., 2017). This would mean that in the case of loss of imprinting by several genes within a cluster corresponding maternally imprinted DMR have low methylation status, whereas the paternally imprinted DMR, in contrast, would be highly methylated. For example, the low methylation level at the maternally imprinted promoter of *GNASXL* and *GNAS-AS1* are correlated with biallelic expression of the *GNAS-AS1* gene. Indeed, as predicted, the same sample has also demonstrated biallelic expression for the promoter of the *NESP55* gene with a high methylation level at paternally imprinted DMR (Bar et al., 2017). This mechanism of the distribution of methylation marks over maternal and paternal DMRs within a cluster is yet to be explained, although knowledge of correlation between gDMRs and imprinted expression within a cluster provides great potential for understanding tissue-specific imprinted expression.

1.5.1. Placenta-specific germline DMRs

The appearance of tissue-specific imprinting could be partially explained by functional classes of DMRs within the ICR regulating the organisation of a cluster. Despite the rapid development of large-scale, base-resolution methylation technologies that allow detection of methylation status of a given allele caused by appearance and preservation of imprinted status, the list of germline DMRs is, as yet, incomplete. To date, nearly 50 ubiquitous imprinted germline DMRs have been described in the human genome with an additional list of placenta-specific DMRs (Court et al., 2014a). One of the intriguing discoveries for many placenta-specific DMRs, with the exception of *ZFAT*, *GPR1-AS*, and *MIR512-1*, is that they do not inherit methylation from the gametes and are devoid of methylation in human embryonic stem cells (Court et al., 2014a). These data provided preliminary evidence that the differential methylation associated with genomic imprinting is a dynamic process that is regulated during fusion of gametes at the fertilization phase. Moreover, evidence of preservation of non-inherited methylation status for placenta-specific germline DMRs across several somatic tissues tested in this study raised the question of an independent and yet unknown mechanism of establishment for these imprints (Figure 5) (Court et al., 2014a).

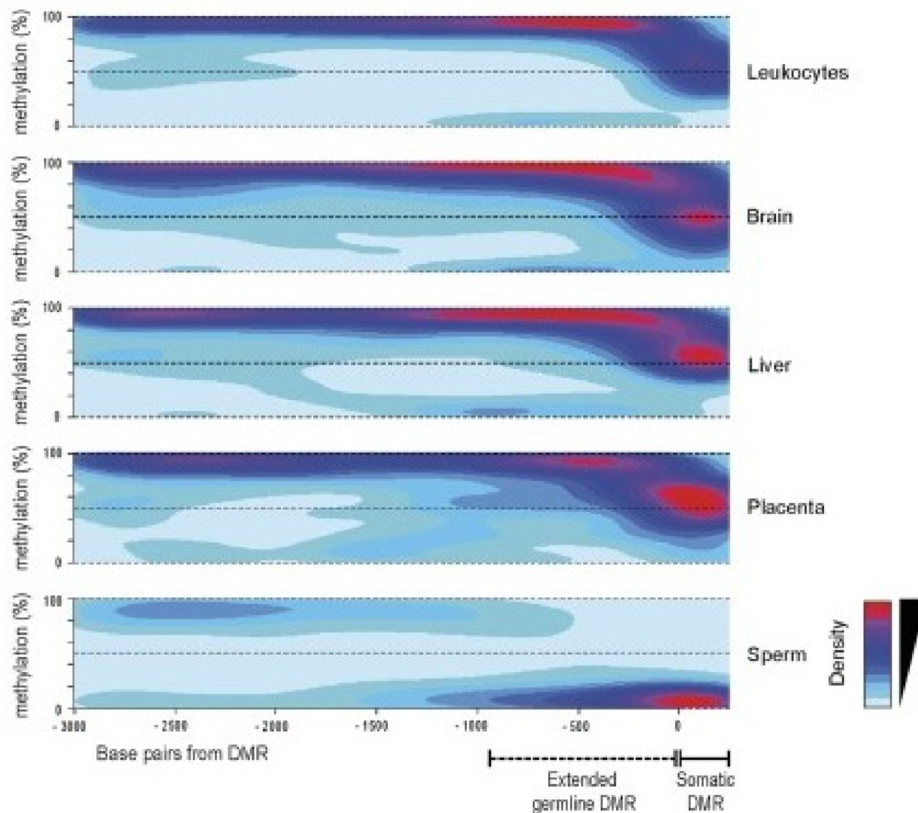


Figure 5. Methylation contour plots for human germline DMRs. The plot is originated from whole genome-wide bisulphite sequencing data sets for all maternally methylated imprinted DMRs and demonstrate preservation of an intermediate methylation status across somatic tissues, including placenta, but with significant extent in sperm. Light blue colour of the density represents an unmethylated state, which is specific for somatic tissues, except placenta. The red colour indicates the fully methylated DMRs. Adapted from Court et al. (2013).

A similar study of imprinted DMRs was conducted later in placenta samples, where the authors were able to identify 72 novel placenta-specific DMRs, and also observed partial loss of methylation for some DMRs (Hanna et al., 2016). Detected DMRs were detected as established in oocytes and these data are in line with the previous study of imprinted genes in mouse, where all placenta-specific genes were denoted as having maternal origin (Prickett and Oakey, 2012).

Although Court et al., had a sample with only a few somatic tissues that include liver-, brain-, and leukocytes samples, they noted a similar switch from intermediate methylation to unmethylated status for placenta-specific DMRs in the tested somatic tissues (Figure 5) (Court et al., 2014a). Later, Hanna et al. attempted to follow this switch in dynamics using isolated embryonic cell types

(Hanna et al., 2016). Extraembryonic tissues were chosen due to their close proximity to the embryo, as they originate later in development. An intermediate methylation status was detected for trophoblast and mesenchymal core of the placental villi, which supports preservation of imprinted status in the trophectoderm (TE) and inner cell mass (ICM) after the blastocyst stage. In contrast, a low methylation level, instead of average methylation, was registered in chorion and amnion. These data suggest that imprinting of placenta-specific DMRs in embryonic lineages starts much earlier than was previously predicted and should be lost by the time of primitive streak formation (Figure 6). As in the example of *Kcnq1* it may indicate a required increase of gene products at specific periods during development or there is another underlying mechanism in which imprinted expression occurs spontaneously due to the specific loss of methylation from the maternal allele. The authors concluded that placental-specific DMRs are either passively or actively losing methylation from the maternal allele, while the rest of the genome undergoes *de novo* methylation required for somatic differentiation (Smith et al., 2014).

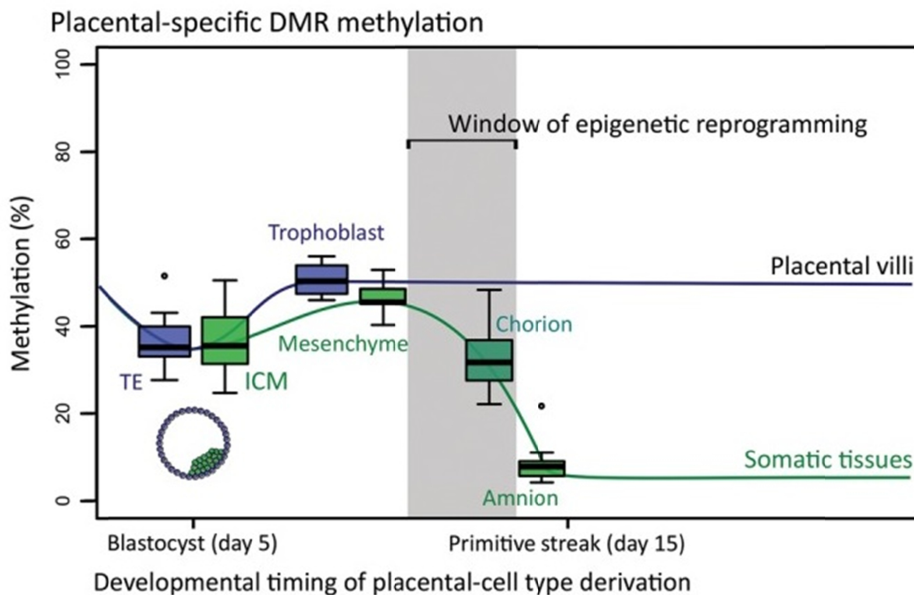


Figure 6. Loss of imprinting from placental-specific DMRs in human embryonic lineages. The distribution of methylation levels across several tissues. Maternal imprinting loss were registered from placental-specific DMRs. Trophectoderm (TE) and inner cell mass (ICM) derived cell lines are coloured in blue and green, respectively. Adapted from Hanna et al. (2016).

1.5.2. Transiently methylated germline DMRs

While some of the placenta-specific gDMRs and imprinted genes lose their methylation during ontogenesis, and thus later exhibit biallelic expression across adult tissues, other placenta-specific DMRs are able to keep and maintain imprinting after birth. At the blastocyst stage the organism survive the wave of methylation erasure process, while some specific sequences located within imprinted regions and certain repeat-type sequences are able to escape such modifications (Messerschmidt et al., 2014; Monk, 2015). Transiently methylated germline DMRs (tDMRs) identified from mouse genome were proposed as one of the possible mechanisms allowing maintenance of genomic imprints across somatic tissues (Proudhon et al., 2012). Transiently methylated gDMRs are themselves not distinguishable from germline DMRs in gametes and the preimplantation embryo, although the main difference is their ability to survive the post-fertilization demethylation process and later gain methylation on their paternal alleles at the time of implantation. The full mechanism by which gametes are able to provide a safe environment for tDMRs is not clear yet, although an elegant explanation of a mechanism associated with TET3-complex was suggested. It is important to note that this suggestion was based on methylation screening analysis in mice and there is no prediction of the number of tDMRs that are expected in the human genome and whether they are tDMRs or gDMRs with unknown mechanism of regulation.

Numerous placenta-specific gDMRs that could fit the criteria for tDMRs were identified in human cell lines (Sanchez-Delgado et al., 2016). Analysis of methyl-seq datasets in blastocysts, placenta and 14 different somatic tissues revealed different patterns of constant transfers.

Another type of pattern was detected for paternally methylated gDMR originated from sperm in humans, *IG*-DMR, which was partially methylated in blastocyst, indicating a partial loss of methylation after the germline stage, and resulting in imprinted expression for five somatic tissues (Sanchez-Delgado et al., 2016). Mouse *H*19 is also known to be protected by the *Dppa3* protein during active demethylation of the male pronucleus (Nakamura et al., 2007). The *Dppa3* protein, also known as Stella, protects the maternal genome from contaminant demethylation and protamine exchange by binding with the H3K9me2 histone (Mihaylova et al., 2012; Wossidlo et al., 2011). At the same time, the paternally methylated IG-DMR was not associated with *Dppa3* protection, presumably because it is not enriched for H3K9me2 nucleosomes (Delaval et al., 2007). Therefore, it is necessary to clarify whether observed patterns are indeed associated with potential tDMRs or became visible due to specific protection during demethylation. Nevertheless, taking into account the chance of the partially methylated state within the cell the authors observed 60 DMRs of oocyte origin with imprinted expression in more than 12 tissues (Sanchez-Delgado et al., 2016).

Analysis of imprinted gene expression in placenta in mice has demonstrated 100% maternal origin and thus DMRs regulating such expression are also expected to be paternally methylated or in other words repressed (Prickett and Oakey, 2012). In the case of assumed placenta-specific tDMRs established in this study, maternal methylation that dictates paternal expression was found for 9 genes in humans and non-human primates that were not previously known to be imprinted, including *AGO1*, *USP4*, *SH3BP2*, *FAM149A*, *MOCS1*, *R3HCC1*, *JMJD1C*, *PAK1* and *PAPLN-AS*. Moreover, attempts to show that such regions and their orthologues correspond to the mouse genome did not provide any successful results (Sanchez-Delgado et al., 2016). Such evidence supports the idea that similar mechanisms exist in human and mouse genomes, while particular sequences that undergo epigenetic modification with following denoting of paternal origin can be species-specific. Discovery of an additional set of placenta-specific germline DMRs that transfer their imprinted expression into adulthood would clearly demonstrate that this process is more widespread and abundant in human genome than expected.

1.6. Post-natal loss of imprinted expression

The switch from imprinted to biallelic expression can indeed be associated with the origin of germline DMRs and the time of their establishment. This observation can be supported by another study, where the loss of imprinted expression was observed in a series of mouse experiments. The authors aimed to detect the presence of autosomal epigenetic allele-specific expression (ASE) effects (Huang et al., 2017). It is important to note that in this case genomic imprinting, cell-specific genomic imprinting and cell-specific random monoallelic effects were combined into one common category called DAEEs – differential allele expression effects. This was preceded by an observation that parental alleles are only weakly correlated or not correlated at all (Figure 7).

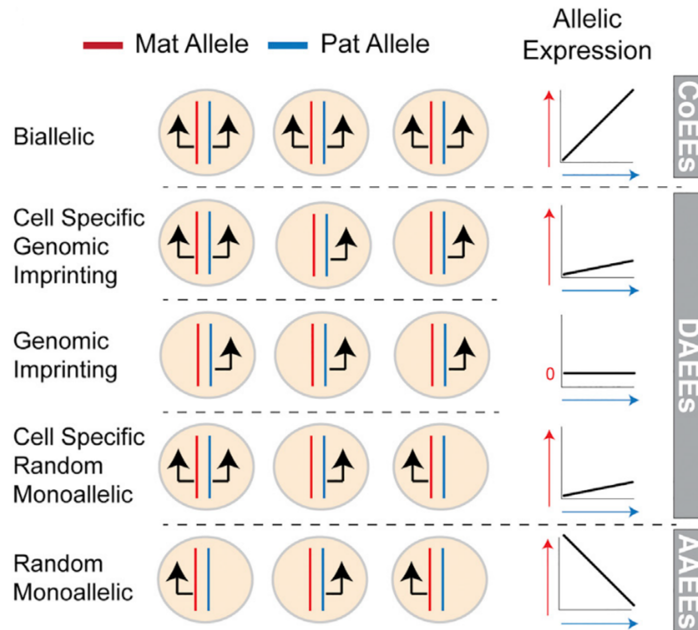
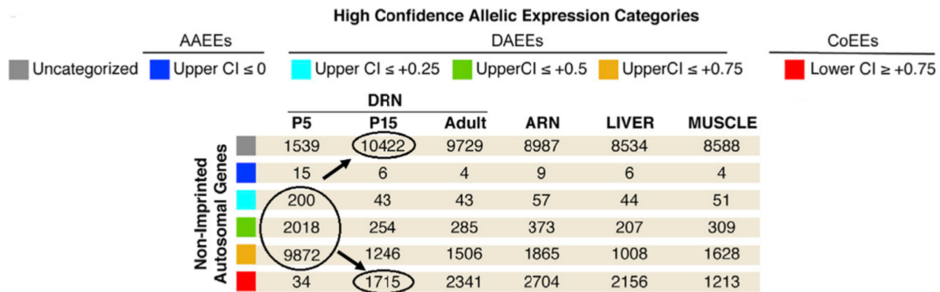


Figure 7. Schematic representation of different allelic effects at the cellular level. Allele-specific effects are divided into three categories used in this study – i) allele co-expression effects (CoEEs) that are expected in case of classical biallelic expression, ii) differential allele expression effects (DAEEs) that are expected in case of genomic imprinting, and iii) antagonistic allele expression effects (AAEEs) that are expected to involve random monoallelic expression. Black arrows correspond to the expression of a parental allele. Linear regressions on the right side represent allelic correlation, where red and blue arrows denote increasing maternal and paternal expression, respectively. Adapted from Huang et al. (2017).

As genomic imprinting is expected to be more common in brain when compared to other somatic tissues (Babak et al., 2015; Perez et al., 2015), analyses that aimed to understand the dynamics of ASE effects were performed on the dorsal raphe nucleus (DRN), which is the largest serotonergic region of forebrain associated with a variety of brain functions and also implicated in several mental illnesses. At the stage of post-natal P5 and P15 juveniles of DRN drastic loss of DAEEs was observed from 88% to 11%, while the number of genes of P15 stage remain the same as in observations from other somatic tissues (Figure 8).



Category	P5 (DRN)	P15 (DRN)	Adult (DRN)	ARN	LIVER	MUSCLE
Uncategorized	1539	10422	9729	8987	8534	8588
Upper CI ≤ 0	15	6	4	9	6	4
Upper CI ≤ +0.25	200	43	43	57	44	51
UpperCI ≤ +0.5	2018	254	285	373	207	309
UpperCI ≤ +0.75	9872	1246	1506	1865	1008	1628
Lower CI ≥ +0.75	34	1715	2341	2704	2156	1213

Figure 8. ASE effects at different developmental stages and somatic tissues of an adult organism. The scale that represent the percentage of all measured expressed genes in a panel of tissues. DAEEs are prevalent in the P5 DRN of neonates, while CoEEs are more prevalent starting from P15 stage. Adapted from Huang et al. (2017).

The number of genes with CoEEs increased ~50 fold between P5 and P15 of DRN (Huang et al., 2017). Situation with the opposite exchange of numbers between P5 and P15 of CoEEs and DAEEs could be viewed from different perspectives. On one hand, the authors have suggested the idea of a profound increase in DAEEs, while on the other hand, the numbers represented in Figure 8 could denote the transition from DAEEs to CoEEs within 10 days between P5 neonates and P15 juveniles. Although the authors report 77% loss of DAEEs within the 10 day period of neonates, there is a 7-fold increase seen for genes with uncategorized ASE effects, suggesting the idea that part of the genes might truly change their type of expression from imprinted to biallelic, while the number of such genes could be much lower due to the increased amount of uncategorized genes.

2. AIMS OF THE STUDY

The aim of the thesis is to investigate the methylation features of imprinted genes, identify novel imprinted genes and study their effects on human complex traits.

The specific objectives of the thesis are as follows:

1. To investigate imprinting-specific methylation status of imprinted genes and germline differentially methylated regions (gDMRs) in a wide panel of somatic tissues of adult humans.
2. To identify parent-of-origin effects for complex phenotypes such as NMR metabolites, body mass index (BMI) and age at menarche.

3. RESULTS AND DISCUSSION

3.1. Imprinted genes and imprinting control regions are stable and show predominant intermediate methylation (Ref. I)

Genomic imprinting is an epigenetic process that starts with establishment in the germline and results in monoallelic expression for specific groups of genes. Human imprinted genes are mostly clustered (Babak et al., 2015), being under *cis*-regulation of an imprinting control region (ICR) (Barlow, 2011), which has at least one germline differentially methylated region (gDMR) (Court et al., 2014a; Prickett and Oakey, 2012). Conserved imprinted patterns between mice and humans do not necessarily reflect the identity of ICRs, therefore validation of regions that were established in human germline and maintain imprinted expression into adulthood is important in the context of understanding their effect on complex traits.

Imprinted genes are predominantly found in placenta and other extra-embryonic tissues, although some genes are also found to be imprinted in adult somatic tissues. The full catalogue of imprinted genes in humans is not yet complete, therefore characterisation of specific methylation patterns in a panel of somatic tissues of an adult human organism would help to identify novel imprinted genes in a specific tissue sample. A recent paper by Baran et al. suggests that the pattern of tissue-specific imprinting in adult somatic tissues is expected to be relatively stable, while major differences could still be found between embryonic and adult tissues (Baran et al., 2015). Hence, it is advantageous to investigate the common methylation pattern in somatic tissues developed from different embryonic layers, including nerve and brain tissue samples, to improve our understanding of the nature of imprinted genes.

3.1.1. Description of cohort and materials

The study was performed using two datasets: 17 post-mortem tissues obtained from four autopsy patients and whole blood samples from 97 healthy individuals (N_{male} = 48, N_{female} = 49) from the Estonian Genome Center, University of Tartu (EGCUT) biobank.

The first dataset consists of four individuals – one female and three males. All tissue samples were collected between four and eight hours after death. The samples were selected to complement the panel that would cover tissues formed from all three germ layers during embryogenesis: mesodermal (e.g., adipose tissue, bone and lymph nodes), endodermal (e.g., gastric mucosa and tonsils), ectodermal (e.g., coronary artery) and even neuroectodermal (e.g. ischiatic nerve and medulla oblongata).

After DNA purification and bisulphite-conversion, all samples were hybridised on the Infinium Human Methylation 450 BeadChip array according to the Infinium HD methylation protocol. As one of the datasets represents post-mortem tissues, we selected negative and positive controls that would allow us to assess the quality of the samples. The Pearson correlation coefficients, which were >0.99, were calculated for both technical and biological replicates, confirming a good level of reproducibility for the array analysis. Methylation normalisation and further quality control was performed using the SWAN method (Maksimovic et al., 2012). To estimate the quality of genotyped CpG sites, detection p-values were calculated for each probe per sample. Probes with detectable methylation levels in <5% of samples with detection p-value <0.01, were cross-reactive, and with SNPs inside the gene body were excluded, which left 353,094 probes for further analysis.

3.1.2. Detection of methylation patterns for imprinted and non-imprinted genes

Imprinted genes and imprinting control regions are expected to have intermediate methylation status, which is explained simply by the distribution of methylated marks over two alleles, where one allele remains unmethylated to perform expression, while the other is methylated to deactivate transcription. Intermediate methylation status as a quantitative outcome is expected to be equal to 0.5. With the rapid development of array and sequence-based methylation profiling methods, the exact definition of an intermediate status became important in developing a standard that could be used uniformly across multiple studies. We therefore aimed to define the exact frame for β-values that would represent intermediate methylation status. To do this, we created the list of well-defined imprinted genes and gDMRs. We used two publicly available catalogues of imprinted genes to create an initial list. A literature search from 'PubMed' was applied to the list of genes to verify that imprinted genes were indeed shown to be imprinted in humans, with results replicated by independent studies. In total, the resulting list contained 76 known human imprinted genes (Suppl. Table 1 in Ref. I). The list of human-specific gDMRs was obtained from the study by Court *et al* (Court et al., 2014a).

To define a better frame for the intermediate methylation level, we calculated the proportions of different CpG methylation ranges and performed a one-tailed chi-squared test with one degree of freedom. By shifting frames in a range of 0.3–0.8, we found that the biggest difference in methylation for imprinted and non-imprinted genes was for β-values corresponding to 0.5–0.7 frame ($p=9.33\times10^{-187}$) (Suppl. Table 2 in Ref. I). Genome-wide visualisation of imprinted (N=76) and non-imprinted genes (N=20,515) revealed that the number of methylated probes with β-value 0.5–0.7 for imprinted genes was higher (25.5%) compared to non-imprinted genes (10%) (Figure 9). Surprisingly, the number of probes with smaller β-value (<0.5), indicating the

unmethylated state, was lower for imprinted genes (28.2%) than for non-imprinted genes (48.1%). Also, significant differences were found for both groups of genes in respect of higher methylation levels (>0.7), showing 46.2% for imprinted and 41.9% for non-imprinted genes (Figure 9).

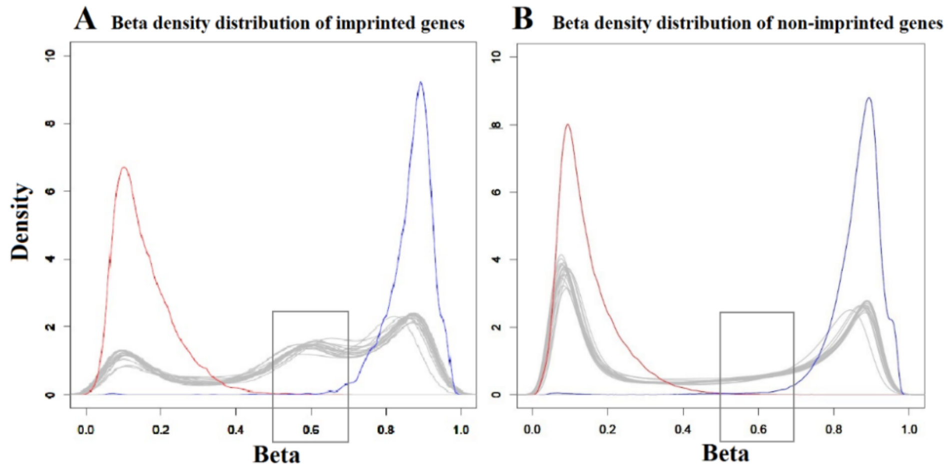


Figure 9. Visualisation of beta-density and methylation level (beta) for imprinted (A) and non-imprinted (B) genes, captured by the methylation array. Positive and negative controls are represented by blue and red lines, respectively. Somatic tissues tested are shown by grey lines. The grey box area depicts the proportion of defined intermediately methylated range which corresponds to β-value of 0.5-0.7.

To take account of the fact that tissue samples were collected after death, we compared the genome-wide methylation pattern of autopsy samples to whole-blood methylation levels measured in healthy individuals. We observed a similar distribution of the proportion of probes in the three categories of methylation status for both groups of genes. Intermediately methylated probes also demonstrated the highest interval between imprinted and non-imprinted genes (21.2% and 12.0%, respectively, $p=1.76\times10^{-121}$). It is well-known that expression patterns are comparable for post-mortem and before death samples, while methylation was expected to change significantly. The global overlap observed in this study clearly indicates that blood is an excellent candidate for both studying methylation patterns and further discovery of imprinted expression.

3.1.3. Equality of variances

Given the fact that imprinted genes indeed contain a higher proportion of intermediately methylated probes, we next hypothesised that the variance of β-values is expected to be smaller for imprinted genes than for non-imprinted

genes. The motivation for this estimation comes from the knowledge that imprinted genes, being subordinated to germline DMRs would keep their imprinted pattern presumably across a variety of somatic tissues in adulthood. Thus, this specific feature should be contrasted by non-imprinted genes, for which expression is highly tissue-dependent, resulting in a more stochastic methylation pattern and therefore in higher variance for β-values that correspond to methylation probes. Levene's test (levenes.test in R, version 3.2.1) was selected as an instrument to compare the variance of methylation levels for imprinted and non-imprinted genes. We compared 76 previously selected imprinted genes with 76 randomly selected genes for each of the 17 post-mortem tissues, iterating the procedure 100 times. An identical procedure was performed for blood samples from healthy individuals. As hypothesized, imprinted genes demonstrated smaller variance of β-values within the 18 tissues calculated out of all replications (e.g., $p=2.31\times10^{-14}$ for brain tissue and $p=4.22\times10^{-12}$ for nerve tissue). Iteration procedure of this test also demonstrated that the number of genes with methylation patterns similar to imprinted genes is expected to be relatively low, as selection of random non-overlapping non-imprinted genes showed comparable variance for β-values.

3.1.4. The location of intermediately methylated probes

The Illumina Infinium 450K methylation array provides access to six regions of a gene: promoter area, 5'UTR, first exon, gene body and 3'UTR, while the promoter area is divided into two regions spanning up to 200 bp and 200-1500 bp from the transcription start site (TSS200 and TSS1500, respectively). We therefore examined whether intermediately methylated probes tend to be grouped in specific gene regions or are equally distributed across the length of a gene. Due to the evidence that germline DMRs are often located within the promoter area, we excluded 49 gDMRs prior to the analysis. We discovered that TSS1500 is the only region that maintains the significant difference after multiple testing correction with an increased proportion of intermediately methylated probes for 22.8% ($p=6.57\times10^{-5}$) (Figure 2 in Ref. I). This finding is clearly in line with the fact that epigenetic gene regulation is mediated through promoter regulation. However, to our knowledge, this is the first evidence that, in the case of imprinted genes, intermediately methylated probes tend to localise in the extended region of the promoter spanning from 200 to 1500 bp from the transcriptional start site.

3.1.5 Germline DMRs are not ubiquitously methylated across somatic tissues

Levene's test provided us with evidence that genes imprinted in humans selected for this study are expected to maintain their imprinted expression

across different somatic tissues. We also tested whether regulatory germline DMRs also preserve their intermediate methylation status. We tested 59 gDMRs including 45 ubiquitous gDMRs and 14 placenta-specific gDMRs (Suppl. Table 4 in Ref. I). As was predicted by other studies, the 45 ubiquitously imprinted gDMRs demonstrated remarkable stability of methylation patterns, maintaining β-values in the range from 49% to 77% across all 18 somatic tissues tested. In contrast, the 12 placenta-specific gDMRs had methylation patterns with β-values ranging from 10% to 20%, except for two regions: *GPR1-AS* and *MIR512*. Both of these regions have a pattern characterized by intermediate methylation status, quantified as 55% and 67%, respectively, while only *GPR1-AS* methylation pattern corresponds to ubiquitously imprinted germline DMRs (Figure 10).

A study of polymorphic methylation in placenta links the unmethylated status of gDMRs with biallelic expression of imprinted genes (Sanchez-Delgado et al., 2016), thus demonstrating a possibility for loss of an imprinted pattern for genes that are regulated by such DMRs. The same set of placenta-specific gDMRs that remained unmethylated in our study was demonstrated to have intermediate methylation status in embryonic and extraembryonic tissues and, being a stochastic trait in placenta, supports the idea that a significant proportion of transcripts are monoallelically expressed in cleavage embryo, escaping further imprinted expression (Court et al., 2014a; Sanchez-Delgado et al., 2016). The other 45 tested gDMRs demonstrated a clear preservation of intermediate methylation levels pointing to the widespread and probable constant genomic imprinting expression across somatic tissues of an adult organism. Intriguingly, all gDMRs selected for this study were identified as established in the germline, while the part of gDMRs did not inherit the methylation from gametes. Changes in their methylation pattern cannot be associated with placental expression, as the identified GPR1-AS exception maintains an intermediate methylation status. Although our study has extended the knowledge of the tissue-specific nature of imprinted genes, further studies are required to provide a deeper understanding of regulatory factors that allow maintenance or loss of imprinted expression throughout the development and aging of an organism.

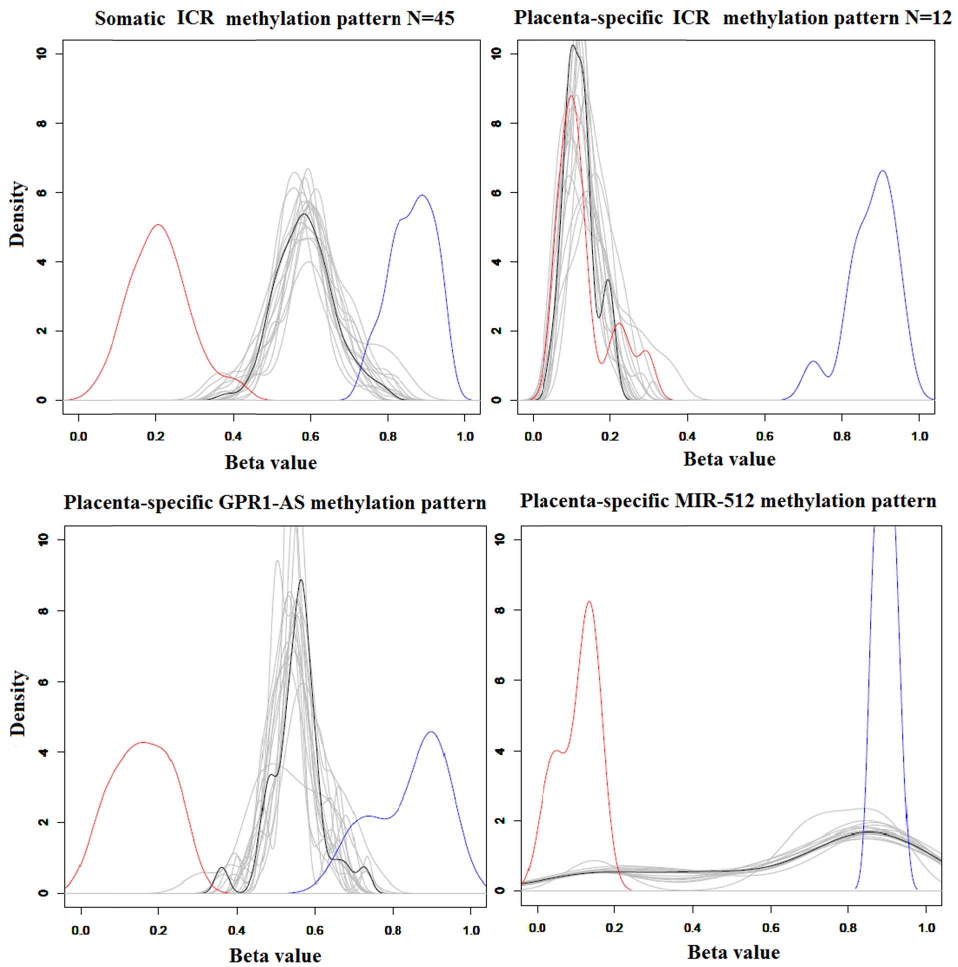


Figure 10. Visualized methylation pattern across 17 human somatic tissues for known germline DMRs. Each tissue is shown by a grey line with the black line representing the mean of all tissues. Controls that correspond to unmethylated and fully methylated status are shown by the red and blue lines. Visualized DMRs can be divided into two categories: ubiquitously imprinted DMRs and placenta-specific DMRs with two exceptions *GPR1-AS* and *MIR-512*.

3.2. Parent-of-origin effects in *PTPRD* gene affect the level of triglycerides in medium very low density lipoprotein NMR metabolite (Ref. II)

Circulating metabolites serve as a bridge connecting genetic variants and associations with chronic diseases, and for this reason they are widely used as an intermediate trait in genome-wide association studies (GWAS). While associations with human metabolites have already been established for hundreds of genetic variants, in only a few cases have the levels of metabolites been found to be impacted by parent-of-origin effects (POEs). Genomic imprinting, as a key driver of POEs, is also known to be strongly associated with a range of metabolic processes in mammals, including glucose metabolism and maintenance of the body temperature of new-borns (Peters, 2014; Yan et al., 2016), although the key observations are focused on the specific group of genes that are already known to be imprinted. In contrast, human circulating metabolites are a large collection of molecules that are often highly correlated, and it is advantageous to view them as a complex to find more gene-phenotype connections. The human nuclear magnetic resonance (^{1}H-NMR) technique, therefore, serves as a powerful tool for measuring metabolites that can be used for a broad range of investigations, including systematic discovery of possible parent-of-origin effects, genome-wide.

In the case of parent-of-origin effects, both the presence of the mutant allele and the allelic origin can affect a phenotype. We were therefore interested in discovering whether the level of human circulating metabolites is modulated by imbalanced expression of alleles. To investigate this hypothesis, we performed a genome-wide scan for 14,815 individuals of European ancestry to identify POEs of common (MAF > 1%) variants in 82 ^{1}H-NMR metabolites.

3.2.1. Description of cohorts and methods

The study was performed on three European cohorts – one Estonian (EGCUT, N=5,861) and two Finnish (FINRISK1997, N=8,438 and DILGOM, N=516). Each cohort was genotyped with a genome-wide commercial array according to standard protocols and further imputed using the 1000 Genomes Project multiethnic reference panel (Suppl. Table 3. In Ref. II).

The quantification of NMR biomarkers (N=135) was performed using proton ^{1}H-NMR spectroscopy of native blood plasma in the Estonian cohort and serum in the Finnish cohorts. All measurements were quantified using the same high-throughput NMR platform in the same analysis laboratory as described by Soininen *et al* (Soininen et al., 2009). As part of the quality control procedures, metabolites with missing data and outliers were discarded (Kettunen et al., 2016), leaving 82 metabolites for the further analysis. Imputation of missing ^{1}H-NMR values was performed using the MICE R package (Azur et al., 2011). In

the final step, residuals for metabolites were calculated adjusting for sex, age, age^2, BMI and the first ten genetic principal components.

To search for possible parent-of-origin effects, we applied a method based on a modified Brown-Forsythe test, comparing the phenotypic variance of the heterozygous group to the variance of both homozygous groups, which allows detection of potential POE effects from population-based data (Hoggart et al., 2014). To measure allelic imbalance, RNA-sequencing libraries were prepared using 482 EGCUT RNA blood samples according to the Illumina manufacturer's instructions. Allele-specific expression (ASE) was further determined using maximum likelihood estimation in combination with a likelihood ratio test (LRT). The statistical power of the ASE was increased by combining multiple SNPs per region and comparing them to the SNP of interest.

3.2.2. Meta-analysis revealed an effect of *PTPRD* gene to VLDL lipoprotein

The approach used in this investigation was chosen because it is sufficiently simple and is effective in scanning for potential POE signals using data from unrelated individuals, which are widely available in many cohorts. This method has been used successfully to identify two genes – *KCNK9* and *SLC2A10* – having POE in the analysis of BMI, and these results were replicated using family trio data.

Genome-wide meta-analysis of 1000G reference panel imputed SNP panel using 14,815 individuals revealed a variant, rs1412727 (MAF=0.17, $p=8.29\times10^{-11}$), located in the intron of the *PTPRD* (protein-tyrosine phosphatase, receptor-type, D) gene (Figure 11) that is associated with the metabolite called 'triglycerides in medium size VLDL'. The designation of the identified molecule as 'triglycerides in medium size VLDL' reflects the number of endogenously secreted triglycerides attached to the core of very low-density lipoprotein molecules. Although triglycerides associated with VLDL molecules are the source of energy for cells of the human body, they have been shown to be causally related to coronary heart disease (Chasman et al., 2009; Sacks, 2015). An additional eight independent SNPs that reached the genome-wide significance threshold, but failed to reach p-value after multiple testing correction, were also selected for further validation (Table 1 in Ref. II).

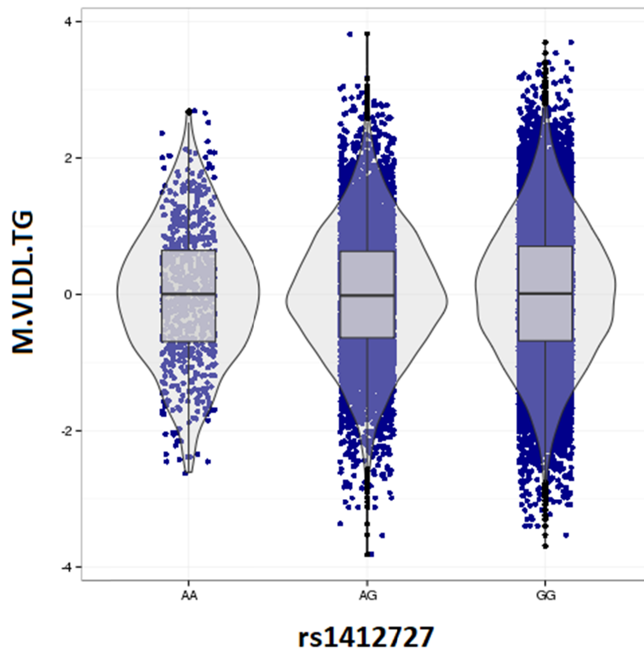


Figure 11. The variability of triglycerides molecules associated with normalized values of medium sized VLDL (M.VLDL.TG) by genotype of rs1412727. The violin plot describes the density of data and the box plot shows median with first and third quartiles, with the whiskers extending from the borders to ±1.5xIQR, where IQR stands for inter-quantile range, the distance between the first and third quantiles.

In addition to genome-wide POE analysis, we were interested whether parent-of-origin effects can also be found for previously established GWAS signals of NMR metabolites (Kettunen et al., 2016). In total, we found five signals that reached a nominal significance threshold ($p<0.05$) (Suppl. Table 1 in Ref. II), but none of them reached this threshold after adjusting for multiple testing.

3.2.3. Allele-specific expression analysis

The POE method enables identification of plausible parent-of-origin effects that may also result in the state of allelic imbalance. We therefore aimed to perform ASE analysis of nine POE signals that were identified in the meta-analysis described above, and five additional nominally significant signals identified from the GWAS study of Kettunen *et al* (Kettunen et al., 2016). As RNA-seq measurements were made using blood samples, it is important to first consider expression levels for identified signals. The main signal that was associated with the *PTPRD* gene was not detected to be expressed in EGCUT blood samples. Indeed, experiments in mice and humans detected *PTPRD* expression in the specific regions of the brain, including the hippocampal CA2 and CA3

regions, B lymphocytes, and in the thymic medulla (Lash et al., 2000; Uetani et al., 2000). Investigation of EGCUT blood samples demonstrated that only two genes, *EEPD1* and *LGI2*, can be considered for further ASE analysis, while others were discarded as they were not expressed in blood samples or due to the lack of informative SNPs to distinguish parental alleles. No SNPs in these two genes were detected to have allelic imbalance.

The *PTPRD* gene is reported by many studies to be associated with several distinct phenotypes, such as age at menarche (the onset of first menstruation in girls that indicates the start of reproductive capacity) (Perry et al., 2014), restless legs syndrome (Winkelmann et al., 2011), type 2 diabetes (Below et al., 2011; Tsai et al., 2010), epilepsy (Speed et al., 2014) and many others (Deming et al., 2017; Johnson et al., 2010; Kim et al., 2013; Mattheisen et al., 2015; Schosser et al., 2013; Traylor et al., 2016), although little is known about this gene and its functional background. *PTPRD* has been reported to belong to the protein tyrosine phosphatase (PTP) family, and knock-out mice exhibited learning impairment associated with increased magnitudes of long-term potentiation in the hippocampal region of the brain. Although we were not able to demonstrate allelic imbalance for *PTPRD*, this gene remains a highly probable candidate for having POE. Through analysis of independent maternal and fetal genetic effects, the *PTPRD* gene was found to affect the level of maternal circulating organohalogens associated with active control of toxicant disposition between the mother and fetus (Traglia et al., 2017). This study gives rise to the idea that POE identified in our study could be driven by maternal of fetal genetic effects, rather than by genomic imprinting. The association signal for age at menarche, rs7865468, that is linked to the *PTPRD* gene, was tested for imprinted expression in the deCODE study, but did not demonstrate either paternal- or maternal-specific association (Perry et al., 2014). Because rs7865468 is effectively independent of the signal identified in our study (r^2=0.0001, D'= 0.0179), it would also be worth testing this signal for the presence of exclusively maternal or paternal expression. Thus, we believe that our results provide great potential both for researchers who are interested in uncovering the nature of genomic imprinting and those who are looking for clarification of the underlying pathophysiology of many diseases.

3.3. Parent-of-origin allelic association affecting BMI and Age at Menarche complex traits (Ref. III and Ref. IV)

Imprinted genes have key regulatory effects on complex traits in both the embryonic state and adulthood (Peters, 2014). Given the fact that imprinted expression is a dynamic process, and assuming the idea of regulation of gene dosage (Korostowski et al., 2011), uncovering imprinted genes associated with BMI and age at menarche traits become important because these phenotypes are associated with the risk of type 2 diabetes (Abbasi et al., 2017), cardiovascular diseases (Twig et al., 2017) and breast cancer (Perry et al., 2014). The causal

relationship between BMI and pubertal timing is yet to be clarified. The vast majority of studies worldwide report that age at menarche highly depends on nutrition and has an inverse relationship with pubertal timing and BMI (Bratke et al., 2017; Talma et al., 2013; Wang et al., 2016). On the other hand, the physiological features of an organism, reflected in increased weight and height at menarche, cannot be ignored, as they indeed have higher BMI than girls of the same age who have not started menstruating (Stark et al., 1989). Phenotypic correlation of these two traits can also occur as a result of a pleiotropic effect from shared genetic background. For example, the *LIN28B* gene, for which imprinted expression is known to be restricted to the placenta, was among the first genes associated with pubertal timing (Ong et al., 2011). Later, it was demonstrated to modulate BMI and body shape of adult organisms (Leinonen et al., 2012). We therefore aimed to identify more imprinted genes that may also effect both BMI and age at menarche phenotypes. Discovery of imprinted genes was one of the projects launched by GIANT and ReproGen consortia using, among others, data from EGCUT, which aimed to find novel loci associated with these two traits.

3.3.1. Description of cohort and materials

For both studies, we used EGCUT cohort samples. For BMI and age at menarche studies, 9,274 and 4,747 individuals were available, respectively. DNA from the samples was genotyped with both the Illumina HumanOmniExpress and the Human370CNV Beadchip. Genotyping data was imputed up to the HapMap2 reference panel (Suppl. Table 3 in Ref. III and Suppl. Table 1 in Ref. IV). Association analysis for BMI was performed using the POE method implemented in the QUICKTEST software, while GWAS analysis for age at menarche was performed using the SNPTESTv2 software.

3.3.2. Association analysis revealed enrichment in imprinted genes

In total, data from 56,092 individuals from 15 studies and 132,989 women from 57 studies were available for BMI and age at menarche analyses, respectively, where EGCUT data was part of these analyses. After imputation, genome-wide data were available in up to 2.6M SNPs for BMI analysis and 2.4M SNPs for age at menarche analysis. BMI analysis was performed separately in men and women, although no sex-specific differences in effect were observed, and therefore summary statistics from sex-stratified analysis were then meta-analysed together. Age at menarche analysis was performed in women only. Parent-of-origin effects affecting BMI were tested as described in reference II, and then validated in family based-studies, whilst for age at menarche, genetic variants

were first tested with linear regression and then replicated in the deCODE Study in 35,377 women with available data regarding parental origins.

As a result, two loci (mapping near *SLC2A10* and *KCNK9* genes) demonstrating parent-of-origin effects were found for the BMI phenotype and three loci (*DLK1*/*WDR25*, *MKRN3*/*MAGEL2* and *KCNK9*) for age at menarche. Association analyses for BMI and age at menarche revealed not only that both phenotypes are enriched in imprinted regions, but also that there are variants in the same gene, *KCNK9,* that are associated with both phenotypes. Although the *KCNK9* variant reported for BMI, rs2471083, study is independent of that reported for age of menarche, rs1469039 (r^2=0.0054, D'=0.2649), the possibility of a shared genetic background deserves further investigation. The *KCNK9* gene exhibits tissue-specific imprinted expression (Luedi et al., 2007; Morcos et al., 2011). It is expressed predominantly in the brain (Kim et al., 2000; Rajan et al., 2000), and has also been found to exclusively have maternal expression in the fetal brain (Luedi et al., 2007), but biallelic expression in lymphoblasts and fibroblast cell lines (Morcos et al., 2011). The *KCNK9* gene encodes a protein that functions as a pH-dependent potassium channel (Kim et al., 2000), whereas the mutation of the maternal allele at nucleotide 770 in exon 2 is causal for Birk-Barel syndrome that is characterized by intellectual disability, hypotonia, hyperactivity, and changes in face structure (Barel et al., 2008). The operation of human *KCNK9* imprinted expression is under regulation of the *PEG13* maternally methylated germline DMR (Court et al., 2014b). *PEG13* binds CTCF-cohesin complex that regulates the separation of sister chromatids during cell division and facilitates downstream DNA repair by recombination (Hanssen et al., 2017). Thus *KCNK9* is a valuable candidate gene for BMI and age at menarche phenotypes, as earlier onset of menstruation was also previously associated with higher risk of cancer (Dall and Britt, 2017). I believe that the effect of gene dosage and expression patterns in tissues related to the examined phenotypes should be viewed more closely to uncover the etiology and association with a pathophysiological pathway.

CONCLUSIONS

Studying of the tissue-specific imprinted expression patterns would bring a better understanding of how these genes can modulate diseases or specific traits in humans and animals. Comprehensive imprinting panels would shed light to the specific tissues and developmental stage of an organism, which would lead to the exploring the allelic effect of a gene in pathophysiological pathways of a disease.

The main conclusions drawn from this thesis are as follows:

- The first aim of the research presented in this thesis was to define the methylation patterns for imprinted genes and germline DMRs in somatic tissues of adult organisms. Using a panel of 18 somatic tissues, we identified an imprinting-specific methylation pattern that is characterized by an increased number of intermediately methylated probes within a specific area of the promoter. Although we were not yet able to show how this distribution of methylation marks affects the expression of imprinted genes, we believe that our results will contribute to further characterization of tissue-specific imprinting patterns.
- The panel of post-mortem tissues used in the first study was complemented with blood samples from living individuals, which confirmed earlier suggestions that blood tissue is a valuable resource for studying imprinted genes. On the other hand, it is important to note that the parent-of-origin effect is less pronounced, and could therefore be overlooked because of the high complexity of this particular tissue.
- We were also able to show that germline DMRs, established in the pre-fertilization period, lose their intermediate methylation status and become unmethylated. Unmethylated status was later linked with the biallelic expression of genes that are regulated by germline DMRs within a cluster. The loss of methylation was detected mainly for placenta-specific germline DMRs. Thus, it is highly probable that genes that were detected as imprinted in placenta have biallelic expression in tissues presented in this study. To date, it is not known when the switch from the monoallelic to the biallelic state occurs, although it is likely that this happens shortly after birth. The mechanism for the loss of imprinting suggested in this study is human-specific, and it is possible that the same conserved imprinted genes in other species have a different expression pattern, as this depends on germline DMRs, which are highly species-specific.
- The concept of dynamic and highly tissue-specific imprinting is important in the light of how imprinted genes modulate diseases and traits. In our study of age at menarche, we showed that there is enrichment for associated genes in imprinted regions of the genome. Further research would be necessary to measure the expression of imprinted genes in relevant reproductive tissues to understand the effect of gene dosage on this particular phenotype.

- We also found that parent-of-origin effects impact BMI and may affect the level of a specific metabolite through the *PTPRD* gene, which was also found in our study to be associated with the age at menarche phenotype. The association of BMI and age at menarche phenotypes is not novel, although our study shows, for the first time, the importance of taking metabolite measurements into account to clarify the causal relationship between these traits.

SUMMARY IN ESTONIAN

Geneetilise vermimise mõju komplekstunnustele

Juba 1980. aastate alguses läbiviidud katsed loomade rakutuumade siirdamisega paljastasid epigeneetilise nähtuse nimega geneetiline vermimine (*genomic imprinting*), mille tõttu tekivad vanemaspetsiifilised erinevused geenide ekspressioonis. Diploidses organismis ekspresseeruvad tavaliselt mõlemalt vanemalt päritud geenikoopiad. Vermitud geenide puhul ekspresseerub ainult üks geenikoopia ja teine on deaktiveeritud.

Nüüdseks on inimestel leitud juba peaaegu 150 vermitud geeni. Need geenid on väga tugeva valikusurve all ning nad mõjutavad väga paljusid bioloogilisi protsesse ning tunnuseid nagu näiteks pre- ja postnataalset arengut, organismi ellujäämist ja kasvu, kognitiivsete võimete ja metaboolsete protsessidega seotud tunnuseid, ning mitmete haiguseid, milledest tuntuimad oleks Angelmani, Prader-Willi ja Beckwith-Weidemanni sündroomid. On oluline mainida, et kuna geneetilise vermimise korral on aktiivne ainult üks geenikoopia, siis iga mutatsioon vermitud geenis võib omada väga suurt mõju tervisele. Seetõttu on geneetilise vermimise mustrite uurimine väga oluline ka kliinilise geneetika seisukohalt, sest see aitab paremini mõista erinevate haiguste tekkemehhanisme.

Kuna geneetiline vermimine pannakse paika sugurakkudes, arvati pikka aega, et see muster püsib sarnasena kõikides kudedes. Uuemad tööd, mis on tehtud nii inimese kui ka katseloomade erinevaid rakke kasutades, on siiski näidanud, et vermitud geenide ekspressioonimuster sõltub nii koest kui ka organismi arengujärgust. Suurimaid erinevusi võiks oodata embrüonaalsete ja ekstrambrüonaalsete kudede ning täiskasvanu somaatiliste kudede vahel. Enamik vermitud geene on leitud platsentast või embrüonaalsetes kudedes ning on vähe informatsooni samade geenide ekspressioonimustrite kohta täiskasvanud organismi kudedes. Seetõttu on oluline uurida iga vermitud geeni puhul eraldi selle vermimismudelit erinevates kudedes eraldi, et oleks võimalik hinnata selle geeni mõju haigusele.

Leidmaks geneetilise vermimise koespetsiifilisust, uurisin juba teadaolevalt vermitud geenide ning nendega seotud vermimise kontrollregioonide (*ICR, imprinting control region*) metülatsioonimustreid kasutades täiskasvanud inimese organismi somaatilisi kudesid. Vermimise kontroll regioonid on suhteliselt väikesed genoomi regioonid, mis on erinevalt metüleeritud sõltuvalt vanema alleelist ning asuvad vermitud geenide lähedal või sees. Nende funktsiooniks on reguleerida vermitud klastris olevate geenide ekspressiooni. Seetõttu uurisin neid ka eraldi oma töös ja näitasin nende spetsiifilist metülatsiooni mustrit, mis erineb vermitud geenide omast. Vermitud geenide ja nendega seotud erinevalt metüleeritud regioonide uurimiseks kasutasime 18 somaatilist autopsia teel saadud kudet ning lisaks ka vereproove tervetelt doonoritelt. Kudede valikul lähtuti sellest, et katta kõik kolm lootelehte: endoderm, ektoderm ja mesoderm.

Kasutades inimese koeproove, leidsin, et vermitud geene iseloomustab promootori spetsiifilises piirkonnas rohkem poolmetüleeritud nukleotiide, mis võiks viidata vanema-spetsiifilisele metülatsioonile. Varasemad epigeneetilised uuringud on andnud vastuolulisi tulemusi, kui on kasutatud inimese täisvere proove, sest veri koosneb paljude erinevate rakkude segust, millede metülatsioonimuster võib olla erinev. Uurides vermitud geenide metülatsioonimustreid näitasin, et inimese verekoe proovid on siiski hea instrument vermitud geenide tuvastamiseks ning et metülatsioonimustrid on enamasti somaatiliste rakkude vahel püsivad. Lisaks leidsin, et ICR-id geeniregioonides, mis oli algselt vermitud sugurakkudes ehk ootsüütides ja spermatosoidides, võivad hilisemalt somaatilistes rakkudes oma poolmetüleeritud oleku kaotada. Sellist muutust nägin ennekõike platsenta-spetsiifiliste ICR-ide puhul. Hilisemad uuringud on seostanud poolmetüleeritud oleku kadu bialleelse ekspressiooni tekkega. Seetõttu võiks oletada, et geenid, mis on vermitud platsentas, võivad hilisemates arengustaadiumites omada bialleelset ekspressiooni.

Töö teises pooles olen otsinud uusi vermitud geene, mis mõjutavad erinevaid inimese fenotüüpe: kehamassiindeksit (KMI), vere metaboliite ja menarhe iga. Menarhe ea analüüsis leidsime assotsiatsioone mitmes teadaolevat vermitud geenide klastris, nagu näiteks *DLK1-WDR25* ja *MKRN3-MAGEL2* klastrid. Lisaks kinnitasime varem teadaolevaid seoseid, et vermitud geenid *KCNK9* ja *SLC2A10* on seotud nii KMI kui ka menarche eaga.

Kuna nii KMI kui ka menarhe iga on seotud ka metaboliitide tasemetega (eriti lipiidide tasemetega), siis viisin läbi uuringu, kus otsisin vanema-spetsiifilisi assotsiatsioone ka vere metaboliitidega. Leidsin, et *PTPRD* geen, mida meie varasem uuring on seostunud menarhe eaga, mõjutab vanemaspetsiifiliselt ka metaboliidi taset. Uuringus leidsin, et metaboliit "triglütseriidid, mis on seotud keskmise suurusega VLDL molekuliga" oli assotsieerunud geneetiliste variantidega *PTPRD* geeni piirkonnas. See metaboliit koosneb kahest molekulist – endogeenselt sekreteeritud triglütseriidist, mis on seotud VLDL molekuli tuumaga. See kompleks on rakkude energia allikaks ja uuringud on näidatud selle metaboliidi seost südamehaigustega. Seetõttu on väga oluline viia läbi edasiseid uuringuid kus uuritakse selle metaboliidi seoseid KMI ja menarhe eaga. Sellised uuringud võimaldavad tulevikus uurida vermitud geenides asuvate geneetiliste variantide mõju geeniekspressioonile konkreetse tunnuse või haiguse seisukohalt olulistes kudedes.

REFERENCES

Aapola, U., Shibuya, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Kawasaki, K., Minoshima, S., Krohn, K., et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics 65, 293–298.

Abbasi, A., Juszczyk, D., van Jaarsveld, C.H., and Gulliford, M.C. (2017). Body Mass Index and Incident Type 1 and Type 2 Diabetes in Children and Young Adults: A Retrospective Cohort Study. J. Endocr. Soc. 1, 524–537.

Abrink, M., Ortiz, J.A., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P., and Losson, R. (2001). Conserved interaction between distinct Kruppel-associated box domains and the transcriptional intermediary factor 1 beta. Proc Natl Acad Sci U S A 98, 1422–1426.

Amarasekera, M., Martino, D., Ashley, S., Harb, H., Kesper, D., Strickland, D., Saffery, R., and Prescott, S.L. (2014). Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. FASEB J. 28, 4068–4076.

Arima, T., Drewell, R. a, Arney, K.L., Inoue, J., Makita, Y., Hata, a, Oshimura, M., Wake, N., and Surani, M. a (2001). A conserved imprinting control region at the HYMAI/ZAC domain is implicated in transient neonatal diabetes mellitus. Hum. Mol. Genet. 10, 1475–1483.

Azur, M.J., Stuart, E.A., Frangakis, C., and Leaf, P.J. (2011). Multiple imputation by chained equations: What is it and how does it work? Int. J. Methods Psychiatr. Res. 20, 40–49.

Babak, T., DeVeale, B., Tsang, E.K., Zhou, Y., Li, X., Smith, K.S., Kukurba, K.R., Zhang, R., Li, J.B., van der Kooy, D., et al. (2015). Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. Nat. Publ. Gr. 47.

Bar, S., Schachter, M., Eldar-Geva, T., and Benvenisty, N. (2017). Large-Scale Analysis of Loss of Imprinting in Human Pluripotent Stem Cells. Cell Rep. 19, 957–968.

Baran, Y., Subramaniam, M., Biton, A., Tukiainen, T., Tsang, E.K., Rivas, M.A., Pirinen, M., Gutierrez-Arcelus, M., Smith, K.S., Kukurba, K.R., et al. (2015). The landscape of genomic imprinting across diverse adult human tissues. Genome Res. 25, 927–936.

Barel, O., Shalev, S.A., Ofir, R., Cohen, A., Zlotogora, J., Shorer, Z., Mazor, G., Finer, G., Khateeb, S., Zilberberg, N., et al. (2008). Maternally Inherited Birk Barel Mental Retardation Dysmorphism Syndrome Caused by a Mutation in the Genomically Imprinted Potassium Channel KCNK9. Am. J. Hum. Genet. 83, 193–199.

Barlow, D.P. (2011). Genomic imprinting: a mammalian epigenetic discovery model. Annu. Rev. Genet. 45, 379–403.

Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian genomic imprinting. Cold Spring Harb. Perspect. Biol. 3, 1–17.

Below, J.E., Gamazon, E.R., Morrison, J. V., Konkashbaev, A., Pluzhnikov, A., McKeigue, P.M., Parra, E.J., Elbein, S.C., Hallman, D.M., Nicolae, D.L., et al. (2011). Genome-wide association and meta-analysis in populations from Starr County, Texas, and Mexico City identify type 2 diabetes susceptibility loci and enrichment for expression quantitative trait loci in top signals. Diabetologia 54, 2047–2055.

Bilanges, B., Varrault, a, Mazumdar, a, Pantaloni, C., Hoffmann, a, Bockaert, J., Spengler, D., and Journot, L. (2001). Alternative splicing of the imprinted candidate tumor suppressor gene ZAC regulates its antiproliferative and DNA binding activities. Oncogene 20, 1246–1253.

Blois, S.M., Sulkowski, G., Tirado-González, I., Warren, J., Freitag, N., Klapp, B.F., Rifkin, D., Fuss, I., Strober, W., and Dveksler, G.S. (2014). Pregnancy-specific glycoprotein 1 (PSG1) activates TGF-β and prevents dextran sodium sulfate (DSS)-induced colitis in mice. Mucosal Immunol. 7, 348–358.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science (80-.). 294, 2536–2539.

Bratke, H., Bruserud, I.S., Brannsether, B., Aßmus, J., Bjerknes, R., Roelants, M., and Júlíusson, P.B. (2017). Timing of menarche in Norwegian girls: associations with body mass index, waist circumference and skinfold thickness. BMC Pediatr. 17, 138.

Breindel, J.L., Skibinski, A., Sedic, M., Wronski-Campos, A., Zhou, W., Keller, P.J., Mills, J., Bradner, J., Onder, T., and Kuperwasser, C. (2017). Epigenetic Reprogramming of Lineage-Committed Human Mammary Epithelial Cells Requires DNMT3A and Loss of DOT1L. Stem Cell Reports.

Cantão, I.H., Tesser, R.B., and Stumpp, T. (2017). An Initial Investigation of an Alternative Model to Study rat Primordial Germ Cell Epigenetic Reprogramming. Biol. Proced. Online 19, 9.

Chasman, D.I., Pare, G., Mora, S., Hopewell, J.C., Peloso, G., Clarke, R., Cupples, L.A., Hamsten, A., Kathiresan, S., Malarstig, A., et al. (2009). Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. PLoS Genet 5, e1000730.

Cheong, C.Y., Chng, K., Ng, S., Chew, S.B., Chan, L., and Ferguson-Smith, A.C. (2015). Germline and somatic imprinting in the nonhuman primate highlights species differences in oocyte methylation. Genome Res. 125, 611–623.

Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev. 23, 105–117.

Coombes, C., Arnaud, P., Gordon, E., Dean, W., Coar, E.A., Williamson, C.M., Feil, R., Peters, J., and Kelsey, G. (2003). Epigenetic properties and identification of an imprint mark in the Nesp-Gnasxl domain of the mouse Gnas imprinted locus. Mol. Cell. Biol. 23, 5475–5488.

Court, F., Martin-Trujillo, A., Romanelli, V., Garin, I., Iglesias-Platas, I., Salafsky, I., Guitart, M., Perez de Nanclares, G., Lapunzina, P., and Monk, D. (2013). Genome-Wide Allelic Methylation Analysis Reveals Disease-Specific Susceptibility to Multiple Methylation Defects in Imprinting Syndromes. Hum. Mutat. 34, 595–602.

Court, F., Tayama, C., Romanelli, V., Martin-Trujillo, A., Iglesias-Platas, I., Okamura, K., Sugahara, N., Sim??n, C., Moore, H., Harness, J. V., et al. (2014a). Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. Genome Res. 24, 554–569.

Court, F., Camprubi, C., Garcia, C., Guillaumet-Adkins, A., Sparago, A., Seruggia, D., Sandoval, J., Esteller, M., Martin-Trujillo, A., Riccio, A., et al. (2014b). The PEG13-DMR and brain-specific enhancers dictate imprinted expression within the 8q24 intellectual disability risk locus. Epigenetics Chromatin 7, 5.

Dall, G.V., and Britt, K.L. (2017). Estrogen Effects on the Mammary Gland in Early and Late Life and Breast Cancer Risk. Front. Oncol. 7.

Davis, T.L., Yang, G.J., McCarrey, J.R., and Bartolomei, M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum. Mol. Genet. 9, 2885–2894.

Delaval, K., Govin, J., Cerqueira, F., Rousseaux, S., Khochbin, S., and Feil, R. (2007). Differential histone modifications mark mouse imprinting control regions during spermatogenesis. EMBO J. 26, 720–729.

Deming, Y., Li, Z., Kapoor, M., Harari, O., Del-Aguila, J.L., Black, K., Carrell, D., Cai, Y., Fernandez, M.V., Budde, J., et al. (2017). Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. Acta Neuropathol. 133, 839–856.

Dent, C.L., and Isles, A.R. (2014). Brain-expressed imprinted genes and adult behaviour: The example of Nesp and Grb10. Mamm. Genome 25, 87–93.

Edwards, C.A., and Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. Curr. Opin. Cell Biol. 19, 281–289.

Fauque, P., Jouannet, P., Lesaffre, C., Ripoche, M.A., Dandolo, L., Vaiman, D., and Jammes, H. (2007). Assisted reproductive technology affects developmental kinetics, H19 imprinting control region methylation and H19 gene expression in individual mouse. BMC Dev. Biol. 7, 19.

Ferguson-Smith, A.C., and Greally, J.M. (2007). Epigenetics: perceptive enzymes. Nature 449, 148–149.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J. (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. 10, 2067–2078.

Garfield, A.S., Cowley, M., Smith, F.M., Moorwood, K., Stewart-Cox, J.E., Gilroy, K., Baker, S., Xia, J., Dalley, J.W., Hurst, L.D., et al. (2011). Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. Nature 469, 534–538.

Gkountela, S., Zhang, K.X., Shafiq, T.A., Liao, W.W., Hargan-Calvopiña, J., Chen, P.Y., and Clark, A.T. (2015). DNA demethylation dynamics in the human prenatal germline. Cell 161, 1425–1436.

Gleason, G., Liu, B., Bruening, S., Zupan, B., Auerbach, A., Mark, W., Oh, J.-E., Gal-Toth, J., Lee, F., and Toth, M. (2010). The serotonin1A receptor gene as a genetic and prenatal maternal environmental factor in anxiety. Proc. Natl. Acad. Sci. U. S. A. 107, 7592–7597.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., et al. (2015). The transcriptome and DNA methylome landscapes of human primordial germ cells. Cell 161, 1437–1452.

Hager, R., Cheverud, J.M., and Wolf, J.B. (2008). Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. Genetics 178, 1755–1762.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science (80-.). 329, 78–82.

Hanna, C.W., Pe??aherrera, M.S., Saadeh, H., Andrews, S., McFadden, D.E., Kelsey, G., and Robinson, W.P. (2016). Pervasive polymorphic imprinted methylation in the human placenta. Genome Res. 26, 756–767.

Hanssen, L.L.P., Kassouf, M.T., Oudelaar, A.M., Biggs, D., Preece, C., Downes, D.J., Gosden, M., Sharpe, J.A., Sloane-Stanley, J.A., Hughes, J.R., et al. (2017). Tissue-

specific CTCF–cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. Nat. Cell Biol.

Hassan, H.M., Kolendowski, B., Isovic, M., Bose, K., Dranse, H.J., Sampaio, A. V., Underhill, T.M., and Torchia, J. (2017). Regulation of Active DNA Demethylation through RAR-Mediated Recruitment of a TET/TDG Complex. Cell Rep. 19, 1685–1697.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 129, 1983–1993.

Hayward, B.E., Kamiya, M., Strain, L., Moran, V., Campbell, R., Hayashizaki, Y., and Bonthron, D.T. (1998). The human GNAS1 gene is imprinted and encodes distinct paternally and biallelically expressed G proteins. Proc Natl Acad Sci U S A 95, 10038–10043.

Hayward, B.E., Barlier, A., Korbonits, M., Grossman, A.B., Jacquet, P., Enjalbert, A., and Bonthron, D.T. (2001). Imprinting of the Gsα gene GNAS1 in the pathogenesis of acromegaly. J. Clin. Invest. 107.

Hikichi, T., Kohda, T., Kaneko-Ishino, T., and Ishino, F. (2003). Imprinting regulation of the murine Meg1/Grb10 and human GRB10 genes; roles of brain-specific promoters and mouse-specific CTCF-binding sites. Nucleic Acids Res. 31, 1398–1406.

Hoggart, C.J., Venturini, G., Mangino, M., Gomez, F., Ascari, G., Zhao, J.H., Teumer, A., Winkler, T.W., T??ernikova, N., Luan, J., et al. (2014). Novel Approach Identifies SNPs in SLC2A10 and KCNK9 with Evidence for Parent-of-Origin Effect on Body Mass Index. PLoS Genet. 10, 1–12.

Huang, W.C., Ferris, E., Cheng, T., Hörndli, C.S., Gleason, K., Tamminga, C., Wagner, J.D., Boucher, K.M., Christian, J.L., and Gregg, C. (2017). Diverse Non-genetic, Allele-Specific Expression Effects Shape Genetic Architecture at the Cellular Level in the Mammalian Brain. Neuron 93, 1094–1109.e7.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science (80-.). 333, 1300–1303.

Johnson, A.D., Yanek, L.R., Chen, M.-H., Faraday, N., Larson, M.G., Tofler, G., Lin, S.J., Kraja, A.T., Province, M.A., Yang, Q., et al. (2010). Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat. Genet. 42, 608–613.

Kabir, N.N., and Kazi, J.U. (2014). Grb10 is a dual regulator of receptor tyrosine kinase signaling. Mol. Biol. Rep. 41, 1985–1992.

Kagami, M., O'Sullivan, M.J., Green, A.J., Watabe, Y., Arisaka, O., Masawa, N., Matsuoka, K., Fukami, M., Matsubara, K., Kato, F., et al. (2010). The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: Hierarchical interaction and distinct functional properties as imprinting control centers. PLoS Genet. 6, 1–13.

Kamiya, M., Judson, H., Okazaki, Y., Kusakabe, M., Muramatsu, M., Takada, S., Takagi, N., Arima, T., Wake, N., Kamimura, K., et al. (2000). The cell cycle control gene ZAC/PLAGL1 is imprinted--a strong candidate gene for transient neonatal diabetes. Hum. Mol. Genet. 9, 453–460.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 429, 900–903.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum. Mol. Genet. 16, 2272–2280.

Kelsey, G. (2010). Imprinting on chromosome 20: Tissue-specific imprinting and imprinting mutations in the GNAS locus. Am. J. Med. Genet. Part C Semin. Med. Genet. 154, 377–386.

Kelsey, G., and Feil, R. (2013a). New insights into establishment and maintenance of DNA methylation imprints in mammals. Philos. Trans. R. Soc. B Biol. Sci. 368.

Kelsey, G., and Feil, R. (2013b). New insights into establishment and maintenance of DNA methylation imprints in mammals. Philos. Trans. R. Soc. B Biol. Sci. 368, 20110336.

Kettunen, J., Demirkan, A., Würtz, P., Draisma, H.H.M.M., Haller, T., Rawal, R., Vaarhorst, A., Kangas, A.J., Lyytikäinen, L.-P., Pirinen, M., et al. (2016). Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat. Commun. 7, 11122.

Keverne, E.B. (2015). Genomic imprinting, action, and interaction of maternal and fetal genomes. Proc Natl Acad Sci U S A 112, 6834–6840.

Kim, H.-N., Roh, S.-J., Sung, Y.A., Chung, H.W., Lee, J.-Y., Cho, J., Shin, H., and Kim, H.-L. (2013). Genome-wide association study of the five-factor model of personality in young Korean women. J. Hum. Genet. 58, 667–674.

Kim, Y., Bang, H., and Kim, D. (2000). TASK-3, a new member of the tandem pore K+ channel family. J. Biol. Chem. 275, 9340–9347.

Korostowski, L., Raval, A., Breuer, G., and Engel, N. (2011). Enhancer-driven chromatin interactions during development promote escape from silencing by a long non-coding RNA. Epigenetics Chromatin 4, 21.

Lash, A.E., Tolstoshev, C.M., Wagner, L., Schuler, G.D., Strausberg, R.L., Riggins, G.J., and Altschul, S.F. (2000). SAGEmap: A public gene expression resource. Genome Res. 10, 1051–1060.

Lawson, H.A., Cheverud, J.M., and Wolf, J.B. (2013). Genomic imprinting and parent-of-origin effects on complex traits. Nat. Rev. Genet. 14, 609–617.

Leinonen, J.T., Surakka, I., Havulinna, A.S., Kettunen, J., Luoto, R., Salomaa, V., and Widén, E. (2012). Association of LIN28B with Adult Adiposity-Related Traits in Females. PLoS One 7.

Li, Y., and Sasaki, H. (2011). Genomic imprinting in mammals: its life cycle, molecular mechanisms and reprogramming. Nat. Publ. Gr. 21, 466–473.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. Dev. Cell 15, 547–557.

Lin, S., Ferguson-Smith, A.C., Schultz, R.M., and Bartolomei, M.S. (2011). Nonallelic Transcriptional Roles of CTCF and Cohesins at Imprinted Loci. Mol. Cell. Biol. 31, 3094–3104.

Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forné, T., Murrell, A., Constância, M., Bartolomei, M., Walter, J., et al. (2003). Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. Hum. Mol. Genet. 12, 295–305.

Luedi, P.P., Dietrich, F.S., Weidman, J.R., Bosko, J.M., Jirtle, R.L., and Hartemink, A.J. (2007). Computational and experimental identification of novel human imprinted genes. Genome Res. 17, 1723–1730.

Mackay, D.J.G., Coupe, A.M., Shield, J.P.H., Storr, J.N.P., Temple, I.K., and Robinson, D.O. (2002). Relaxation of imprinted expression of ZAC and HYMAI in a patient with transient neonatal diabetes mellitus. Hum. Genet. 110, 139–144.

Maksimovic, J., Gordon, L., and Oshlack, A. (2012). SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol. 13, R44.

Mantovani, G., Ballare, E., Giammona, E., Beck-Peccoz, P., and Spada, A. (2002). The Gsα gene: Predominant maternal origin of transcription in human thyroid gland and gonads. J. Clin. Endocrinol. Metab. 87, 4736–4740.

Mantovani, G., Bondioni, S., Locatelli, M., Pedroni, C., Lania, A.G., Ferrante, E., Filopanti, M., Beck-Peccoz, P., and Spada, A. (2004). Biallelic expression of the Gsalpha gene in human bone and adipose tissue. J. Clin. Endocrinol. Metab. 89, 6316–6319.

Market-Velker, B.A., Zhang, L., Magri, L.S., Bonvissuto, A.C., and Mann, M.R.W. (2009). Dual effects of superovulation: Loss of maternal and paternal imprinted methylation in a dose-dependent manner. Hum. Mol. Genet. 19, 36–51.

Martinet, C., Monnier, P., Louault, Y., Benard, M., Gabory, A., and Dandolo, L. (2016). H19 controls reactivation of the imprinted gene network during muscle regeneration. Development 143, 962–971.

Mattheisen, M., Samuels, J.F., Wang, Y., Greenberg, B.D., Fyer, A.J., McCracken, J.T., Geller, D.A., Murphy, D.L., Knowles, J.A., Grados, M.A., et al. (2015). Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. Mol. Psychiatry 20, 337–344.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. 28, 812–828.

Mihaylova, B., Emberson, J., Blackwell, L., Keech, A., Simes, J., Barnes, E.H., Voysey, M., Gray, A., Collins, R., Baigent, C., et al. (2012). The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: Meta-analysis of individual data from 27 randomised trials. Lancet 380, 581–590.

Mohammad, F., Mondal, T., Guseva, N., Pandey, G.K., and Kanduri, C. (2010). Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development 137, 2493–2499.

Monk, D. (2015). Germline-derived DNA methylation and early embryo epigenetic reprogramming: The selected survival of imprints. Int. J. Biochem. Cell Biol. 67, 128–138.

Monk, D., Arnaud, P., Apostolidou, S., Hills, F.A., Kelsey, G., Stanier, P., Feil, R., and Moore, G.E. (2006). Limited evolutionary conservation of imprinting in the human placenta. Proc. Natl. Acad. Sci. 103, 6623–6628.

Monk, D., Arnaud, P., Frost, J., Hills, F.A., Stanier, P., Feil, R., and Moore, G.E. (2009). Reciprocal imprinting of human GRB10 in placental trophoblast and brain: Evolutionary conservation of reversed allelic expression. Hum. Mol. Genet. 18, 3066–3074.

Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., and Dandolo, L. (2013). H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. Proc. Natl. Acad. Sci. 110, 20693–20698.

Morcos, L., Ge, B., Koka, V., Lam, K.C., Pokholok, D.K., Gunderson, K.L., Montpetit, A., Verlaan, D.J., and Pastinen, T. (2011). Genome-wide assessment of imprinted expression in human cells. Genome Biol. 12, R25.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: Implications for epigenetic reprogramming. J. Biol. Chem. 279, 52353–52360.

Mukhopadhyay, A., Ravikumar, G., Dwarkanath, P., Meraaj, H., Thomas, A., Crasta, J., Thomas, T., Kurpad, A. V., and Sridhar, T.S. (2015). Placental expression of the insulin receptor binding protein GRB10: Relation to human fetoplacental growth and fetal gender. Placenta 36, 1225–1230.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., et al. (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat. Cell Biol. 9, 64–71.

Ong, K.K., Elks, C.E., Wills, A.K., Wong, A., Wareham, N.J., Loos, R.J.F., Kuh, D., and Hardy, R. (2011). Associations between the pubertal timing-related variant in LIN28B and BMI vary across the life course. J. Clin. Endocrinol. Metab. 96.

Pauler, F.M., and Barlow, D.P. (2006). Imprinting mechanisms--it only takes two. Genes Dev. 20, 1203–1206.

Pearson, C.E. (2003). Slipping while sleeping? Trinucleotide repeat expansions in germ cells. Trends Mol. Med. 9, 490–495.

Perez, J.D., Rubinstein, N.D., Fernandez, D.E., Santoro, S.W., Needleman, L.A., Ho-Shing, O., Choi, J.J., Zirlinger, M., Chen, S.K., Liu, J.S., et al. (2015). Quantitative and functional interrogation of parent-of-origin allelic expression biases in the brain. Elife 4.

Perry, J.R., Day, F., Elks, C.E., Sulem, P., Thompson, D.J., Ferreira, T., He, C., Chasman, D.I., Esko, T., Thorleifsson, G., et al. (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 514, 92–97.

Peters, J. (2014). The role of genomic imprinting in biology and disease: an expanding view. Nat. Publ. Gr. 15, 517–530.

Peters, J., and Williamson, C.M. (2008). Control of imprinting at the Gnas cluster. Adv. Exp. Med. Biol. 626, 16–26.

Peters, J., Wroe, S.F., Wells, C.A., Miller, H.J., Bodle, D., Beechey, C. V, Williamson, C.M., and Kelsey, G. (1999). A cluster of oppositely imprinted transcripts at the Gnas locus in the distal imprinting region of mouse chromosome 2. Proc. Natl. Acad. Sci. U. S. A. 96, 3830–3835.

Peters, J., Holmes, R., Monk, D., Beechey, C. V., Moore, G.E., and Williamson, C.M. (2006). Imprinting control within the compact Gnas locus. In Cytogenetic and Genome Research, pp. 194–201.

Piccolo, F.M., Bagci, H., Brown, K.E., Landeira, D., Soza-Ried, J., Feytout, A., Mooijman, D., Hajkova, P., Leitch, H.G., Tada, T., et al. (2013). Different Roles for Tet1 and Tet2 Proteins in Reprogramming-Mediated Erasure of Imprints Induced by EGC Fusion. Mol. Cell 49, 1023–1033.

Piras, G., El Kharroubi, a, Kozlov, S., Escalante-Alcalde, D., Hernandez, L., Copeland, N.G., Gilbert, D.J., Jenkins, N. a, and Stewart, C.L. (2000). Zac1 (Lot1), a potential tumor suppressor gene, and the gene for epsilon-sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. Mol. Cell. Biol. 20, 3308–3315.

Plasschaert, R.N., and Bartolomei, M.S. (2015). Tissue-specific regulation and function of Grb10 during growth and neuronal commitment. Proc. Natl. Acad. Sci. 112, 6841–6847.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature 463, 1101–1105.

Prickett, A.R., and Oakey, R.J. (2012). A survey of tissue-specific genomic imprinting in mammals. Mol. Genet. Genomics 287, 621–630.

Proudhon, C., Duffié, R., Ajjan, S., Cowley, M., Iranzo, J., Carbajosa, G., Saadeh, H., Holland, M.L., Oakey, R.J., Rakyan, V.K., et al. (2012). Protection against De Novo Methylation Is Instrumental in Maintaining Parent-of-Origin Methylation Inherited from the Gametes. Mol. Cell 47, 909–920.

Rajan, S., Wischmeyer, E., Liu, G.X., Preisig-Müller, R., Daut, J., Karschin, A., and Derst, C. (2000). TASK-3, a novel tandem pore domain acid-sensitive K+ channel. An extracellular histidine as pH sensor. J. Biol. Chem. 275, 16650–16657.

Reik, W., and Dean, W. (2001). DNA methylation and mammalian epigenetics. Electrophoresis 22, 2838–2843.

Reik, W., and Walter, J. (2001). Genomic Imprinting : Parental Influence on the Genome. 2.

Renfree, M.B., Suzuki, S., Kaneko-Ishino, T., and B, P.T.R.S. (2013). The origin and evolution of genomic imprinting and viviparity in mammals. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 20120151.

Riso, V., Cammisa, M., Kukreja, H., Anvar, Z., Verde, G., Sparago, A., Acurzio, B., Lad, S., Lonardo, E., Sankar, A., et al. (2016). ZFP57 maintains the parent-of-origin-specific expression of the imprinted genes and differentially affects non-imprinted targets in mouse embryonic stem cells. Nucleic Acids Res. 44, 8165–8178.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A., and Jones, P.A. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: Coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res. 27, 2291–2298.

Sacks, F.M. (2015). The crucial roles of apolipoproteins E and C-III in apoB lipoprotein metabolism in normolipidemia and hypertriglyceridemia. Curr. Opin. Lipidol. 26, 56–63.

Sakamoto, A., Liu, J., Greene, A., Chen, M., and Weinstein, L.S. (2004). Tissue-specific imprinting of the G protein Gsα is associated with tissue-specific differences in histone methylation. Hum. Mol. Genet. 13, 819–828.

Sanchez-Delgado, M., Court, F., Vidal, E., Medrano, J., Monteagudo-Sánchez, A., Martin-Trujillo, A., Tayama, C., Iglesias-Platas, I., Kondova, I., Bontrop, R., et al. (2016). Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting. PLoS Genet. 12.

Sanz, L. a, Chamberlain, S., Sabourin, J.-C., Henckel, A., Magnuson, T., Hugnot, J.-P., Feil, R., and Arnaud, P. (2008). A mono-allelic bivalent chromatin domain controls tissue-specific imprinting at Grb10. EMBO J. 27, 2523–2532.

Schosser, A., Butler, A.W., Uher, R., Ng, M.Y., Cohen-Woods, S., Craddock, N., Owen, M.J., Korszun, A., Gill, M., Rice, J., et al. (2013). Genome-wide association study of co-occurring anxiety in major depression. World J. Biol. Psychiatry 14, 611–621.

Schulz, R., Woodfine, K., Menheniott, T.R., Bourc'his, D., Bestor, T., and Oakey, R.J. (2008). Wamidex: A web atlas of murine genomic imprinting and differential expression. Epigenetics 3, 89–96.

Scott, R.A., Scott, L.J., Mägi, R., Marullo, L., Gaulton, K.J., Kaakinen, M., Pervjakova, N., Pers, T.H., Johnson, A.D., Eicher, J.D., et al. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes 66, 2888–2902.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014). Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell 15, 459–470.

Smith, R.J., Arnaud, P., Konfortova, G., Dean, W.L., Beechey, C. V., and Kelsey, G. (2002). The mouse Zac1 locus: Basis for imprinting and comparison with human ZAC. Gene 292, 101–112.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human pre-implantation embryo. Nature 511, 611–615.

Soininen, P., Kangas, A.J., Wurtz, P., Tukiainen, T., Tynkkynen, T., Laatikainen, R., Jarvelin, M.-R., Kahonen, M., Lehtimaki, T., Viikari, J., et al. (2009). High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 134, 1781–1785.

Speed, D., Hoggart, C., Petrovski, S., Tachmazidou, I., Coffey, A., Jorgensen, A., Eleftherohorinou, H., De Iorio, M., Todaro, M., De, T., et al. (2014). A genome-wide association study and biological pathway analysis of epilepsy prognosis in a prospective cohort of newly treated epilepsy. Hum. Mol. Genet. 23, 247–258.

Stark, O., Peckham, C.S., and Moynihan, C. (1989). Weight and age at menarche. Arch Dis Child 64, 383–387.

Stringer, J.M., Suzuki, S., Pask, A.J., Shaw, G., and Renfree, M.B. (2012). GRB10 imprinting is eutherian mammal specific. In Molecular Biology and Evolution, pp. 3711–3719.

Surani, M. a, Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nat. Educ. 308, 548–550.

Svensson, A.C., Sandin, S., Cnattingius, S., Reilly, M., Pawitan, Y., Hultman, C.M., and Lichtenstein, P. (2009). Maternal effects for preterm birth: A genetic epidemiologic study of 630,000 families. Am. J. Epidemiol. 170, 1365–1372.

Takahashi, N., Gray, D., Strogantsev, R., Noon, A., Delahaye, C., Skarnes, W.C., Tate, P.H., and Ferguson-Smith, A.C. (2016). ZFP57 and the targeted maintenance of postfertilization genomic imprints. Cold Spring Harb. Symp. Quant. Biol. 80, 177–187.

Talma, H., Schönbeck, Y., van Dommelen, P., Bakker, B., van Buuren, S., and HiraSing, R.A. (2013). Trends in Menarcheal Age between 1955 and 2009 in the Netherlands. PLoS One 8.

Tang, W.W.C., Dietmann, S., Irie, N., Leitch, H.G., Floros, V.I., Bradshaw, C.R., Hackett, J.A., Chinnery, P.F., and Surani, M.A. (2015). A unique gene regulatory network resets the human germline epigenome for development. Cell 161, 1453–1467.

Tomé, S., Panigrahi, G.B., Castel, A.L., Foiry, L., Melton, D.W., Gourdon, G., and Pearson, C.E. (2011). Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. Hum. Mol. Genet. 20, 2131–2143.

Traglia, M., Croen, L.A., Lyall, K., Windham, G.C., Kharrazi, M., DeLorenze, G.N., Torres, A.R., and Weiss, L.A. (2017). Independent Maternal and Fetal Genetic Effects on Mid-gestational Circulating Levels of Environmental Pollutants. G3 (Bethesda). g3.117.039784.

Traylor, M., Zhang, C.R., Adib-Samii, P., Devan, W.J., Parsons, O.E., Lanfranconi, S., Gregory, S., Cloonan, L., Falcone, G.J., Radmanesh, F., et al. (2016). Genome-wide meta-analysis of cerebral white matter hyperintensities in patients with stroke. Neurology 86, 146–153.

Tsai, F.-J.J., Yang, C.-F.F., Chen, C.-H.C.H.C.-C.C.H., Chuang, L.-M.M., Lu, C.-H.H., Chang, C.-C.T.C.-T.C., Wang, T.-Y.Y., Chen, R.-H.H., Shiu, C.-F.F., Liu, Y.-M.M., et al. (2010). A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet 6, e1000847.

Twig, G., Ben-Ami Shor, D., Furer, A., Levine, H., Derazne, E., Goldberger, N., Haklai, Z., Levy, M., Afek, A., Leiba, A., et al. (2017). Adolescent body mass index and cardiovascular disease-specific mortality by midlife. J. Clin. Endocrinol. Metab. 102, 3011–3020.

Ueda, T., Abe, K., Miura, A., Yuzuriha, M., Zubair, M., Noguchi, M., Niwa, K., Kawase, Y., Kono, T., Matsuda, Y., et al. (2000). The paternal methylation imprint of the mouse H19 locus is acquired in the gonocyte stage during foetal testis development. Genes to Cells 5, 649–659.

Uetani, N., Kato, K., Ogura, H., Mizuno, K., Kawano, K., Mikoshiba, K., Yakura, H., Asano, M., and Iwakura, Y. (2000). Impaired learning with enhanced hippocampal long-term potentiation in PTPδ-deficient mice. Eur. Mol. Biol. Organ. J. 19, 2775–2785.

Valleley, E.M., Cordery, S.F., and Bonthron, D.T. (2007). Tissue-specific imprinting of the ZAC/PLAGL1 tumour suppressor gene results from variable utilization of monoallelic and biallelic promoters. Hum. Mol. Genet. 16, 972–981.

Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Aknin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., et al. (2006). Zac1 Regulates an Imprinted Gene Network Critically Involved in the Control of Embryonic Growth. Dev. Cell 11, 711–722.

Varrault, A., Dantec, C., Le Digarcher, A., Chotard, L., Bilanges, B., Parrinello, H., Dubois, E., Rialle, S., Severac, D., Bouschet, T., et al. (2017). Identification of Plagl1/Zac1 binding sites and target genes establishes its role in the regulation of extracellular matrix genes and the imprinted gene network. Nucleic Acids Res. 45, 10466–10480.

von Meyenn, F., and Reik, W. (2015). Forget the Parents: Epigenetic Reprogramming in Human Germ Cells. Cell 161, 1248–1251.

Vu, T.H., Li, T., and Hoffman, A.R. (2004). Promoter-restricted histone code, not the differentially methylated DNA regions or antisense transcripts, marks the imprinting status of IGF2R in human and mouse. Hum. Mol. Genet. 13, 2233–2245.

Vu, T.H., Jirtle, R.L., and Hoffman, A.R. (2006). Cross-species clues of an epigenetic imprinting regulatory code for the IGF2R gene. In Cytogenetic and Genome Research, pp. 202–208.

Wang, Z., Dang, S., and Xing, Y. (2016). Correlation of body mass index levels with menarche in adolescent girls in Shaanxi, China: a cross sectional study. BMC Womens. Health 16, 61.

Weinstein, L.S., Yu, S., Warner, D.R., and Liu, J. (2001). Endocrine manifestations of stimulatory G protein α subunit mutations and the role of genomic imprinting. Endocr. Rev. 22, 675–705.

Williamson, C.M., Ball, S.T., Nottingham, W.T., Skinner, J. a, Plagge, A., Turner, M.D., Powles, N., Hough, T., Papworth, D., Fraser, W.D., et al. (2004). A cis-acting

control region is required exclusively for the tissue-specific imprinting of Gnas. Nat. Genet. 36, 894–899.
Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., et al. (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. Nat. Genet. 38, 350–355.
Winkelmann, J., Czamara, D., Schormair, B., Knauf, F., Schulte, E.C., Trenkwalder, C., Dauvilliers, Y., Polo, O., Högl, B., Berger, K., et al. (2011). Genome-Wide association study identifies novel restless legs syndrome susceptibility loci on 2p14 and 16q12.1. PLoS Genet. 7.
Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethyl-cytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. 2, 241.
Wu, S.C., and Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. Nat. Rev. Mol. Cell Biol. 11, 607–620.
Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet. 18, 517–534.
Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., and Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. Nature 504, 460–464.
Yamashita, R., Sugano, S., Suzuki, Y., and Nakai, K. (2012). DBTSS: DataBase of Transcriptional Start Sites progress report in 2012. Nucleic Acids Res. 40.
Yan, X., Himburg, H.A., Pohl, K., Quarmyne, M., Tran, E., Zhang, Y., Fang, T., Kan, J., Chao, N.J., Zhao, L., et al. (2016). Deletion of the Imprinted Gene Grb10 Promotes Hematopoietic Stem Cell Self-Renewal and Regeneration. Cell Rep. 17, 1584–1594.
Yu, S., Yu, D., Lee, E., Eckhaus, M., Lee, R., Corria, Z., Accili, D., Westphal, H., and Weinstein, L.S. (1998). Variable and tissue-specific hormone resistance in heterotrimeric Gs protein alpha-subunit (Gsalpha) knockout mice is due to tissue-specific imprinting of the gsalpha gene. Proc. Natl. Acad. Sci. U. S. A. 95, 8715–8720.
Zheng, H.Y., Tang, Y., Niu, J., Li, P., Ye, D.S., Chen, X., Shi, X.Y., Li, L., and Chen, S.L. (2013). Aberrant DNA methylation of imprinted loci in human spontaneous abortions after assisted reproduction techniques and natural conception. Hum. Reprod. 28, 265–273.

WEB pages:

MouseBook Imprinting Catalogue:
http://www.mousebook.org/mousebook-catalogs/imprinting-resource

ACKNOWLEDGMENTS

I would like to thank my great supervisor Dr Reedik Mägi for introducing me into the terrific world of genomics and bioinformatics. I am very grateful for your invaluable guidance and help over these years. I am very happy that I found not only the greatest supervisor in this universe, but also a wonderful friend with opposite views on life (except cats). I know this thesis is also a big thing for you and I hope you get a chance to open THAT bottle very soon.

My sincere thanks also go to my supervisor Prof Andres Metspalu. Thank you for being a fantastic director of the Estonian Genome Center, making the atmosphere here comfortable and cosy, and at the same time keeping the highest scientific level.

My deepest gratitude goes to my supervisor, Prof Andrew Paul Morris. I can barely express how much I appreciate your help, your trust and patience, all the warm words and the incredible English language lessons.

The Estonian Genome Center would be nothing without people who work here. I remember the time when I was the only Russian-speaking student in the Center who almost did not speak Estonian, but I always felt welcome and a part of a big friendly family. I am really happy to thank each of you for your support that you gave me every time I needed it.

I am very grateful to Prof Markus Perola and his group for all the great moments in Finland. I really miss you guys, these three and a half years together went as one day.

I would also thank Prof Andres Salumets and Prof Pärt Peterson for introducing me into the world of methylation and first steps of academic writing.

My incredible supervisor Dr Inga Prokopenko, thank you for guiding me through the human side of the scientific world, thank you for taking care of me. I hope we will have more great days in Oxford and London soon with you and your girls.

My dearest friends, you are the people with whom we can discuss any topic while having at least 3 pots of tea. I am very thankful for your support and the superpower to find a funny side in any difficult situation. This planet is in safe hands as long as the world has us. My 25 sorts of tea are waiting for more exciting talks in the kitchen at 3 am.

Last but not least, I would like to thank my family. My special thanks go to my lovely parents, who taught me to be strong, ambitious and pragmatic, to my older brother, who is always very proud of his little sister, to all our dogs and cats for supportive tail-waving and purr-purr moments. During these years I also got amazing parents-in-law. Now we can have even more fascinating conversations about natural history, evolution, and space exploration.

And of course all my deepest and warmest words belong to you, my dear husband, Juri. Thank you for making me so happy and sharing this journey with me.

PUBLICATIONS

CURRICULUM VITAE

Name:	Natalia Pervjakova
Date of birth:	November 27, 1987
Contact:	Estonian Genome Center, Institute of Genomics, University of Tartu
E-mail:	natalia.pervjakova@ut.ee

Education:

2012–Present	PhD studies, Molecular Biomedicine, Molecular and Cell Biology, Faculty of Science and technology, University of Tartu, Estonia
2010–2012	Master's studies, Biotechnology and Biomedicine, Faculty of Science and Technology, University of Tartu, Estonia
2007–2010	Bachelor's studies, Gene Technology, Faculty of Science and Technology, University of Tartu, Estonia

Professional employment:

2017–Present	Visiting researcher in Imperial College of Science, Dept. of Genomics of Common Disease, London, UK
2017–2017	Research assistant in Imperial College of Science, Dept. of Genomics of Common Disease, London, UK
2011–Present	Specialist in Estonian Genome Center, Institute of Genomics, University of Tartu, Estonia
2012–2015	Researcher in National Institute for Health and Welfare, Helsinki, Finland
2008–2010	Lab Assistant, Department of Biochemistry, Institute of Molecular and Cell Biology, University of Tartu, Estonia

Administrative work:

2016–Present	Member of American Society of Human Genetics
2015–Present	Teaching staff of the course "Introduction to the statistical analysis of genetic association studies", Imperial College of Science, Technology and Medicine, London, UK.
2014–2014	Teaching staff of the 'Omic's practical course' in National Institute for Health and Welfare, Helsinki, Finland

Publications:

Pervjakova N., Kukushkina V., Haller T., Kasela S., Joensuu A., Kristiansson K., Annilo T., Perola M., Salomaa V., Jousilahti P., Metspalu A., Mägi R., (2017). Genome-wide analysis of NMR metabolites revealed parent-of-origin effect on triglycerides in medium VLDL in *PTPRD* gene. Biomark Med. 2018 Mar 14. doi: 10.2217/bmm-2018-0020.

Pervjakova N, Prokopenko I. (2017). The TAD-pathway for GWAS signals. Eur J Hum Genet 25.

Rosticci M*., **Pervjakova N***., Kaakinen M*., Cicero AF., Pfeufer A., Marullo L., Mägi R., Fischer K., Jiang L., D'Addato S., et al., (2017). Common variants in *HMGCR* affect blood Apolipoprotein B levels – accepted to Biomarkers in Medicine

Nuotio ML., **Pervjakova N**., Joensuu A., Karhunen V., Milani L., Kettunen J., Järvelin MR., Metspalu A., Salomaa V., Kristiansson K., Perola M. (2017). Epigenome-wide association study (EWAS) in metabolic syndrome and its components – under review in Diabetes

Scott RA., Scott LJ., Mägi R., Marullo L., Gaulton KJ., Kaakinen M., **Pervjakova N**., Pers TH., Johnson AD., Eicher JD., et al. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes. 2017 Nov;66(11):2888–2902. doi: 10.2337/db16-1253. Epub 2017 May 31.

Day FR., Thompson DJ., Helgason H., Chasman DI., Finucane H., Sulem P., Ruth KS., Whalen S., Sarkar AK., Albrecht E.,…, **Pervjakova N**., et al. (2017). Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nat Genet. 2017 Jun;49(6):834-841. doi: 10.1038/ng.3841. Epub 2017 Apr 24.

Justice AE., Winkler TW., Feitosa MF., Graff M., Fisher VA., Young K., Barata L., Deng X., Czajkowski J., Hadley D.,…, **Pervjakova N**., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat Commun. 2017 Apr 26;8:14977. doi: 10.1038/ncomms14977.

Wahl S., Drong A., Lehne B., Loh M., Scott WR., Kunze S., Tsai P-C., Ried JS., Zhang W., Yang Y.,…,**Pervjakova N**., et al. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017 Jan 5;541(7635):81-86. doi: 10.1038/nature20784. Epub 2016 Dec 21.

Okbay A., Beauchamp JP., Fontana MA., Lee JJ., Pers TH., Rietveld CA., Turley P., Chen G-B., Emilsson V., Meddens SFW.,…**Pervjakova N**., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. Nature. 2016 May 26;533(7604):539-42. doi: 10.1038/nature17671. Epub 2016 May 11.

Pervjakova N, Kasela S, Morris AP, Kals M, Metspalu A, Lindgren CM, Salumets A, Mägi R. (2016). Imprinted genes and imprinting control regions show predominant intermediate methylation in adult somatic tissues. Epigenomics. 2016 Jun;8(6):789-99. doi: 10.2217/epi.16.8. Epub 2016 Mar 23

Winkler TW., Justice AE., Graff M., Barata L., Feitosa MF., Chu S., Czajkowski J., Esko T., Fall T., Kilpelainen TO.,…,**Tsernikova N**., et al. 2015. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS

Genet. 2016 Jun 29;12(6):e1006166. doi: 10.1371/journal.pgen.1006166. eCollection 2016 Jun.

Nikpay M., Goel A., Won H-H., Hall LM., Willenborg C., Kanoni S., Saleheen D., Kyriakou T., Nelson CP., Hopewell JC.,…, **Pervjakova N**., et al. (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015 Oct;47(10):1121-1130. doi: 10.1038/ng.3396. Epub 2015 Sep 7.

Day FR., Ruth KS., Thompson DJ., Lunetta KL., **Pervjakova N.,** Chasman DI., Stolk L., Finucane HK., Sulem P., Bulik-Sullivan B., et al. (2015). Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. Nat Genet. 2015 Nov;47(11):1294-1303. doi: 10.1038/ng.3412. Epub 2015 Sep 28.

Horikoshi M., Mägi R., van de Bunt M., Surakka I., Sarin AP., Mahajan A., Marullo L., Thorleifsson G., Hägg S., Hottenga JJ.,…,**Pervjakova N**., et al. (2015). Discovery and Fine-Mapping of Glycaemic and Obesity-Related Trait Loci Using High-Density Imputation. PLoS Genet. 2015 Jul 1;11(7):e1005230. doi: 10.1371/journal.pgen.1005230. eCollection 2015 Jul.

Joshi PK., Esko T., Mattsson H., Eklund N., Gandin I., Nutile T., Jackson AU., Schurmann C., Smith A V., Zhang W., et al.,…,**Tsernikova N**., (2015). Directional dominance on stature and cognition in diverse human populations. Nature. 2015 Jul 23;523(7561):459-462. doi: 10.1038/nature14618. Epub 2015 Jul 1.

Surakka I., Horikoshi M., Mägi R., Sarin AP., Mahajan A., Lagou V., Marullo L., Ferreira T., Miraglio B., Timonen S.,…,**Tsernikova N**., et al. (2015). The impact of low-frequency and rare variants on lipid levels. Nat Genet. 2015 Jun;47(6):589-97. doi: 10.1038/ng.3300. Epub 2015 May 11.

Perry JR., Day F., Elks CE., Sulem P., Thompson DJ., Ferreira T., He C., Chasman DI., Esko T., Thorleifsson G.,…,**Tsernikova N**., et al. (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature. 2014 Oct 2;514(7520):92-97. doi: 10.1038/ nature13545. Epub 2014 Jul 23.

Hoggart CJ., Venturini G., Mangino M., Gomez F., Ascari G., Zhao JH., Teumer A., Winkler TW., **Tsernikova N.,** Luan J., et al. (2014). Novel Approach Identifies SNPs in *SLC2A10* and *KCNK9* with Evidence for Parent-of-Origin Effect on Body Mass Index. PLoS Genet. 2014 Jul 31; 10(7):e1004508. doi: 10.1371/journal.pgen.1004508. eCollection 2014 Jul.

Tammiste A., Jiang T., Fischer K., Mägi R., Krjutškov K., Pettai K., Esko T., Li Y., Tansey KE., Carroll LS.,…,**Tsernikova N**., et al. (2013). Whole-exome sequencing identifies a polymorphism in the BMP5 gene associated with SSRI treatment response in major depression. J Psychopharmacol. 2013 Oct;27(10):915-20. doi: 10.1177/0269881113499829. Epub 2013 Aug 7.

Awards and stipends:

2016 fellowship for persons from Central and Eastern European countries, European Human Genetics Conference 2016

2015 fellowship for the 4th Sardinian Summer School Course, “From GWAS to Function”, Polaris Technology Park, Pula, Italy.

2014 Graduate School in Biomedicine and Biotechnology Scholarship for developing academic cooperation including universities or other relevant institutions, Tartu, Estonia

2014 Kristjan Jaak Traveling Scholarship, Ministry of Education and Research of Estonia

2014 Kristjan Jaak Scholarship for Study and research activities of Doctoral students in foreign universities and research institutions, Ministry of Education and Research, Estonia

2013 DoRa T6 Scholarship for Study and research activities of Doctoral students in foreign universities and research institutions, Ministry of Education and Research, Estonia

ELULOOKIRJELDUS

Nimi:	Natalia Pervjakova
Sünniaeg:	27. november 1987
Aadress:	Estonian Genome Center, Institute of Genomics, University of Tartu
E-post:	natalia.pervjakova@ut.ee

Haridus:

2012–Praegu	Doktoriõpe, Molekulaarne biomeditsiin, loodus- ja täppisteaduste valdkond, Tartu Ülikool, Tartu, Eesti
2010–2012	Magistriõpe, Biotehnoloogia ja Biomeditsiin, loodus- ja täppisteaduste valdkond, Tartu Ülikool, Tartu, Eesti
2007–2010	Bakalaureuseõpe, geenitehnoloogia, loodus- ja täppisteaduste valdkond, Tartu Ülikool, Tartu, Eesti

Teenistuskäik:

2017–Praegu	Külalisteadlane, Imperial College of Science, Genomics of Common Disease õppetool, London, UK
2017–2017	Teadusassistent, Imperial College of Science, Genomics of Common Disease õppetool, London, UK
2011–Praegu	Specialist, Eesti Geenivaramu, Genoomika instituut, Tartu Ülikool, Eesti
2012–2015	Teadlane, National Institute for Health and Welfare, Helsinki, Soome
2008–2010	Labori assistent, Biokeemia õppetool, Molekulaar- ja rakubioloogia instituut, Tartu Ülikool, Eesti

Teadusorganisatsiooniline ja- administratiivne tegevus:

2016–Present	Ameerika inimegeneetika ühingu liige
2015–Present	'Introduction to the statistical analysis of genetic association studies' kursuse õppejõud, Imperial College of Science, Technology and Medicine, London, UK
2014–2014	'Omic's practical course' kursuse õppejõud, National Institute for Health and Welfare, Helsinki, Soome

Teaduspublikatsioonid:

Pervjakova N., Prokopenko I. (2017). The TAD-pathway for GWAS signals. Eur J Hum Genet 25.

Pervjakova N., Kukushkina V., Haller T., Kasela S., Joensuu A., Kristiansson K., Annilo T., Perola M., Salomaa V., Jousilahti P., Metspalu A., Mägi R., (2017). Genome-wide analysis of NMR metabolites revealed parent-of-origin effect on triglycerides in medium VLDL in PTPRD gene – in press in Biomarkers in Medicine

Rosticci M*., **Pervjakova N***., Kaakinen M*., Cicero AF., Pfeufer A., Marullo L., Mägi R., Fischer K., Jiang L., D'Addato S., et al., (2017). Common variants in HMGCR affect blood Apolipoprotein B levels – accepted to Biomarkers in Medicine

Nuotio ML., **Pervjakova N**., Joensuu A., Karhunen V., Milani L., Kettunen J., Järvelin MR., Metspalu A., Salomaa V., Kristiansson K., Perola M. (2017). Epigenome-wide association study (EWAS) in metabolic syndrome and its components – under review in Diabetes

Scott RA., Scott LJ., Mägi R., Marullo L., Gaulton KJ., Kaakinen M., **Pervjakova N.,** Pers TH., Johnson AD., Eicher JD., et al. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes. 2017 Nov;66(11):2888-2902. doi: 10.2337/db16-1253. Epub 2017 May 31.

Day FR., Thompson DJ., Helgason H., Chasman DI., Finucane H., Sulem P., Ruth KS., Whalen S., Sarkar AK., Albrecht E.,…, **Pervjakova N**., et al. (2017). Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nat Genet. 2017 Jun;49(6):834-841. doi: 10.1038/ng.3841. Epub 2017 Apr 24.

Justice AE., Winkler TW., Feitosa MF., Graff M., Fisher VA., Young K., Barata L., Deng X., Czajkowski J., Hadley D.,…, **Pervjakova N**., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat Commun. 2017 Apr 26;8:14977. doi: 10.1038/ncomms14977.

Wahl S., Drong A., Lehne B., Loh M., Scott WR., Kunze S., Tsai P-C., Ried JS., Zhang W., Yang Y.,…,**Pervjakova N**., et al. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017 Jan 5;541(7635):81–86. doi: 10.1038/nature20784. Epub 2016 Dec 21.

Okbay A., Beauchamp JP., Fontana MA., Lee JJ., Pers TH., Rietveld CA., Turley P., Chen G-B., Emilsson V., Meddens SFW.,…**Pervjakova N**., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. Nature. 2016 May 26;533(7604):539-42. doi: 10.1038/nature17671. Epub 2016 May 11.

Pervjakova N, Kasela S, Morris AP, Kals M, Metspalu A, Lindgren CM, Salumets A, Mägi R. (2016). Imprinted genes and imprinting control regions show predominant intermediate methylation in adult somatic tissues. Epigenomics. 2016 Jun;8(6):789-99. doi: 10.2217/epi.16.8. Epub 2016 Mar 23

Winkler TW., Justice AE., Graff M., Barata L., Feitosa MF., Chu S., Czajkowski J., Esko T., Fall T., Kilpelainen TO.,…,**Tsernikova N**., et al. 2015. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet. 2016 Jun 29;12(6):e1006166. doi: 10.1371/journal.pgen.1006166. eCollection 2016 Jun.

Nikpay M., Goel A., Won H-H., Hall LM., Willenborg C., Kanoni S., Saleheen D., Kyriakou T., Nelson CP., Hopewell JC.,…, **Pervjakova N**., et al. (2015).

A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015 Oct;47(10):1121-1130. doi: 10.1038/ng.3396. Epub 2015 Sep 7.
Day FR., Ruth KS., Thompson DJ., Lunetta KL., **Pervjakova N.,** Chasman DI., Stolk L., Finucane HK., Sulem P., Bulik-Sullivan B., et al. (2015). Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. Nat Genet. 2015 Nov;47(11):1294-1303. doi: 10.1038/ng.3412. Epub 2015 Sep 28.
Horikoshi M., Mägi R., van de Bunt M., Surakka I., Sarin AP., Mahajan A., Marullo L., Thorleifsson G., Hägg S., Hottenga JJ.,…,**Pervjakova N**., et al. (2015). Discovery and Fine-Mapping of Glycaemic and Obesity-Related Trait Loci Using High-Density Imputation. PLoS Genet. 2015 Jul 1; 11(7):e1005230. doi: 10.1371/journal.pgen.1005230. eCollection 2015 Jul.
Joshi PK., Esko T., Mattsson H., Eklund N., Gandin I., Nutile T., Jackson AU., Schurmann C., Smith A V., Zhang W., et al.,…,**Tsernikova N**., (2015). Directional dominance on stature and cognition in diverse human populations. Nature. 2015 Jul 23;523(7561):459-462. doi: 10.1038/nature14618. Epub 2015 Jul 1.
Surakka I., Horikoshi M., Mägi R., Sarin AP., Mahajan A., Lagou V., Marullo L., Ferreira T., Miraglio B., Timonen S.,…,**Tsernikova N**., et al. (2015). The impact of low-frequency and rare variants on lipid levels. Nat Genet. 2015 Jun;47(6):589-97. doi: 10.1038/ng.3300. Epub 2015 May 11.
Perry JR., Day F., Elks CE., Sulem P., Thompson DJ., Ferreira T., He C., Chasman DI., Esko T., Thorleifsson G.,…,**Tsernikova N**., et al. (2014). Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature. 2014 Oct 2;514(7520):92-97. doi: 10.1038/nature13545. Epub 2014 Jul 23.
Hoggart CJ., Venturini G., Mangino M., Gomez F., Ascari G., Zhao JH., Teumer A., Winkler TW., **Tsernikova N.,** Luan J., et al. (2014). Novel Approach Identifies SNPs in *SLC2A10* and *KCNK9* with Evidence for Parent-of-Origin Effect on Body Mass Index. PLoS Genet. 2014 Jul 31;10(7):e1004508. doi: 10.1371/journal.pgen.1004508. eCollection 2014 Jul.
Tammiste A., Jiang T., Fischer K., Mägi R., Krjutškov K., Pettai K., Esko T., Li Y., Tansey KE., Carroll LS.,…,**Tsernikova N**., et al. (2013). Whole-exome sequencing identifies a polymorphism in the BMP5 gene associated with SSRI treatment response in major depression. J Psychopharmacol. 2013 Oct;27(10):915-20. doi: 10.1177/0269881113499829. Epub 2013 Aug 7.

Stipendiumid:

2016 Stipendium inimestele kesk- ja ida Euroopast, Euroopa Inimgeneetika Ühing, 2016
2015 4. Sardiinia Suvekooli stipendium. 'From GWAS to Function', Polaris Tehnoloogia Park, Pula, Itaalia.
2014 Biomeditsiini ja Biotehnoloogia doktorikooli stipendium, Tartu, Eesti

2014 Kristjan Jaagu välissõidu stipendium, teadusministeerium, Eesti
2014 Kristjan Jaagu osalise õppe stipendium, teadusministeerium, Eesti
2013 DoRa T6 Rahvusvaheliste koostöövõrgustike arendamine läbi Eesti doktorantide õpirändestipendium, teadusministeerium, Eesti

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 serotonin. 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 intra-cellular 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 zoo-plankton 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 boreo-nemoral 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, 141 p.
256. **Liina Nagirnaja**. Global and fine-scale genetic determinants of recurrent pregnancy loss. Tartu, 2014, 211 p.
257. **Triin Triisberg**. Factors influencing the re-vegetation of abandoned extracted peatlands in Estonia. Tartu, 2014, 133 p.
258. **Villu Soon**. A phylogenetic revision of the *Chrysis ignita* species group (Hymenoptera: Chrysididae) with emphasis on the northern European fauna. Tartu, 2014, 211 p.
259. **Andrei Nikonov**. RNA-Dependent RNA Polymerase Activity as a Basis for the Detection of Positive-Strand RNA Viruses by Vertebrate Host Cells. Tartu, 2014, 207 p.
260. **Eele Õunapuu-Pikas**. Spatio-temporal variability of leaf hydraulic conductance in woody plants: ecophysiological consequences. Tartu, 2014, 135 p.
261. **Marju Männiste**. Physiological ecology of greenfinches: information content of feathers in relation to immune function and behavior. Tartu, 2014, 121 p.
262. **Katre Kets**. Effects of elevated concentrations of CO_2 and O_3 on leaf photosynthetic parameters in *Populus tremuloides*: diurnal, seasonal and interannual patterns. Tartu, 2014, 115 p.

263. **Külli Lokko**. Seasonal and spatial variability of zoopsammon communities in relation to environmental parameters. Tartu, 2014, 129 p.
264. **Olga Žilina**. Chromosomal microarray analysis as diagnostic tool: Estonian experience. Tartu, 2014, 152 p.
265. **Kertu Lõhmus**. Colonisation ecology of forest-dwelling vascular plants and the conservation value of rural manor parks. Tartu, 2014, 111 p.
266. **Anu Aun**. Mitochondria as integral modulators of cellular signaling. Tartu, 2014, 167 p.
267. **Chandana Basu Mallick**. Genetics of adaptive traits and gender-specific demographic processes in South Asian populations. Tartu, 2014, 160 p.
268. **Riin Tamme**. The relationship between small-scale environmental heterogeneity and plant species diversity. Tartu, 2014, 130 p.
269. **Liina Remm**. Impacts of forest drainage on biodiversity and habitat quality: implications for sustainable management and conservation. Tartu, 2015, 126 p.
270. **Tiina Talve**. Genetic diversity and taxonomy within the genus *Rhinanthus*. Tartu, 2015, 106 p.
271. **Mehis Rohtla**. Otolith sclerochronological studies on migrations, spawning habitat preferences and age of freshwater fishes inhabiting the Baltic Sea. Tartu, 2015, 137 p.
272. **Alexey Reshchikov**. The world fauna of the genus *Lathrolestes* (Hymenoptera, Ichneumonidae). Tartu, 2015, 247 p.
273. **Martin Pook**. Studies on artificial and extracellular matrix protein-rich surfaces as regulators of cell growth and differentiation. Tartu, 2015, 142 p.
274. **Mai Kukumägi**. Factors affecting soil respiration and its components in silver birch and Norway spruce stands. Tartu, 2015, 155 p.
275. **Helen Karu**. Development of ecosystems under human activity in the North-East Estonian industrial region: forests on post-mining sites and bogs. Tartu, 2015, 152 p.
276. **Hedi Peterson**. Exploiting high-throughput data for establishing relationships between genes. Tartu, 2015, 186 p.
277. **Priit Adler**. Analysis and visualisation of large scale microarray data, Tartu, 2015, 126 p.
278. **Aigar Niglas**. Effects of environmental factors on gas exchange in deciduous trees: focus on photosynthetic water-use efficiency. Tartu, 2015, 152 p.
279. **Silja Laht**. Classification and identification of conopeptides using profile hidden Markov models and position-specific scoring matrices. Tartu, 2015, 100 p.
280. **Martin Kesler**. Biological characteristics and restoration of Atlantic salmon *Salmo salar* populations in the Rivers of Northern Estonia. Tartu, 2015, 97 p.
281. **Pratyush Kumar Das**. Biochemical perspective on alphaviral nonstructural protein 2: a tale from multiple domains to enzymatic profiling. Tartu, 2015, 205 p

282. **Priit Palta**. Computational methods for DNA copy number detection. Tartu, 2015, 130 p.
283. **Julia Sidorenko**. Combating DNA damage and maintenance of genome integrity in pseudomonads. Tartu, 2015, 174 p.
284. **Anastasiia Kovtun-Kante**. Charophytes of Estonian inland and coastal waters: distribution and environmental preferences. Tartu, 2015, 97 p.
285. **Ly Lindman**. The ecology of protected butterfly species in Estonia. Tartu, 2015, 171 p.
286. **Jaanis Lodjak**. Association of Insulin-like Growth Factor I and Corticosterone with Nestling Growth and Fledging Success in Wild Passerines. Tartu, 2016, 113 p.
287. **Ann Kraut**. Conservation of Wood-Inhabiting Biodiversity – Semi-Natural Forests as an Opportunity. Tartu, 2016, 141 p.
288. **Tiit Örd**. Functions and regulation of the mammalian pseudokinase TRIB3. Tartu, 2016, 182. p.
289. **Kairi Käiro**. Biological Quality According to Macroinvertebrates in Streams of Estonia (Baltic Ecoregion of Europe): Effects of Human-induced Hydromorphological Changes. Tartu, 2016, 126 p.
290. **Leidi Laurimaa**. *Echinococcus multilocularis* and other zoonotic parasites in Estonian canids. Tartu, 2016, 144 p.
291. **Helerin Margus**. Characterization of cell-penetrating peptide/nucleic acid nanocomplexes and their cell-entry mechanisms. Tartu, 2016, 173 p.
292. **Kadri Runnel**. Fungal targets and tools for forest conservation. Tartu, 2016, 157 p.
293. **Urmo Võsa**. MicroRNAs in disease and health: aberrant regulation in lung cancer and association with genomic variation. Tartu, 2016, 163 p.
294. **Kristina Mäemets-Allas**. Studies on cell growth promoting AKT signaling pathway – a promising anti-cancer drug target. Tartu, 2016, 146 p.
295. **Janeli Viil**. Studies on cellular and molecular mechanisms that drive normal and regenerative processes in the liver and pathological processes in Dupuytren's contracture. Tartu, 2016, 175 p.
296. **Ene Kook**. Genetic diversity and evolution of *Pulmonaria angustifolia* L. and *Myosotis laxa sensu lato* (Boraginaceae). Tartu, 2016, 106 p.
297. **Kadri Peil**. RNA polymerase II-dependent transcription elongation in *Saccharomyces cerevisiae*. Tartu, 2016, 113 p.
298. **Katrin Ruisu**. The role of RIC8A in mouse development and its function in cell-matrix adhesion and actin cytoskeletal organisation. Tartu, 2016, 129 p.
299. **Janely Pae**. Translocation of cell-penetrating peptides across biological membranes and interactions with plasma membrane constituents. Tartu, 2016, 126 p.
300. **Argo Ronk**. Plant diversity patterns across Europe: observed and dark diversity. Tartu, 2016, 153 p.

301. **Kristiina Mark**. Diversification and species delimitation of lichenized fungi in selected groups of the family Parmeliaceae (Ascomycota). Tartu, 2016, 181 p.
302. **Jaak-Albert Metsoja**. Vegetation dynamics in floodplain meadows: influence of mowing and sediment application. Tartu, 2016, 140 p.
303. **Hedvig Tamman**. The GraTA toxin-antitoxin system of *Pseudomonas putida*: regulation and role in stress tolerance. Tartu, 2016, 154 p.
304. **Kadri Pärtel**. Application of ultrastructural and molecular data in the taxonomy of helotialean fungi. Tartu, 2016, 183 p.
305. **Maris Hindrikson**. Grey wolf (*Canis lupus*) populations in Estonia and Europe: genetic diversity, population structure and -processes, and hybridization between wolves and dogs. Tartu, 2016, 121 p.
306. **Polina Degtjarenko**. Impacts of alkaline dust pollution on biodiversity of plants and lichens: from communities to genetic diversity. Tartu, 2016, 126 p.
307. **Liina Pajusalu**. The effect of CO_2 enrichment on net photosynthesis of macrophytes in a brackish water environment. Tartu, 2016, 126 p.
308. **Stoyan Tankov**. Random walks in the stringent response. Tartu, 2016, 94 p.
309. **Liis Leitsalu**. Communicating genomic research results to population-based biobank participants. Tartu, 2016, 158 p.
310. **Richard Meitern**. Redox physiology of wild birds: validation and application of techniques for detecting oxidative stress. Tartu, 2016, 134 p.
311. **Kaie Lokk**. Comparative genome-wide DNA methylation studies of healthy human tissues and non-small cell lung cancer tissue. Tartu, 2016, 127 p.
312. **Mihhail Kurašin**. Processivity of cellulases and chitinases. Tartu, 2017, 132 p.
313. **Carmen Tali**. Scavenger receptors as a target for nucleic acid delivery with peptide vectors. Tartu, 2017, 155 p.
314. **Katarina Oganjan**. Distribution, feeding and habitat of benthic suspension feeders in a shallow coastal sea. Tartu, 2017, 132 p.
315. **Taavi Paal**. Immigration limitation of forest plants into wooded landscape corridors. Tartu, 2017, 145 p.
316. **Kadri Õunap**. The Williams-Beuren syndrome chromosome region protein WBSCR22 is a ribosome biogenesis factor. Tartu, 2017, 135 p.
317. **Riin Tamm**. In-depth analysis of factors affecting variability in thiopurine methyltransferase activity. Tartu, 2017, 170 p.
318. **Keiu Kask**. The role of RIC8A in the development and regulation of mouse nervous system. Tartu, 2017, 184 p.
319. **Tiia Möller**. Mapping and modelling of the spatial distribution of benthic macrovegetation in the NE Baltic Sea with a special focus on the eelgrass *Zostera marina* Linnaeus, 1753. Tartu, 2017, 162 p.
320. **Silva Kasela**. Genetic regulation of gene expression: detection of tissue- and cell type-specific effects. Tartu, 2017, 150 p.

321. **Karmen Süld**. Food habits, parasites and space use of the raccoon dog *Nyctereutes procyonoides*: the role of an alien species as a predator and vector of zoonotic diseases in Estonia. Tartu, 2017, p.
322. **Ragne Oja**. Consequences of supplementary feeding of wild boar – concern for ground-nesting birds and endoparasite infection. Tartu, 2017, 141 p.
323. **Riin Kont**. The acquisition of cellulose chain by a processive cellobiohydrolase. Tartu, 2017, 117 p.
324. **Liis Kasari**. Plant diversity of semi-natural grasslands: drivers, current status and conservation challenges. Tartu, 2017, 141 p.
325. **Sirgi Saar**. Belowground interactions: the roles of plant genetic relatedness, root exudation and soil legacies. Tartu, 2017, 113 p.
326. **Sten Anslan**. Molecular identification of Collembola and their fungal associates. Tartu, 2017, 125 p.
327. **Imre Taal**. Causes of variation in littoral fish communities of the Eastern Baltic Sea: from community structure to individual life histories. Tartu, 2017, 118 p.
328. **Jürgen Jalak**. Dissecting the Mechanism of Enzymatic Degradation of Cellulose Using Low Molecular Weight Model Substrates. Tartu, 2017, 137 p.
329. **Kairi Kiik**. Reproduction and behaviour of the endangered European mink (*Mustela lutreola*) in captivity. Tartu, 2018, 112 p.
330. **Ivan Kuprijanov**. Habitat use and trophic interactions of native and invasive predatory macroinvertebrates in the northern Baltic Sea. Tartu, 2018, 117 p.
331. **Hendrik Meister**. Evolutionary ecology of insect growth: from geographic patterns to biochemical trade-offs. Tartu, 2018, 147 p.
332. **Ilja Gaidutšik.** Irc3 is a mitochondrial branch migration enzyme in *Saccharomyces cerevisiae.* Tartu, 2018, 161 p.
333. **Lena Neuenkamp**. The dynamics of plant and arbuscular mycorrhizal fungal communities in grasslands under changing land use. Tartu, 2018, 241 p.
334. **Laura Kasak.** Genome structural variation modulating the placenta and pregnancy maintenance. Tartu, 2018, 181 p.
335. **Kersti Riibak.** Importance of dispersal limitation in determining dark diversity of plants across spatial scales. Tartu, 2018, 133 p.
336. **Liina Saar.** Dynamics of grassland plant diversity in changing landscapes. Tartu, 2018, 206 p.
337. **Hanna Ainelo.** Fis regulates *Pseudomonas putida* biofilm formation by controlling the expression of *lapA*. Tartu, 2018, 143 p.