

TARTU ÜLIKOOL
LOODUS- JA TEHNOLOOGIATEADUSKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOTEHNOLOOGIA ÕPPETOOL

30 EAP

Diana Nõmmemees

Vanemaspetsiifiline geeniekspressioon inimese platsentas

Magistritöö

Juhendaja prof. Maris Laan

TARTU 2016

Vanemaspetsiifiline geeniekspressioon inimese platsentas

Genoomne vermitus on epigeneetiline nähtus, mille puhul osadel geenidel ekspresseerub ainult ühe vanema alleel ning teise vanema alleel on vaigistatud. Grupp vermitud gene ekspresseerub platsentas, kus nad reguleerivad platsenta arengut ja loote kasvu. Vermituse häired võivad viia rasedustüsistusteni ning mitmeid gene on seostatud üsasisesse kasvupeetusega. Käesoleva töö eesmärk oli kirjanduses varem avaldatud vermitud geenide vanemaspetsiifilise avaldumismustri tuvastamine inimese platsentakoes normaalse raseduse kolme trimestri ja rasedustüsistuste korral. Minu magistritöö lähtematerjalideks oli 55 inimese platsenta koeproovide transkriptomid ning uuritavate perede ülegenoomsed genotüübid. Kasutades rangeid kriteeriume ja mitut filtreerimise etappi, tuvastasin platsentas 11 vermitud geeni. Käesolevas magistritöös tuvastatud 10 vermitud geeni vastasid eelnevalt kirjanduses publitseeritule, välja arvatud *RTL1*, mis näitas vastupidist alleelset vanemlikku päritolu.

Imprinted genes are exclusively expressed from either the paternally or maternally derived allele and the other allele is silenced via epigenetic mechanisms. A number of imprinted genes express in placenta, where they regulate placental and fetal growth. The aims of this study were: 1) To determine imprinted gene transcript allele specific expression in placenta based on previously published articles and Geneimprint database 2) To identify imprinted gene expression in three pregnancy trimesters and in pregnancy complications. In this study we identified 11 imprinted genes. 10 genes validated previously reported parent specific allelic expression, but contrary to previously described paternal expression *RTL1* exhibited maternal expression in our study. We also detected 41 genes with biallelic expression and 3 with random monoallelic expression. 14 genes showed characteristics similar to random monoallelic expression and biallelic expression.

Märksõnad: vermitus, platsenta, RNA-seq, monoalleelne ekspressioon, RMAE, rasedustüsistus

B220 Geneetika, tsütogeneetika

B726 Kliiniline bioloogia

B790 Kliiniline geneetika

B110 Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika

Sisukord

Vanemaspetsiifiline geeniekspressioon inimese platsentas.....	2
Sisukord.....	3
Kasutatud lühendid.....	5
Sissejuhatus.....	7
1. KIRJANDUSE ÜLEVAADE.....	8
1.1. Geenide monoalleelne ekspressioon genoomis.....	8
1.1.1. Genoomne vermimine.....	8
1.1.2. Dünaamiline ja fikseeritud RMAE.....	11
1.1.3. Geenide vermimise hindamine.....	12
1.1.4. RNA sekveneerimine vermimise tuvastamiseks.....	13
1.2. Inimese vermitud geenid.....	13
1.2.1. Patoloogiad.....	15
1.2.2. Rasedustüsistused ja platsentas vermitud geenid.....	16
2. EKSPERIMENTAALOSA.....	18
2.1. Töö eesmärgid.....	18
2.2. Materjal ja meetodika.....	20
2.2.1. Valim.....	20
2.2.2. Genoomse DNA eraldamine.....	23
2.2.3. Platsentast RNA eraldamine ja sekveneerimine.....	25
2.2.4. Perede ülegenoomsete SNP-de genotüpiseerimine.....	26
2.2.5. Potentsiaalselt vermitud geenide valik uuringusse.....	26
2.2.6. Mõisted.....	26
2.2.7. Käesolevas magistritöös analüüsitava andmestiku filtrid.....	27
2.2.8. Andmeanalüüs vermitud geenide oleku määramiseks.....	29
2.3. Tulemused.....	30
2.3.1. Ainult pooled kirjanduses raporteeritud vermitud geenidest on avaldunud platsentas.....	30
2.3.2. Vermitud geenid.....	34
2.3.3. Geenid, mis ei ole platsentas vermitud.....	41
2.4. Arutelu.....	43
2.4.1. Käesoleva magistritöö kattuvus kirjanduses raporteeritud andmetega.....	43
2.4.2. RNA sekveneerimise tugevad ja nõrgad küljed.....	45
KOKKUVÕTE.....	46
Monoallelic parent-of-origin gene expression in human placenta.....	47

Diana Nõmmemees	47
SUMMARY	47
Kasutatud kirjanduse loetelu	48
Kasutatud veebiaadressid	55
LISAD	56
Lisa 1.....	56
Lisa 2.....	58
Lisa 3.....	65
Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	66

Kasutatud lühendid

aRMAE – autosomatic RMAE, autosomaalne RMAE

BAM – SAM (järjestuse joondamise info tekstifailina) faili binaarne versioon

Bp – base pair, aluspaar

BWS – Beckwith-Wiedemann sündroom

Chr – kromosoom

CTCF – CCCTC-seonduv faktor

EDTA – etüleendiamiintetraatseethape

FISH – fluorescence in situ hybridization, fluorestsents in situ hübridisatsioon

FPKM – Fragments Per Kilobase per Million mapped fragments, fragmente kb kohta miljoni joondatud fragmendi kohta

fRMAE – fixed RMAE, fikseerunud RMAE

GD – gestatsioonidiabeet

ICR – (Imprinting Control Region, ICR).

IUGR – intrauterine growth retardation, üsisisene kasvupeetus

KMI – kehamassiindeks

LGA – large for gestational age, makrosoomia ehk gestatsiooniaja kohta suur laps

lncRNA – long non-coding RNA, pikk mittekodeeriv RNA

MAF – minor allele frequency, minoorne alleeli sagedus

miRNA – mikroRNA

NGEM – Normaliseeritud geeniekspressiooni mediaan

PE – preeklampsia

RCF – Relative Centrifugal Force, relatiivne tsentrifugaaljõud

REPROMETA – REPROgrammed fetal and/or maternal METAbolism –
REPROgrammeeritud loote ja/või ema METAbolism

RMAE – random monoallelic expression, juhuslik monoalleelne ekspressioon

RNA-seq – RNA sügavsekveneerimine

RT-PCR – Reverse Transcriptase PCR, pöördtranskriptaas PCR

SD – standard deviation, standardhälve

SDS – anioonne surfaktant

SGA – small for gestational age, hüpotoofia ehk gestatsiooniaja kohta väike laps

SNP – single nucleotide polymorphism, ühenukleotiidne polümorfism

SRS – Silver-Russell sündroom

TBE – Tris-boorhape-EDTA puhvriühendus

TE – puhvriühendus, mis koosneb Tris-ist ja EDTA-st

TNDM1 – Transient neonatal diabetes mellitus type 1, mööduv neonataalne tüüp 1 diabeet

XCI – X-Chromosome Inactivation, X-kromosoomi inaktivatsioon

Sissejuhatus

Diploidsed organismid pärivad mõlemalt vanemalt ühe geenikoopia ning tavaliselt toimub ekspressioon mõlemal alleelil. Kõige tuntum erand on naissoost imetajate X-kromosoomi vaigistamine, kus ühelt vanemalt pärit kromosoom vaigistatakse täielikult, kuid tavaliselt on mittepatoloogiline monoalleelne ekspressioon ühe geeni piires. Genoomne vermitus on epigeneetiline nähtus, mille puhul osadel geenidel ekspresseerub ainult ühe vanema alleel ning teise vanema alleel on vaigistatud.

Vermitud geenidega on seostatud mitmeid haigusi, nagu Prader-Willi ja Angelmani sündroom. Grupp vermitud gene ekspresseerub platsentas, kus nad reguleerivad platsenta arengut ja loote kasvu. Üsasisest kasvupeetust on seostatud vermituse häirega ning seetõttu tuleks vermitud gene uurida ka teiste raseduspatoloogiate puhul põhjalikumalt.

Käesoleva magistritöös fookuses oli kirjanduses varem avaldatud vermitud geenide vanemaspetsiifilise avaldumismustri tuvastamine inimese platsentakoes normaalse raseduse kolme trimestri ja rasedustüsistuste korral. Minu magistritöö lähtematerjalideks oli 55 inimese platsenta koeproovide transkriptoomid ning uuritavate perede ülegenoomsed genotüübid.

1. KIRJANDUSE ÜLEVAADE

1.1. Geenide monoalleelne ekspressioon genoomis

Enamus imetajate autosomaalseid gene ekspresseerub kodominantselt mõlema vanema kromosoomilt. Kuid mõnedes lookustes on ühe vanema alleeli ekspressioon alla surutud epigeneetiliste mehhanismidega. Kõige tuntum näide on naissoost imetajate X-kromosoomi vaigistamine, kus ühelt vanemalt pärit kromosoom vaigistatakse täielikult.

1.1.1. Genoomne verimine

Vermitud geenid ekspresseeruvad enamasti somaatilistes rakkudes eksklusiivselt kas emalt või isalt päritud alleele (joonis 1. A). Vermituse olekut reguleerivad naise või mehe sugurakkudes toimuvad epigeneetilised nähtused. (Reinius ja Sandberg, 2015)

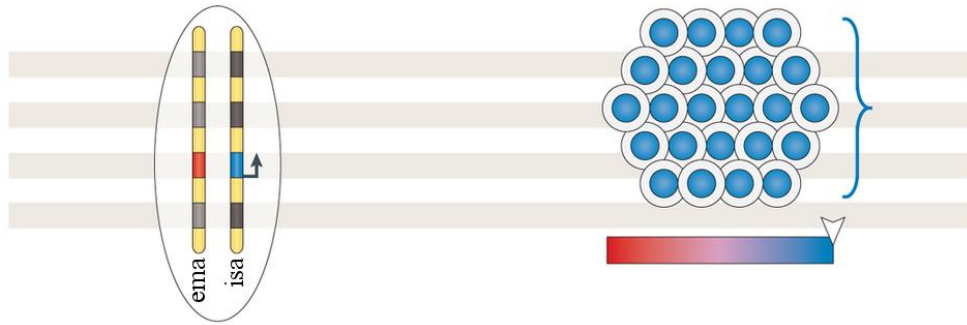
Uuringud hiirtel on näidanud, et >80% teadaolevatest vermitud geenidest asuvad klastrites. Enamasti on hiire ja inimese ortoloogsed klastrid sarnase paigutusega ning kõik klastrid sisaldavad nii emapoolse kui isapoolse alleelse ekspressiooniga gene.

Vanemaspetsiifilist alleelset ekspressiooni reguleerib klastris mitme geeni suhtes *cis*-positsioonis olev vermituse kontrollregioon (*Imprinting Control Region*, ICR). ICR regioon on ühe vanema alleelspetsiifiline DNA metülatsioon ja kromatiini modifikatsioonid. Metülatsioon ICR-des on omandatud kas ema või isa sugurakkudes. Sugurakkude metülatsioon on robustne ja resistentne ulatuslikule genoomi metülatsiooni ümberprogrammeerimisele, mis toimub pärast viljastumist, kuid see eemaldatakse uute sugurakkude arengu käigus. (ülevaateartikkel Peters, 2014)

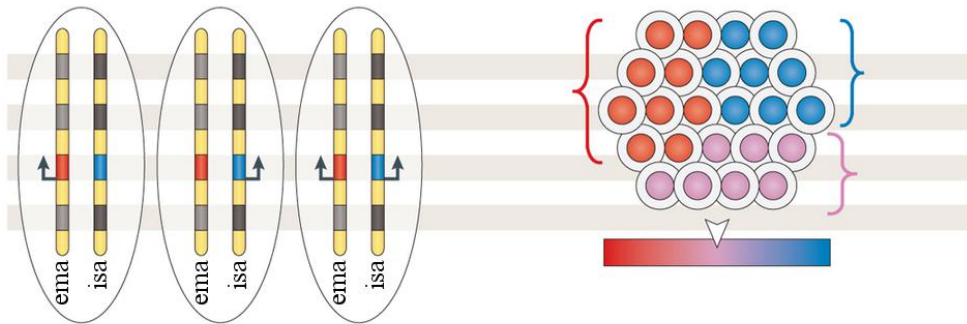
Enamus ICR-e omandavad metülatsiooni naise sugurakkudes oogeneesi ajal ja nendes ICR-des on tavaliselt pikkade mittekodeerivate RNA-de (lncRNA) promootorid, mida transkribeeritakse vastasahelalt kodeerivate geenide klastris (joonis 2. A). ICR-id, mis omandavad metüleerituse mehe sugurakkudes, paiknevad intergeensetes regioonides (joonis x. B ja C). ICR on aktiivne mittemetüleeritult ning inaktiivne metülatsiooni korral. ICR-id reguleerivad vermitust kahe mehhanismiga:

- 1) lncRNA-d ICR-dest vaigistavad *cis*-positsioonis olevad vermitud geenid (joonis 2. A)
- 2) Aktiivne ICR moodustab isolaatori, seondudes CCCTC-seonduva faktorile (CTCF) ning blokeerides 3'suunas olevate geenide promootorite *enhancerid* ja vaigistades need geenid. (ülevaateartikkel Peters, 2014) (joonis 2.C)

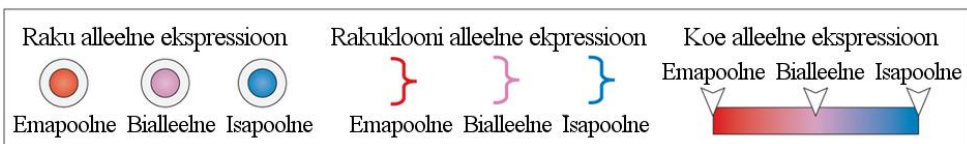
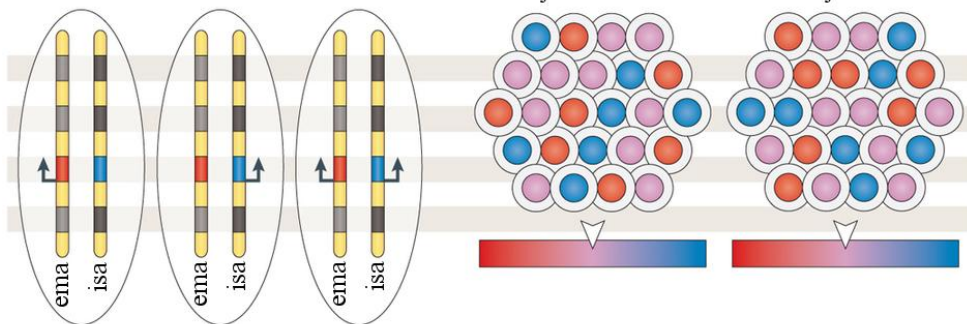
A Genoomne vermimine



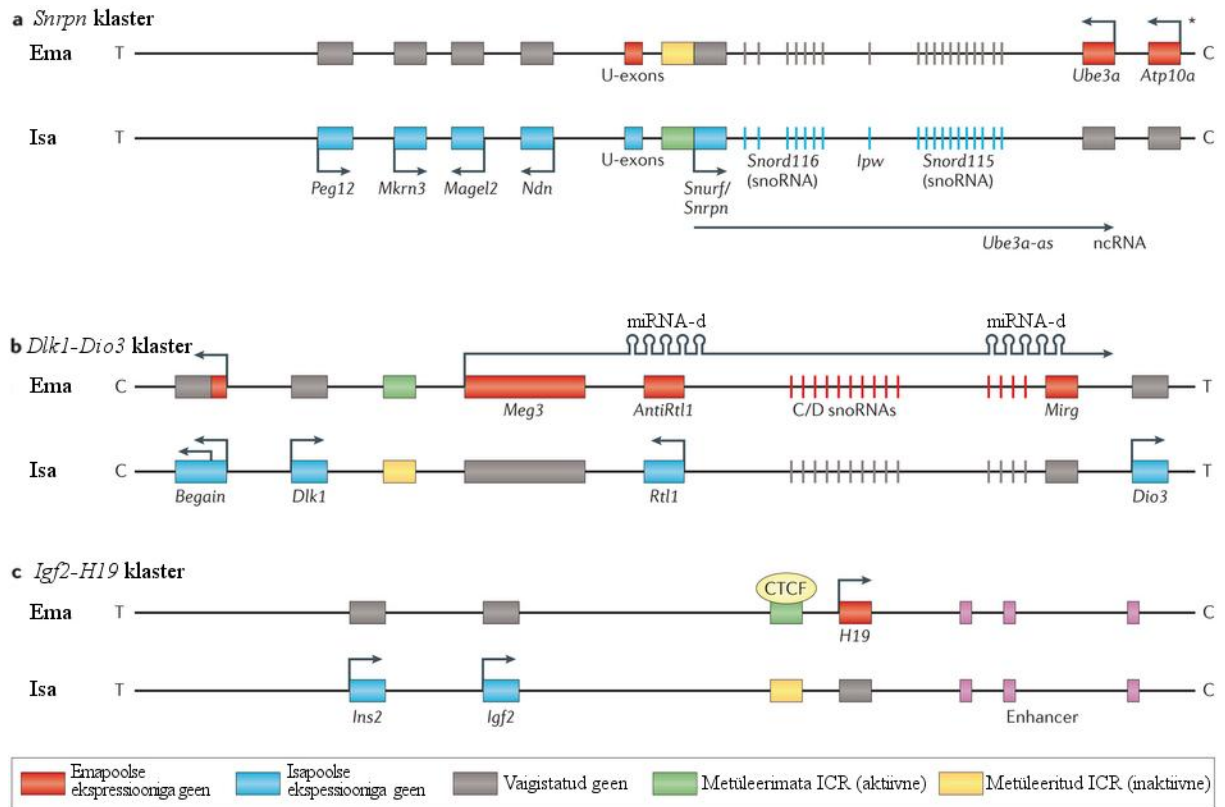
B Fikseeritud aRMAE



C Dünaamiline aRMAE



Joonis 1. Monoalleelse geeniekspressiooni tüübid. Vasakul pool on võimalikud ühe raku alleelse ekspressiooni olekud, kus transkriptsioon (märgitud musta noolega) on märgitud vastavalt ema või isa geenikoopialt. Paremal pool on välja toodud alleelne ekspressioon rakupopulatsioonis (ringikesed, erinevate värvidega tähistatud avalduv alleel). Värvilised loogelised sulud tähistavad klonaalsetid rakke, kus fikseerunud monoalleelne või bialleelne ekspressioon on kandunud edasi raku jagunemisel. Iga rakupopulatsiooni all on valge nool ja gradiendi ribal näidatud alleelne ekspressioon kui ta oleks detekteeritud terve rakupopulatsiooni peale. Eeldatavad vermituse mustrid on toodud: A) genoomse vermituse, B) fikseeritud aRMAE ja C) dünaamilise aRMAE kohta. Joonis kohandatud Reinius ja Sandberg, 2015 järgi.



Joonis 2. Hiire vermitud geenide klastrite näited. A) *Snrpn* klastris on emapoolse alleelse ekspressiooniga *Ube3a* geen ja viis isapoolse alleelse ekspressiooniga geeni, samuti mitu mittekodeerivat RNA-d. B) *Dlk1 –Dio3* klastris on neli isapoolse alleelse ekspressiooniga geeni ja mitu emapoolse alleelse ekspressiooniga mittekodeerivat RNA-d. ICR on intergeenne ja aktiivne emapoolsel alleelil. C) *Igf2 –H19* klastris on isapoolse alleelse ekspressiooniga *Igf2* ja *Ins2* ning emapoolse alleelse ekspressiooniga *H19*. ICR on intergeenne. Aktiivne emapoolne ICR seondub CCCTC-seonduvale faktorile (CTCF), moodustades isolaatori, mis vaigistab *Igf2* blokeerides *Igf2 enhancer*itele ligipääsu. Nooltega on tähistatud transkribeeritavad alleelid ja geenid. Tärniga on tähistatud geen, mille vermitus inimestes ja hiirtes on vastuoluline. C- tsentromeer, Ipw- vermitud regiooni Prader-Willi sündroomi regioonis, miR- miRNA, *Mirg*- miRNA-t kodeeriv geen, T- telomeer. Kohandatud Peters, 2014 põhjal.

1.1.2. Dünaamiline ja fikseeritud RMAE

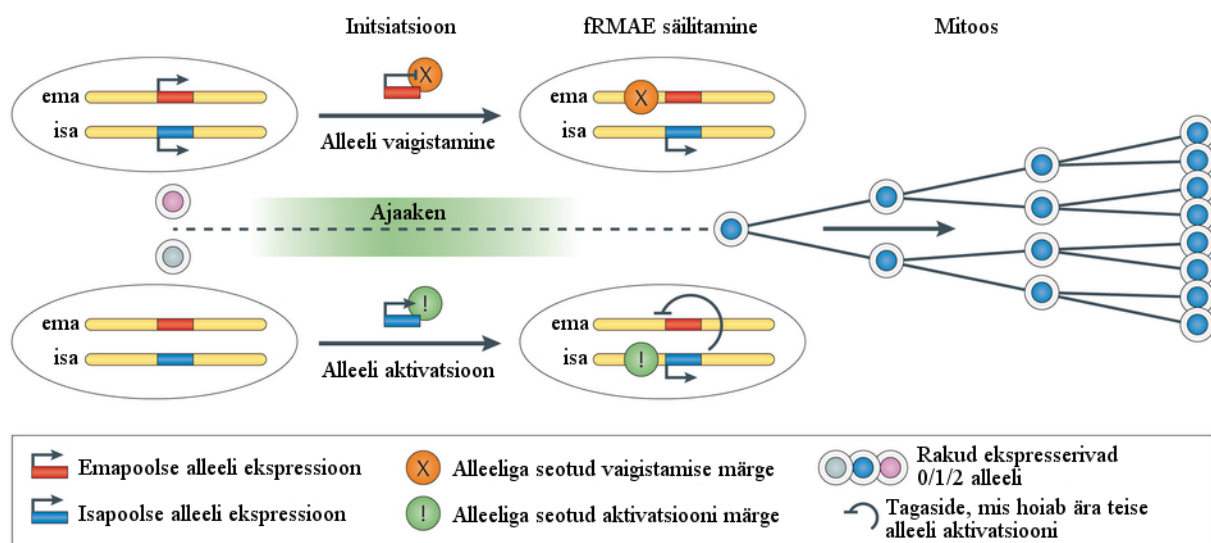
aRMAE (*autosomal Random MonoAllelic Expression*) on juhuslik monoalleelne ekspressioon, mis on defineeritud kui ühe alleeli ekspressioon, kus juhuslik alleelivalik toimub somaatilisel, nii et erinevad organismi rakud ekspresseerivad erinevaid allelele. aRMAE uurimine on paras väljakutse, kuna erinevad meetodid ja tingimused toodavad konfliktseid tulemusi ning ebatäiuslik RNA alleelne detektsioon ühes rakus suurendab monoalleelse ekspressiooni vaatlusi. (Reinius ja Sandberg, 2015)

RMAE jaguneb omakorda:

2. fikseeritud RMAE (fRMAE)- alleelispetsiifiline ekspressioon pärandub edasi tütararakule pärast mitoosi (joonis 1. B).
3. dünaamiline RMAE- ajutine alleelispetsiifiline ekspressioon, mis ei kandu rakujagunemisel edasi (joonis 1. C)

fRMAE tekib bialleelselt ekspresseeritud geenilt kas ühe alleeli vaigistamise või eelnevalt vaigistatud geeni ühe alleeli aktivatsiooniga (joonis 3). Alleeli aktivatsiooni puhul võib olemas olla ka tagasisideme mehhanism, mis hoiab ära teise alleeli aktivatsiooni. Teise võimalusena eksisteerib selline ajaaken, kus madala tõenäosusega initsiatsioon võib saavutada kõrge sagedusega rakke monoalleelse ekspressiooniga.

Kuna fRMAE alleelne valik on mitoosi teel edasi kanduv, peavad reguloorsed modifikatsioonid olema seotud *cis* positsioonis olevate alleelidega. (Reinius and Sandberg, 2015)



Joonis 3. Fikseeritud RMAE teke. Joonis kohandatud Reinius ja Sandberg, 2015 järgi.

1.1.3. Geenide verimise hindamine

Alleelset geeniekspressiooni võib määrata 8 erineva meetodiga (tabel 1). Ühe geeni põhistest lähenemistest kasutatakse kõige enam RT-PCR-i. Ühe geeni põhisel lähenemisel on puuduseks rakukultuurist tingitud mittefüsioloogilised DNA aberratsioonid. Samuti võib ebatäiuslik RNA detektsioon suurendada monoalleelse ekspressiooni hinnangut. Ülegenoomsete lähenemiste puhul kasutatakse enim suure statistilise võimsusega SNP-tundlike mikrokiipe ning aina enam RNA sekveneerimist (RNA-seq). Kahjuks ei saa nende kahe meetodiga tuvastada dünaamilist aRMAE-d. Ühe raku RNA sekveneerimise puhul tuleb hinnata alleelide tehnilistel põhjustel välja langemise määra.

Tabel 1. Alleelspetsiifilise geeniekspressiooni hindamise meetodid. Kohandatud Reinius ja Sandberg, 2015 põhjal.

Meetod	Resolutsioon	Kasutamine
Kandidaatgeeni-põhine lähenemine		
RT-PCR	Rakk või klonaalne populatsioon	Kasutatakse koos alleelspetsiifiliste praimeritega, SNP- tundlike hübriidisatsiooniproovide, alleelspetsiifiliste restriktioonisaitide või cDNA produktide sekveneerimisega
Äsjasünteesitud RNA FISH	Rakk	<i>in situ</i> äsjasünteesitud RNA visualiseerimine.
Ühe raku RNA FISH	Rakk	SNP- tundlik meetod <i>in situ</i> RNA lokaliseerimiseks
Rakkude sorteerimine	Rakk	Sõltub markeritest: pinnavalgud koos alleeltundlike antikehadega või fluorestseeruva alleelse produktiga. Detektsioon valgu tasemel. Alleelse ekspressiooni stabiilsust saab uurida samu rakke uuesti sorteerides
Elusraku visualiseerimine	Rakk	Ekspressiooni dünaamika ajas, vajab geneetiliste konstruktide sisestamist. Detektsioon valgu tasemel
Ülegenoomsed lähenemised		
SNP-tundlikud mikrokiibid	Klonaalne populatsioon	Rakkude populatsiooni põhjal hinnang eelnevalt defineeritud proovide hulga transkriptidele
RNA sekveneerimine	Klonaalne populatsioon	Rakkude populatsiooni põhjal hinnang transkriptidele
Ühe raku RNA sekveneerimine	Rakk	Ühe raku põhjal hinnang transkriptidele. Tuvastab nii dünaamilise kui fikseeritud RMAE

FISH- fluorestsents *in situ* hübriidisatsioon; RMAE- juhuslik monoalleelne ekspressioon; RT-PCR – pöördtranskriptaas PCR

1.1.4. RNA sekveneerimine vermituse tuvastamiseks

Alleelse ekspressiooni määramiseks saab kasutada RNA sekveneerimist, kasutades lugemeid, mis katavad heterosügootseid SNP-sid, mille genotüüp on teada. RNA sekveneerimine on palju tundlikum meetod ning võimaldab vaadata palju lookusi korraga ja seetõttu leitakse kergemini uusi vermitud transkripte.

RNA sekveneerimist on seni palju kasutatud hiire vermitud geenide tuvastamiseks. RNA-seq-iga leiti hiire ajust 1300 uue vermitud lookuse, sealhulgas 484 mittekodeerivat RNA-d ja 347 soospetsiifilist geeni (Gregg *et al.*, 2010a; Gregg *et al.*, 2010b). Kuid vaatamata tehnilisest küljest täpse eksperimendi replikatsioonile ei õnnestunud korduskatsetel kinnitada enamikke uusi vermitud gene. Bioloogiline varieeruvus ja katse tehniline variatsioon tekitavad rohkem müra kui varem eeldati ning tulemusi on keeruline replitseerida. RNA-seqi puhul ei saa kasutada statistilisi meetodeid, mis eeldavad et iga lugem on juhuslikult ja üksteisest sõltumatult valitud ja vaba süstemaatilistest veast. Seetõttu tuleb müra piir selgelt defineerida. (DeVeale *et al.*, 2012)

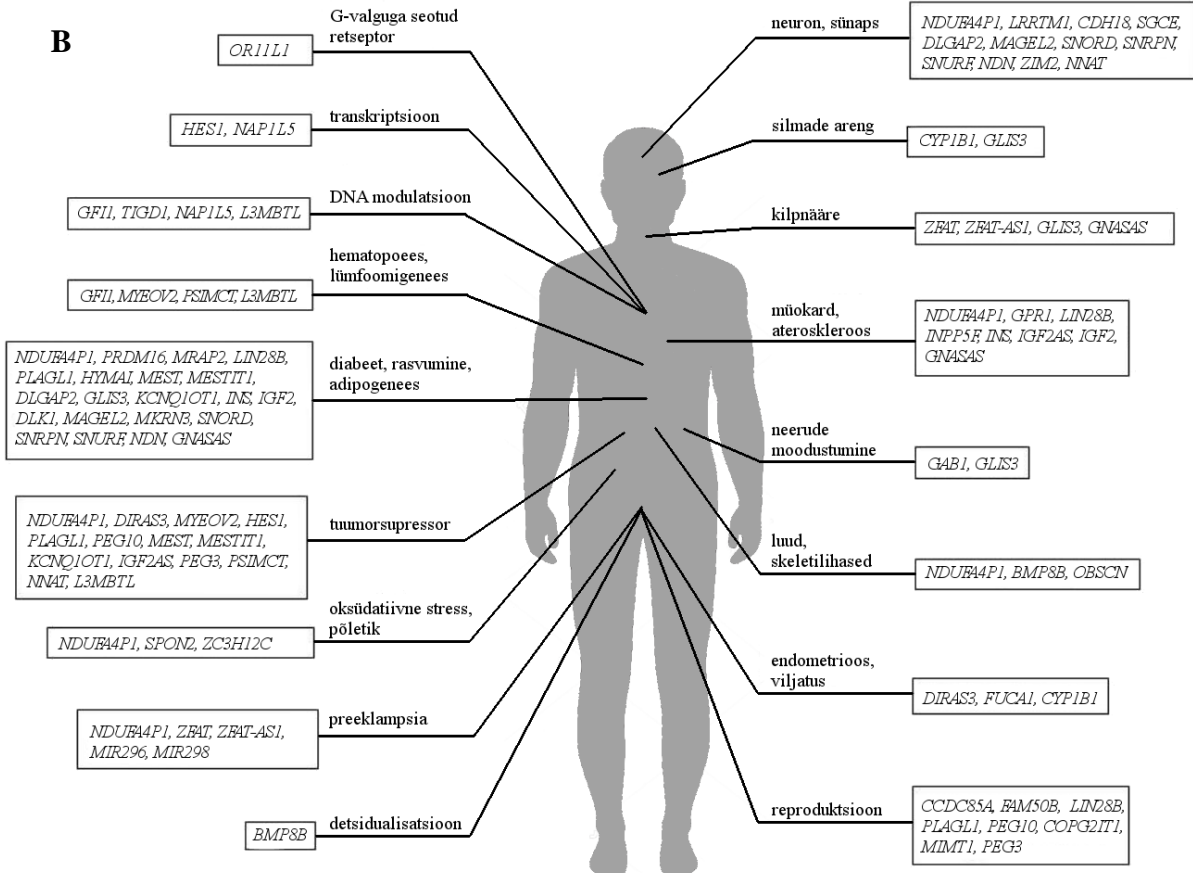
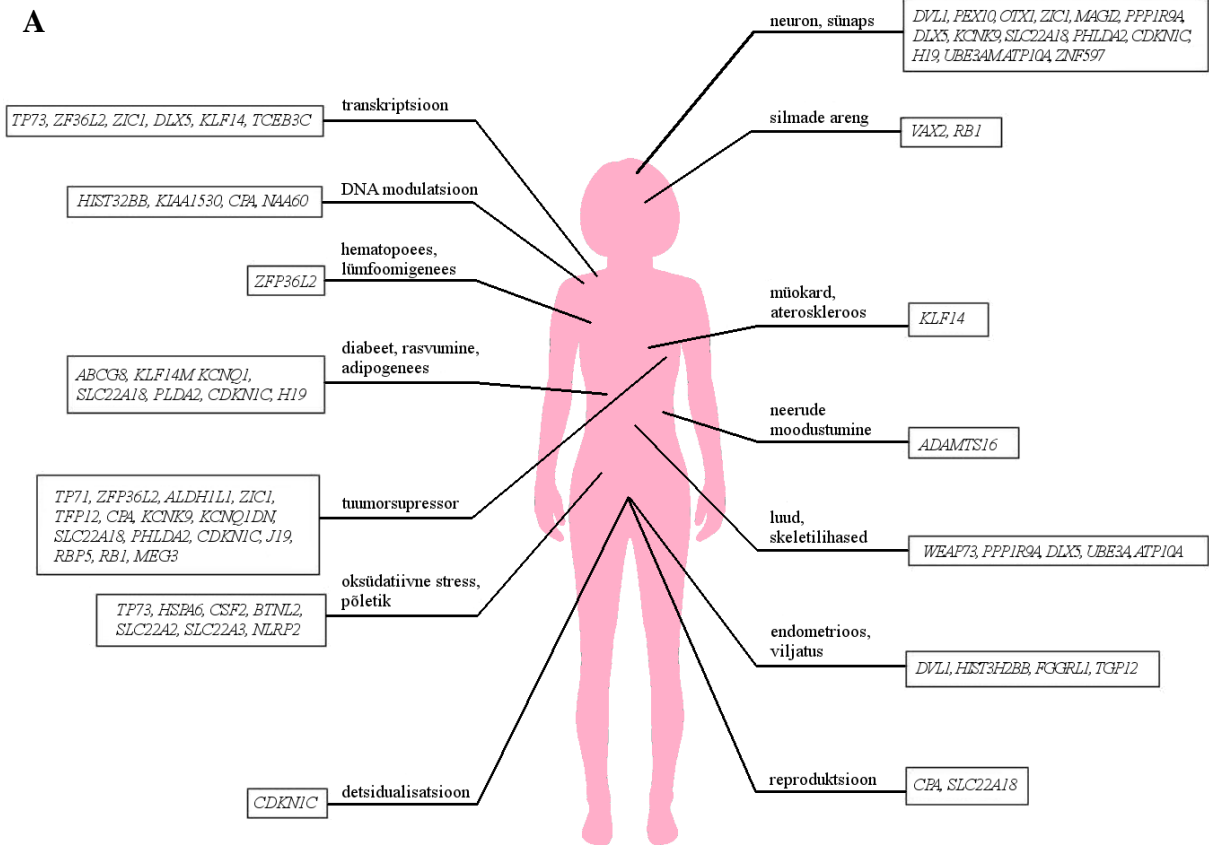
Siiani ainuke RNA sekveneerimise töö inimese platsentas on tehtud Metsalu *et al.* poolt (Metsalu *et al.*, 2014). Selles töös kinnitati nelja geeni vermitust platsentas (*AIM1*, *PEG10*, *RHOBTB3* ja *ZFAT-ASI*), tuvastasid neli uut potentsiaalset vermitud geeni (*LGALS8*, *LGALS14*, *PAPPA2* ja *SPTLC3*) ning leidsid neli geeni (*ABP1*, *BCLAF1*, *IFI30* ja *ZFAT*), millel esines juhuslik alleelne kallutus ekspresseerides ainult ühe vanema alleeli.

1.2. Inimese vermitud geenid

Geneimprint andmebaasi järgi on inimesel 213 vermitud geeni. Allpool oleval joonisel on välja toodud vermitud geenidega seostatud teadaolevad molekulaarsed protsessid ja organid, kus nad ekspresseeruvad (joonis 4).

Molekulaarsed protsessid

Organid, koed, viljakus



Joonis 4. Geneimprint andmebaasi põhjal inimeses vermitud geenid. A) Emapoolse alleelse ekspressiooniga geenid; B) Isapoolse alleelse ekspressiooniga geenid. Kohandatud Kobayashi, 2014 põhjal.

1.2.1. Patoloogiad

Nii geneetilised kui epigeneetilised mehhanismid võivad viia vermitud geenide ekspresseerumise häireteni, mis omakorda võib põhjustada haigusi. Geneetilisteks mehhanismideks on vermitud geeni funktsiooni kaotus või deletsioon ning uniparentaalne disoomia, epigeneetilised mehhanismid on DNA metülatsiooni muutumine vermitud geenide klastris (tabel 2). (Peters, 2014)

Tabel 2. Vermitud geenide ekspressiooniga seotud inimese sündroomid ja vastavad hiire mudelid.

Inimese patoloogiad	Peamised sümptomid	Põhjused	Hiire mudel
Prader-Willi sündroom	Arengu mahajäämus, raskused imetamisel, rasvumine, hüpogonadism, kognitiivne hälve, raevuhood ja obsessiiv-kompulsiivne häire	11 geeni isapoolse ekspressiooni kadumine 15q11-13 regioonis. Peamiselt isapoolse kromosoomi osalise deletsiooni või MatUPD15 tõttu. Harva ka vermimise häire tõttu (<i>SNURF/SNRPN</i>)	Olemas 30 hiire mudelit. Esineb neonataalset suremust, raskused imetamisel, sünnijärgne kasvupeetus, rasvumine.
Angelmani sündroom	Arengu mahajäämus, mikrotsefaalia, raske intellektuaalne puue, kõne- ja unehäired, iseloomulik EEG ja õnnelik käitumisviis (sage naermine)	Deletsioonid 15q11-13 regioonis emapoolselt kromosoomilt, <i>UBE3A</i> emapoolse ekspressiooni kadumine või PatUPD15. Harva ka vermimise häire tõttu.	4 hiire mudelit. Esinevad kognitiivsed häired, motoorikahäired ja ebanormaalne EEG.
Beckwith-Wiedemanni sündroom	makrosoomia, gigantism, suur keel, kõhuseina defektid, eelsoodumus embrüonaalseteks kasvajateks	ICR1 hüpermetülatsioon, ICR2 hüpometülatsioon <i>CDKN1C</i> vaigistamine või <i>IGF2</i> biallelne ekspressioon ja <i>H19</i> vaigistamine; PatUPD11.	<i>Igf2</i> transgeensed ja <i>Cdkn1c-knockout</i> hiire mudelid. Esineb loote ja vastsündinu suremust.
Silver-Russeli sündroom	Düsmorfism, IUGR ja sünnijärgne kasvupeetus.	ICR1 hüpometülatsioon, viib <i>IGF2</i> vaigistamiseni ja <i>H19</i> bialleelse ekspressioonini; MatUPD7	Spetsiifiline hiire mudel puudub
Mööduv neonataalne tüüp 1 diabeet	Vastsündinu hüperglütseemia ja IUGR	<i>PLAGL1</i> ja <i>HYMAI</i> üleekspresseerumine. Tervetel indiviididel monoalleelsed, patsientidel bialleelsed	Hiire mudelil esineb neonataalne hüperglütseemia, ei ole kasvupeetust

MatUPD15- Emapoolne 15. kromosoomi uniparentaalne disoomia; PatUPD11- Isapoolne 11. kromosoomi uniparentaalne disoomia; MatUPD7- Emapoolne 7. kromosoomi uniparentaalne disoomia; ICR1- Vermitud kontrollregioon 1, reguleerib isapoolse ekspressiooniga *IGF2* ja emapoolse ekspressiooniga *H19* vermitust; ICR2- Vermitud kontrollregioon 2, reguleerib isapoolse ekspressiooniga *KCNQ1OT1* ja emapoolse ekspressiooniga *CDKN1C* ja *PHLDA2* vermitust

1.2.2. Rasedustüsistused ja platsentas vermitud geenid

Imetajate puhul on näidatud mitmete vermitud geenide kriitiline roll platsenta arengu ja loote kasvu regulaatoritena (näited tabelis 2 ja tulemuste osas tabelis 7).

1.2.2.1. Emapoolsed: preeklampsia ja gestatsioonidiabeet

Emapoolsete rasedusaegsete komplikatsioonide gruppi kuuluvad raske preeklampsia (PE) ja gestatsioonidiabeet (GD). Gestatsioonidiabeet on raseduse jooksul esmakordselt diagnoositud diabeet. Preeklampsia on süsteemne haigus, mille klassikalisteks sümptomiteks loetakse raseduse ajal kõrgenenud vererõhku (vähemalt 140/90 mmHg) eelneva hüpertensioonita ning proteiinuuria vähemalt 0,3 g valku 24 h uriinist (English *et al.*, 2015). Preeklampsiat peetakse põhiliseks ema ja loote surma põhjustajaks.

Gestatsioonidiabeediga emade platsental on näidatud 6,2-7 protsenti madalamat keskmist metülatiooni *MEST* geenil kui kontrollidel. Üsasisesel kokkupuutel gestatsioonidiabeediga on pikaajalised efektid järeltulijate epigenoomis, näiteks *MEST* geeni epigeneetiline väärtusprogrammeerimine võib soodustada rasvumist läbi elu. (El Hajj *et al.*, 2013)

MEG3 seostatakse emaka spiraalarterite remodelleerimise häirega, mis võib viia preeklampsiani. MEG3 on madalama ekspressiooniga 80% preeklampsia platsentadel võrreldes kontrollidega. (Zhang *et al.*, 2015)

ZFAT on isapoolse ekspressiooniga monoalleelne geen platsentas, kuid bialleelne teistes kudedes. *ZFAT* ekspressioon on kolm korda madalam preeklampsia platsentades kui kontrollides ($p = 0,002$) (Barboux *et al.*, 2012).

Yu *et al.* on näidanud, et *H19* ekspresseerub bialleelselt esimesed 10 gestatsiooninäädalat normaalse raseduse jooksul ning seejärel muutub monoalleelseks. Pooltel preeklampsia juhtudel esines *H19* vermituse kadumist ning *H19-1* oli bialleelne ekspressioon raseduse lõpuni. Nendel patsientidel oli raske hüpertensioon, samas kui monoalleelselt ekspresseeritud *H19* patsientidel oli kergem haiguse vorm (Yu *et al.*, 2009).

PEG10 oli preeklampsia patsientide platsentas alla reguleeritud (Liang *et al.*, 2014).

PHLDA2 ekspressioonitase preeklampsia patsientides oli statistiliselt kõrgem kui kontrollides (Jin *et al.*, 2015).

Seerumis esinev DLK1 on madalama kontsentratsiooniga GD patsientidel võrreldes kontrollidega (P=0,001) (Li *et al.*, 2014b).

1.2.2.2. Lapsepoolsed: hüpotroofia ja makrosoomia

Rasedustüsistuste hulka kuuluvad ka loote kasvuga seotud häired. Gestatsiooniaja kohta suured (LGA, *Large for Gestational Age*) lapsed on defineeritud sünnikaaluga >90 protsentiili populatsiooni kasvukõveral. Gestatsiooniaja kohta väikesed (SGA, *Small for Gestational Age*) lapsed on defineeritud sünnikaaluga <10 protsentiili populatsiooni kasvukõveral. (Sildver *et al.*, 2015). Üsasisene kasvupeetus (*intrauterine growth restriction*, IUGR) on SGA alavorm, mille puhul loode ei saavuta oma kasvupotentsiaali geneetilise haiguse, infektsiooni või muul põhjustel. IUGR vastündinud on tavaliselt sünnikaaluga alla 10. kasvuprotsentiili, kuid leidub ka normaalpiiridesse jääva kaaluga imikuid. Selle tõttu on SGA ja IUGR eristamine keeruline. (Bamberg and Kalache, 2004)

DLK1 on alla reguleeritud SGA platsentades (P = 0,005) (Díaz *et al.*, 2013).

Platsenta *H19* geeniekspressioon on loote kasvupeetusega grupis oluliselt kõrgem kui kontrollidel ning patsientidel esines *H19* promotoorpiirkonna hüpomütülatsioon. *H19* mRNA tase jäi samaks nii vermituse kui selle kaotuse puhul kontrollides (Koukoura *et al.*, 2011).

Mööduv neonataalne tüüp 1 diabeet (*Transient neonatal diabetes mellitus type 1*, TNDM1) on harvaesinev haigus, mis möödub esimestel elukuudel, kuid võib tagasi tulla puberteedieas (Shield *et al.*, 2004). Lisaks on >95% TNDM1 patsientidel üsasisene kasvupeetus (*intrauterine growth retardation*, IUGR) (Flanagan *et al.*, 2007). Enamus TNDM1 juhtumitest on põhjustatud vermitud lookuse 6q24, kus asub kaks geeni *PLAGL1* ja *HYMAI*, üleekspressioonist (Mackay ja Temple, 2010). Nendel geenidel ekspresserub tavaliselt ainult isapoolne alleel ja emapoolne alleel on metüleeritud ning vaigistatud (Kamiya *et al.*, 2000). Tervetel indiviididel on need geenid monoalleelsed, TNDM1 puhul bialleelsed (Mackay *et al.*, 2002).

PHLDA2 ja *CDKN1C* on kõrgema ekspressioonitasemega IUGR platsentades, samas kui *MEST*, *MEG3*, *GNAS*, *PLAGL1*, *GATM* ja *IGF2* on madalama ekspressioonitasemega võrreldes mitte-IUGR platsentadega (McMinn *et al.*, 2006). Diplas *et al.* kinnitas *PHLDA2*, *NNAT* ja *PEG10* kõrgemat ekspressiooni ja *PLAGL1* ja *ZNF331* madalamat ekspressiooni (Diplas *et al.*, 2009).

2. EKSPERIMENTAALOSA

2.1. Töö eesmärgid

Antud magistritöö otsis vastust järgmisele võtmeküsimusele:

Kas kirjanduses avaldatud vermitud geenid omavad vanemaspetsiifilist avaldumismustrit ka inimese platsentakoes?

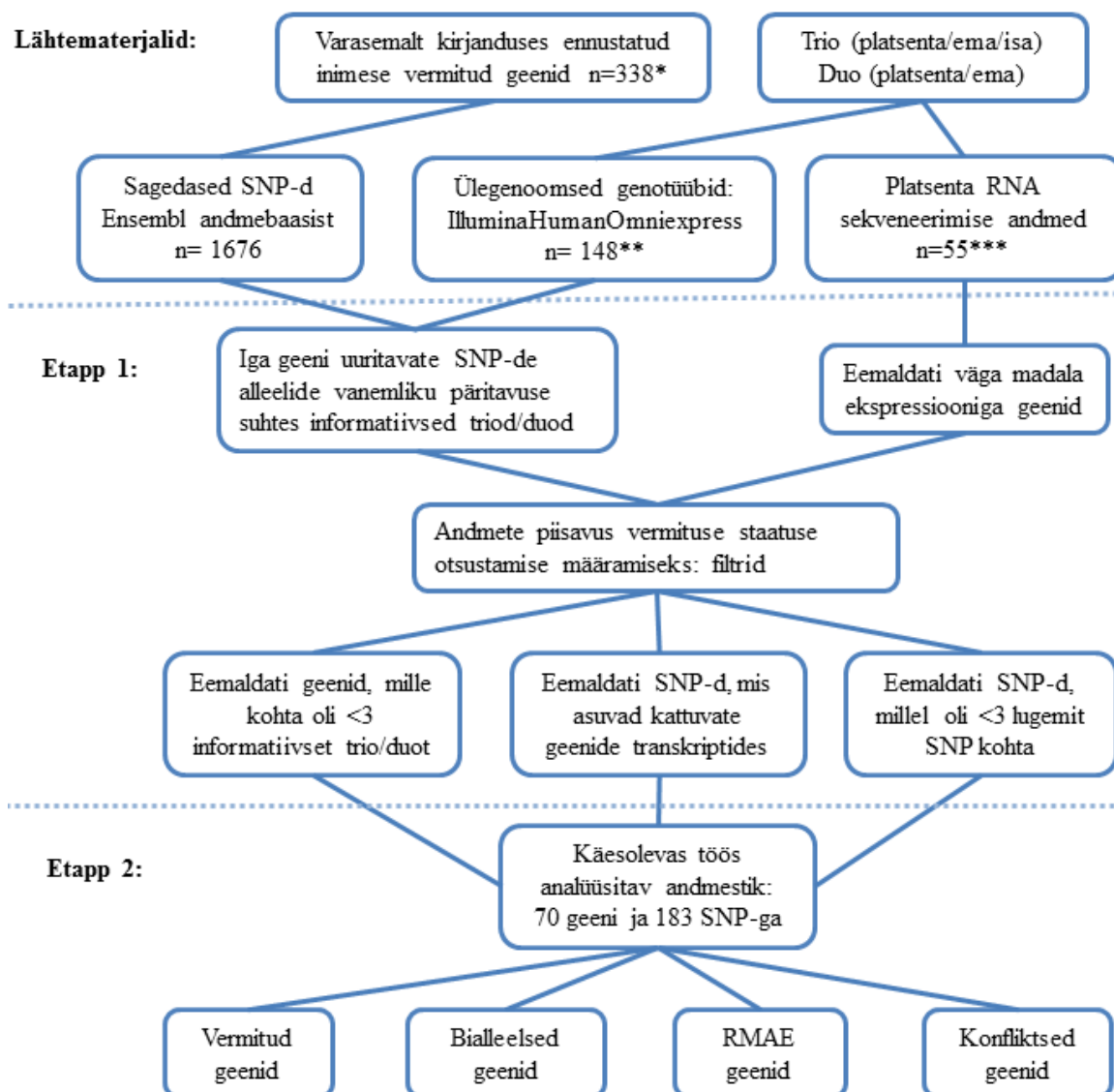
Kitsamad eesmärgid:

1. Määrata platsentas avaldunud ja kirjanduses raporteeritud potentsiaalselt vermitud geenide transkriptide vanemaspetsiifiline päritolu (Joonis 5).
2. Tuvastada vermitud geenide ekspressiooniolekut läbi raseduse kolme trimestri ning rasedustüsistuste puhul.

Magistritöö lähtematerjalideks oli kaks ülegenoomset analüüsi, mis kombineerides moodustasid tervikandmestiku ema-isa-platsenta triod/duo kohta:

1. 55 inimese platsenta koeproovide transkriptoomid (RNA-seq)
2. Ülegenoomsed genotüübid koos uuritavate perede (isa ja ema) veredest eraldatud DNA variantide andmetega.

Bioloogilised materjalid on kogutud SA TÜK Naistekliinikus ning tulemused on saadud TÜ MRI Inimese molekulaargeneetika uurimisrühmas.



Joonis 5. Käesoleva magistr tööplaan. RMAE on juhuslik monoalleelne ekspressioon, mis on defineeritud kui ühe alleeli ekspressioon kus juhuslik alleelivalik toimub somaatilisel, nii et erinevad organismi rakud ekspresseerivad erinevaid allelele (Reinius and Sandberg, 2015). Sage SNP defineeriti kui Ensembl-i andmebaasis transkriptis asuv SNP, mille $MAF \geq 0,1$. Madala ekspressiooniga geen defineeriti kui geen, millel on keskmiselt <50 lugemit kõigil indiviididel. Informatiivseks trioks/duoks lugesin pere, kellel saab määrata alleelset pärilikkust (platsenta heterosügoot ning vähemalt üks vanematest homosügoot). Vermitud ja bialleelsete geenide definitsioonid on antud detailsemalt metoodikas ja tulemustes.

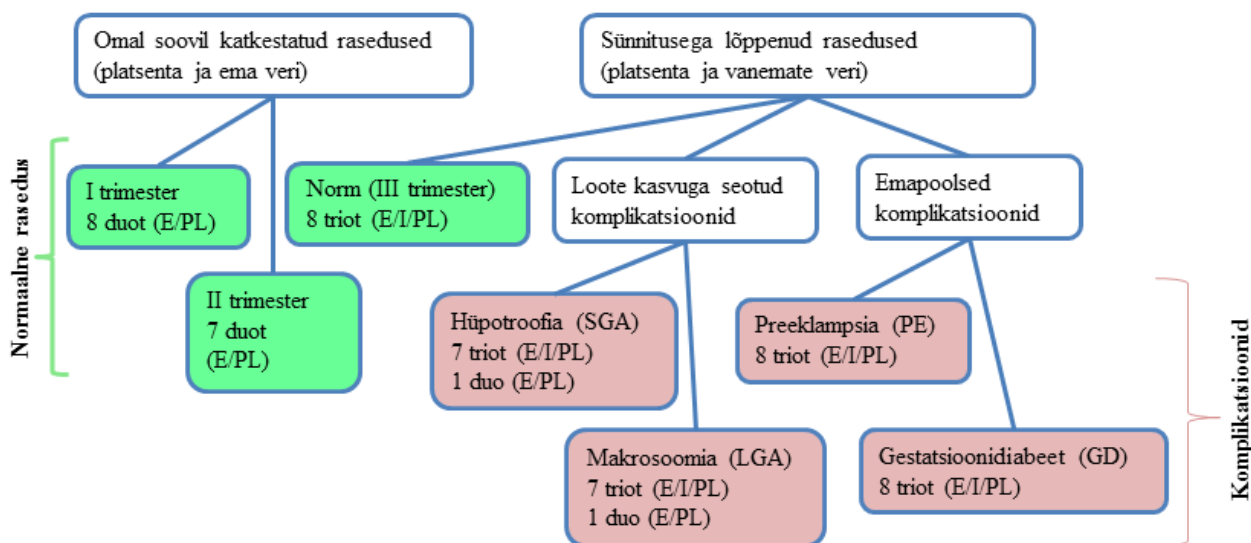
* – Allikad: Geneimprint koduleht (veebiaadress 3), Metsalu *et al.*, 2014, Court *et al.*, 2014;

** – Kasak *et al.*, 2015; *** – Söber *et al.*, 2015.

2.2. Materjal ja metoodika

2.2.1. Valim

Potentsiaalselt vermitud geenide dünaamika uurimiseks läbi raseduse oli käesoleva magistritöö fookuses kolm raseduse trimestrit. Lisaks võrdlesime õigeaegsete sünnituste platsenta vermitud geenide ekspressiooni nelja tüsistusega rasedusega (Joonis 6).



Joonis 6. Valimi alagrupid. E - ema, I - isa, PL- platsenta. Kolm trimestrit on märgitud rohelise taustaga. Tüsistustega rasedused märgitud punase taustaga. Alagruppe iseloomustatakse põhjalikult järgnevatel peatükkides.

2.2.1.1. I ja II trimester

Valimit on põhjalikult kirjeldatud Kristiina Rulli doktoritöös (Rull, 2009). Platsenta ja ema vereproovid koguti aastatel 2003-2006 Tartu Ülikooli Kliinikumis Kristiina Rulli poolt naistelt, kes läbisid a) omal soovil (kirurgilise) aborti esimesel trimestril (8 juhtumit); b) terapeutilise meditsiiniliselt indutseeritud aborti teisel trimestril ema riskifaktorite tõttu (7 juhtumit, loote anomaaliaid ei tuvastatud). SA Tartu Ülikooli Kliinikumi Ühendlaboris puhastati mikroskoobi all I trimestri platsenta koorionihatud Piret Illisoni poolt. Ema ja loote parameetrid on toodud tabelis 3. Välistati suured kromosomaalsed ümberkorraldused, näiteks aneuploidiaid ja kromosomaalsed anomaaliaid (Ühendlabor, Tartu Ülikooli Kliinikum). Varasemalt on antud valimit kasutatud ja detailselt kirjeldatud Kasak *et al.*, 2015 teadusartiklis.

Tabel 3. Katkestatud raseduse valimi kliinilised parameetrid, keskmine \pm SD.

	Ühik	I trimester	II trimester
Emade parameetrid			
Vanus	aasta	25,5 \pm 5,88	25,7 \pm 8,65
KMI enne rasedust	kg/m ²	22,9 \pm 3,71	21,6 \pm 2,87
Eelnevaid rasedusi	0/1/ \geq 2	2/5/1	4/1/2
Gestatsiooniaeg	nädal \pm päev	9 \pm 12	19 \pm 9
Loote parameetrid			
Sugu	N/M	5/3	4/3

2.2.1.2. Tüsistusteta ja tüsistustega rasedused

REPROMETA (REPROgrammeeritud loote ja/või ema METAbolismi) uuringu valim koguti aastatel 2006-2011 Tartu Ülikooli Kliinikumi Naistekliinikus platsenta geeniekspressiooni profiili määramiseks. Proovide kogumist koordineerisid Maris Laan, Kristiina Rull, Pille Vaas ning Pille Teesalu. REPROMETA valim sisaldab kliinilist informatsiooni ning bioloogilist materjali üksikrasedustest gestatsiooniajaga 37-41 nädalat. Raseda kaardist ning sünnitusloost koguti kliiniline ja epidemioloogiline informatsioon ema haiguste, kehamõõtude ning eelnevate sünnituste kohta. Loote sünniinformatsioon sisaldas gestatsiooniaega, sünnikaalu- ja pikkust, sugu ja platsenta kaalu. Valimist jäeti välja juhtumid a) loote- ja kromosoomianomaalitega; b) perekonnad, kus esines pärilikke haigusi (süsteemne erütematoosne luupus, reumatoidartriit, trombofiilia – V faktori geeni Leideni mutatsioon); c) patsiendid, kellel esines enne rasedust diabeeti, kroonilist hüpertensiooni ja kroonilisi neeruhaigusi, vältimaks eelsoodumusi rasedustüsistusteks. Hiljuti on antud valimit kasutatud Söber *et al.*, 2015 teadusartiklis.

Valimis olevad inividid on jaotatud alagruppideks sünnikaalu ning rasedustüsistuse puudumise/olemasolu alusel. Kontrollgrupp koosnes tüsistusteta rasedustest, mille tulemusena sündis normaalkaalus laps (sünnikaal 10-90 protsentiili vahel; platsenta n=8). Loote kasvuga seotud tüsistustega gruppi kuulusid gestatsiooniaja kohta väikesed (SGA, *Small for Gestational Age*; <10 protsentiili; platsenta n=8) ja gestatsiooniaja kohta suured vastsündinud (LGA, *Large for Gestational Age*; >90 protsentiili; platsenta n=8). Kaaluprotsentiilid SGA ja LGA defineerimiseks on arvatud Eesti sünniregistris oleva informatsiooni alusel (Sildver *et al.*, 2015). Vastsündinud olid kasvupeetusega (IUGR; *IntraUterine Growth Restriction*) kui viimane ultraheli näitas häirunud Doppleri signaali

nabanööri või keskmises tserebraalses arteris ja/või vastsündinu kaal ja/või kõhuümbermõõt oli <10 protsentiili aga peaümbermõõt >10 protsentiili. SGA ja LGA alagrupi emadel ei olnud muid rasedustüsistusi. Emapoolsete rasedusaegsete komplikatsioonide gruppi kuulusid raske preeklampsia (PE; platsenta n=8) ja gestatsioonidiabeet (GD; platsenta n=8). Raske preeklampsia defineeriti kui hüpertensioon (süstoolne vererõhk ≥ 160 mmHg ja/või diastoolne vererõhk ≥ 110 mmHg), lisaks võis esineda proteiinuuria (≥ 5 g valke uriinis 24 h jooksul). Kõigi juhtumite puhul oli tegemist hilise preeklampsiaga (hüpertensioon ja/või proteiinuuria diagnoositi pärast 34. gestatsiooninädalat). GD diagnoositi kui 75 g suukaudse glükoosi tolerantsuse testil 24-28 gestatsiooninädalaloli vereplasmas glükoositase üle 5,1 mmol/l ja/või 1h ja 2h plasma glükoositaset vastavalt >10 mmol/l ja >8,5 mmol/l. Gestatsiooniaeg ning mees- ja naissoost laste suhe püüti hoida võimalikult võrdsena. Kasutatud alavalimite vastsündinute ning vanemate kliiniliste parameetrite keskmised koos standardhällbega on toodud tabelis 4.

Tabel 4. Indiviidide kliinilised parameetrid, keskmine \pm SD. Modifitseeritud Söber *et al.*, 2015 ja Kasak *et al.*, 2015 põhjal.

	Ühik	Norm	SGA	LGA	PE	GD
Emade parameetrid						
Vanus	aasta	29,3 \pm 7,85	24,5 \pm 3,51	30,9 \pm 5,06	27,4 \pm 7,39	30,3 \pm 5,15
KMI enne rasedust	kg/m ²	23,9 \pm 3,77	20,6 \pm 2,92	24,8 \pm 4,61	26,3 \pm 5,00	26,3 \pm 7,85
Rasedusaegne kaaluiive	kg	17,2 \pm 3,52	13,3 \pm 3,47*	21,7 \pm 7,25	10,9 \pm 3,36*	14,8 \pm 5,76
Eelnevaid rasedusi	0/1/ \geq 2	3/4/1	7/1/0*	2/2/4	6/1/1	3/3/2
Gestatsiooniaeg	nädal \pm päev	40 \pm 12	39 \pm 10	40 \pm 4	38 \pm 4	39 \pm 6
Keisrilõige/ vaginaalne sünnitus		5/3	6/2	3/5	2/6	3/5
Isade parameetrid						
Vanus	aasta	31,8 \pm 5,99	27,1 \pm 5,64	35,8 \pm 8,28	32,8 \pm 8,94	33,5 \pm 7,05
KMI	kg/m ²	25,1 \pm 3,83	22,7 \pm 2,04	29,9 \pm 5,65	28,8 \pm 6,34	27,8 \pm 4,34
Lapse parameetrid						
Kaal	g	3703 \pm 392	2442 \pm 235*	4726 \pm 208*	2794 \pm 488*	4269 \pm 238*
Pikkus	cm	51,3 \pm 1,89	46,3 \pm 1,04*	53,4 \pm 1,18*	47,6 \pm 1,51*	52,4 \pm 1,30
Platsenta kaal	g	571,3 \pm 115,8	397,5 \pm 89,9*	816,2 \pm 116,0*	476,9 \pm 119,9	658,1 \pm 185,2
IUGR	arv	0	5	0	4	0
Sugu	N/M	3/5	5/3	4/4	4/4	5/3

SGA – *small for gestational age*, väike gestatsiooniaja kohta; LGA – *large for gestational age*, suur gestatsiooniaja kohta; PE – preeklampsia; GD – gestatsioonidiabeet; * – p-väärtus <0,05 võrreldes kontrollidega; IUGR- *intrauterine growth retardation*, üsisisene kasvupeetus.

2.2.2. Genoomse DNA eraldamine

Uuringus analüüsitud platsenta ja vererakkudest eraldatud genoomsed DNA-d pärinevad TÜ MRI Inimese Molekulaargeneetika uurimisrühma bioloogiliste materjalide kogust. DNA-d on eraldatud verest aastal 2012, platsentast aastatel 2009-2013. Käesoleva magistr töö autor on biopanga loomises kaasa löönud aastatel 2012-2013.

2.2.2.1. DNA eraldamine verest

Inimese genoomset DNA-d eraldati plasmafraktsioonita EDTA verest, mida hoiti -80°C juures. Külmunud vered soojendati toatemperatuuril, seejärel pipeteeriti vastavalt vere kogusele 25ml või 50ml tuubi. Lüüsietaapid tehti jääl. Järgnevad kogused on 1 ml vere kohta.

Verele lisati 4 ml puhvrit A (Tabel 5). Vere ja puhvri korralikuks segunemiseks loksutati õrnalt tuubi. Saadud segu tsentrifugeeriti 30 min 4°C ja 1942 RCF (*relative centrifugal force*) juures. Supernatant eemaldati ning vererakkude vähese lüüsumise puhul korraldati eelnevat lüüsetappi. Järgmises pesuetapis lisati 2 ml puhvrit A. Sade loksutati põhjast lahti ning tsentrifugeeriti 20 min 4°C ja 1942 RCF juures. Supernatant valati ära, nõrutati ning sademele lisati 0,9 ml puhvrit B (Tabel 5). Sade toksiti lahusesse. Tilkhaaval lisati 100 µl puhvrit C (Tabel 5) ja segu inkubeeriti üleöö loksutil 37°C juures. Järgmisel päeval lisati 270 µl küllastatud NaCl lahust. Tuubi raputati tugevalt 15 sek ning tsentrifugeeriti toatemperatuuril 30 min 3452 RCF juures. Supernatant pipeteeriti uude tuubi ning lisati topeltkoguses filtreeritud 96% etanooli. Tuubi keerati segamiseks. Klaaskonksuga eemaldati DNA tuubist ning pandi 500 µl TE lahusesse. Eraldatud DNA-d säilitati -80°C juures. Eraldatud DNA-de kontsentratsioonid mõõdeti NanoDrop ND-1000 spektrofotomeetriga. Genoomse DNA terviklikkust kontrolliti 1% agarosgeelil (0,5x Tris/boorhape/EDTA (TBE)).

Tabel 5. Kasutatud puhvrite koostis.

Puhver A	Puhver B	Puhver C
10,95 % sahharoosi	400mM NaCl	10% SDS
1mM Tris-HCl	2mM EDTA	2mg/ml Proteinaas K
5mM MgCl ₂	10mM Tris-HCl	
1% Triton X-100		

2.2.2.2. DNA eraldamine platsentast

Platsentasid hoitakse -80 °C juures ning enne DNA eraldamist tõstetakse proov -20°C juurde. Igast platsentast murtakse ühekaupa 0,02 g suurune tükk ning DNA eraldamine toimub NucleoSpin® Tissue kitiga (Machery-Nagel, *Genomic DNA from tissue*). Seejärel lisatakse 180 µl puhver T1-e ja 28 µl proteinaas K-d. Platsentatükk pidi olema kaetud lahusega. Pärast platsentatüki segamist Vortexil tuleb teda inkubeerida 4h 56 °C juures ning iga tunni tagant vorteksida. Lüüsimise lõpuks ei tohiks platsentatükki näha olla. Pärast Vortexil segamist lisati 200 µl puhver B3 ning inkubeeriti 10 min 70 °C juures. Seejärel tuleb uuesti Vortexil segada ja lisada 210 µl 100% etanooli ning uuesti segada. Järgmiseks tuleb kogu lahus panna NucleoSpin® koekoloni ning tsentrifugeerida 1 min toatemperatuuril 11 200 RCF juures. Lävivool valati välja ning lisati 500 µl puhver BW-d. Tsentrifugeeriti uuesti 1 minut 11 200 RCF juures ja valati lävivool ära. Lisati 600 µl puhver B5-te ning tsentrifugeeriti 11 200 RCF juures 1 minut. Valati lävivool ära ning fuugiti uuesti 3 min 11 200 RCF juures. Seejärel asetati kolonn lõplikusse epsi nr I ning lisati 100 µl puhver BE-d pipeteerides võimalikult

membraani vastu. Pärast 5 minutit toatemperatuuril inkubeerimist tsentrifuugiti 2 minutit 11 200 RCF juures. Läbivool tõsteti eemale ning kolonn tõsteti epsi nr II. Lisati 50 µl puhver BE-d ning inkubeeriti 5 minutit. Seejärel tsentrifuugiti 2 minutit 11 200 RCF juures. Lisati uuesti 50 µl puhver BE-d ning inkubeeriti 5 minutit. Pärast 2 minutit 11 200 RCF juures fuugimist visati kolonnid ära. Läbivooludes I ja II sisalduva DNA kontsentratsioonid mõõdeti NanoDrop ND-1000 spektrofotomeetriga. Eraldatud DNA säilitati -80 °C juures.

2.2.3. Platsentast RNA eraldamine ja sekveneerimine

Pärast vaginaalset sünnitust või keisrilõiget võeti ühe tunni jooksul +4 °C juures hoitud platsentast proov. Platsenta keskmisest regioonist võeti täispaks 2-3 cm suurune blokk, mis pesti 1 x PBS lahuses ema vere kontaminatsiooni eemaldamiseks ning seejärel asetati RNAlater lahusesse (Ambion Inc, Life Technologies). Proovi hoiti -80 °C juures RNA eraldamiseni. Totaalne RNA eraldati 200-300 mg homogeniseeritud platsentakoest kasutades TRIzol reagenti (Invitrogen, Life Technologies) ning puhastati RNeasy MinElute kolonnidega (Qiagen, Holland). RNA puhtuse taset ja kontsentratsioon määrati NanoDrop® ND-1000 UV-Vis spektrofotomeetris (Thermo Fisher Scientific Inc., USA) ja RIN (RNA intaktsuse number) hinnati Agilent 2100 bioanalüsaatoriga (Agilent Technologies, USA). RIN väärtused olid vahemikus 5,9-8,9 (mediaan 6,9). Platsentast eraldati RNA Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudis Kristiina Rull, Rait Kivi, Jaana Männik ja Mario Reimani poolt.

Raamatukogude valmistamine, RNA sekveneerimine ja esmane bioinformaatiline töötlus viidi läbi FIMM-is (Institute for Molecular Medicine Finland; veebiaadress 1). Raamatukogu valmistamise etapis osalesid Siim Sõber ja Mario Reiman, ülejäänud tehti FIMM sekveneerimise tuumiklabori poolt. Kõrge kvaliteediga DNA-vabast RNA-st (5 µg) eemaldati rRNA kasutades Ribo-Zero™ rRNA Removal kitti (Epicentre). Raamatukogu valmistati Nextera™ Technology (Illumina) abil. Totaalne RNA sekveneerimine tehti Illumina HiSeq2000-ga kasutades 46 aluspaari (bp) ja 100 bp paarislugemeid. Esialgne andmeanalüüs ja geeniekspressioonide hindamine teostati RNA-Seq pipeline v2.4 (FIMM sekveneerimise tuumiklabor) abil. Kvaliteedi kontroll viidi läbi FastQC versioon 0.10.0-ga. Pythoni skriptidega eemaldati adaptorite, rRNA, mitokondriaalse DNA ja homopolümeeride pikad järjestused. Lugemid joondati inimese referentsgenoomile (GRCh37.p7/hg19) TopHat versiooniga 2.0.3 (Kim *et al.*, 2013). Joondamise edukus oli 46 bp paarislugemite puhul keskmiselt 83,6 % (76,5%-87,3%) ning 100 bp paarislugemite puhul 57,8% (56,3%-59,3%). Pärast filtreerimist oli 46 bp paarislugemitega lugemeid igas transkriptomis keskmiselt 37,6 miljonit ($25,2 \times 10^6$ - $50,5 \times 10^6$) ning 100 bp paarislugemitega 57,6 miljonit ($45,2 \times 10^6$ - $70,0 \times 10^6$) lugemit. Iga proovi kohta on 17- kordne kattuvus üle eksoonse ala.

Transkriptid kvantifitseeriti cufflinks versioon 2.0.2-ga (Trapnell *et al.*, 2010) ja geenide ekspressiooni HTSeq-iga (Anders *et al.*, 2015). RNA-seq andmete geenide diferentsiaalset ekspressiooni testiti DESeq (Anders and Huber, 2010) ja DESeq2 (Love *et al.*, 2014) R (R Development Core Team, 2011) pakettidega. Statistilist analüüsi on täpsemalt kirjeldatud Söber *et al.*, 2015 teadusartiklis. Käesolevas töös kasutatavad SNP-d on RNA-seqist tuvastatud ainult minoorse alleeli olemasolul ning SNP-d ja kõrge kvaliteediga lugemite arvud mõlemalt alleelilt määrati SAMtools v0.1.18 (Li *et al.*, 2009) FIMM poolt. 10 bp kaugusel ekson-intron piirist asuvad introonsed SNP-d filtreeriti välja, kuna suure tõenäosusega oli tegemist joonduse veaga.

2.2.4. Perede ülegenoomsete SNP-de genotüpiseerimine

Platsenta ja vanemate vere genoomne DNA genotüpiseeriti Illumina HumanOmniExpress--12-v1/24-v1 BeadChips kiibiga (715 000 markerit, mille keskmine vahekaugus 2,1 kb) Eesti Geenivaramu poolt (veebiaadress 2). Proovide genotüpiseerimise keskmine tuvastamismäär oli 99,6% (mediaan 99,7%). Antud valimit on kasutatud Kasak *et al.*, 2015 teadusartiklis.

2.2.5. Potentsiaalselt vermitud geenide valik uuringusse

Uuritavad inimese geenid valiti Geneimprint andmebaasist (n=286 ;veebilink 3), kus on välja toodud vermitud geenide kandidaadid ning ekspresseeritud alleeli pärituolu. Välja jäid geenid, millel puudus Ensembl andmebaasis sissekanne (n=7). Lisaks kaasasime uuringusse hiljuti ilmunud Court *et al.*, 2014 (n=23 geeni) ja Metsalu *et al.*, 2014 (n=29 geeni) artiklid, kus on tuvastatud uusi platsentas vermitud geene. Lisas 1 on välja toodud vermitud geenide kandidaatide loetelu.

2.2.6. Mõisted

Lugemite käsitlemisel on kasutatud järgmisi põhidefinitsioone:

FPKM on fragmente tuhande aluspaari kohta miljoni joondatud fragmendi kohta (*Fragments Per Kilobase per Million mapped fragments*).

Toorlugem on RNA-seqi BAM failist saadud lugemite arv antud positsioonis.

Kõrge kvaliteediga lugem on SAMtools-i poolt filtreeritud kvaliteetsed õigesti joondatud lugemid.

Normaliseeritud geeniekspressiooni mediaan (NGEM) on DESeq ja DESeq2-ga normaliseeritud HTSeqi lugemite arvude mediaan geeni kaupa üle indiviidide.

2.2.7. Käesolevas magistritöös analüüsitava andmestiku filtrid

Enne uuritavate geenide RNA lugemite vanemlikku päritolu määramist läbis analüüsiv andmestik mitu kvaliteedi- ja piisavuse filtrit. See oli vajalik selleks, et kinnitada või ümber lükata nende geenide vanemspetsiifilise avaldumise staatus usaldusväärsete andmetele toetudes.

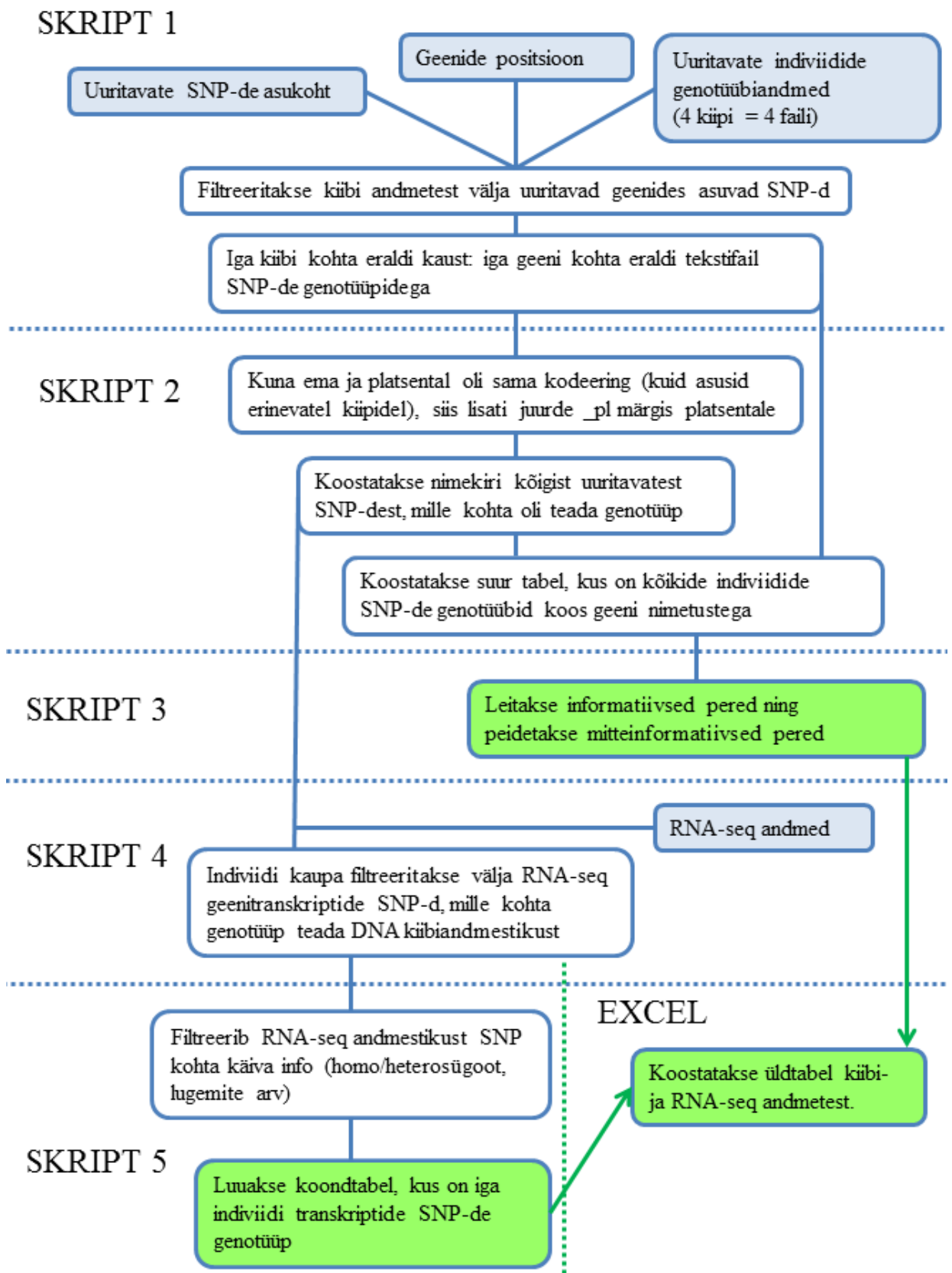
Esmane filtreerimine toimus geenide avaldumise taseme järgi (Söber *et al.*, 2015). Eemaldati geenid, mille normaliseeritud geeniekspressiooni mediaan jäi alla 51 lugemi, kuna selliste geenide puhul on vermituse olekut raske määrata vähese sekveneerimise sügavuse või madala geeniekspressiooni tõttu.

Teises etapis teostati potentsiaalselt uuritavate SNP-de valik Ensembl andmebaasist. Vermitud kandidaatgeenide 5' ja 3' UTR-is ning eksonis asuvad SNP-d valiti Ensembl Biomart (Ensembl versioon 83) tööriistaga. Lisaks filtreeriti välja haruldasemad SNP-d (MAF $\leq 10\%$), et uuritavad SNP-d oleks suurema tõenäosusega antud valimis olemas.

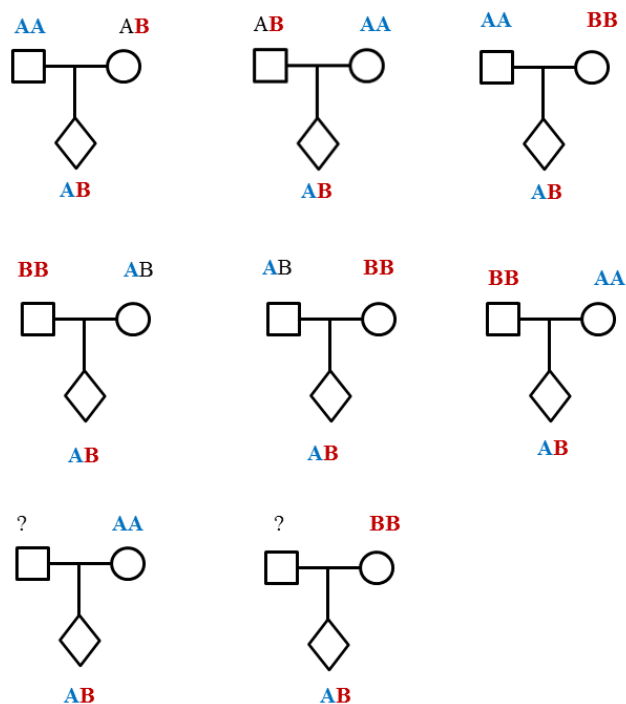
Kolmandas andmetöötamise etapis kirjutati Pythoni v3.4.1. (Python Software Foundation, veebiaadress 4) programmeerimiskeeles järgnevad skriptid (joonis 7) :

1. Otsiti kandidaatgeenide positsiooni järgi DNA kiibilt neis asuvate SNP-de genotüübid. Seejärel võrreldi neid uuritavate Ensembl SNP-dega. Loodi iga kiibi kohta üks kaust, mis sisaldas iga geeni kohta tekstifaili SNP-de genotüüpidega.
2. Eristati ema ja platsenta kodeering üksteisest, sest alguses asusid nad erinevatel kiipidel kuid sama koodiga. Seejärel koostati nimekiri kõigist informatiivsetest SNP-dest, mille kohta on kiibil olemas genotüüp. Loodi suur tabel, kus on kõik invidiidid perede kaupa kõrvuti ning iga indiviidi all uuritavate geenide SNP-de genotüübid. Seejuures arvestati erinevate inimese referentsgenoomi versioonidega.
3. Seejärel määrati igale SNP-le informatiivsed pered, kus platsenta oli heterosügootne ning vähemalt üks vanematest homosügootne selles positsioonis (joonis 8).
4. Kasutades eespoolt loodud informatiivsete SNP-de nimekirja, filtreeriti need SNP-d RNA-seq andmetest.
5. RNA-seq andmestikust filtreeriti välja iga transkripti SNP kohta käiva olulise info (hetero/homosügootsus, kõrge kvaliteediga lugemite arv). Luuakse koondtabel, kus on iga platsenta kohta informatiivsete SNP-de transkriptide genotüüp.

Seejärel ühendati EXCELis skriptis 3 saadud informatiivsete SNP/perede tabel ja skriptist 5 saadud RNA-seq transkriptide SNP-de genotüübi tabel.



Joonis 7. Selgitav skeem töös kasutatud skriptide tööpõhimõtetest. Sinisel taustal on esile toodud algandmed, rohelisel taustal andmeanalüüsi tulemusena saadud koondtabelid.



Joonis 8. Informatiivsed pered, kus platsenta on heterosügootne ning vähemalt üks vanematest on homosügoot. Tumepunase ja sinise värviga on esile toodud alleelne päritolu.

Neljandas etapis eemaldati veel ebapiisava informatsiooniga geenid ja SNP-d:

- geenid, millel oli ≤ 3 informatiivset pere;
- käsitsi SNP-d, mis asusid intronites või kattuvates geenides;
- SNP-d, mille lugemite mediaan oli ≤ 3 toorlugemi/SNP kohta üle indiviidide (vastavalt Pinter *et al.*, 2015 kriteeriumitele). Tarmo Puurand (TÜ MRI Bioinformaatika osakonna ja Geenivaramu programmeerija) abistas lugemite arve positsiooni kohta saada.

2.2.8. Andmeanalüüs vermitud geenide oleku määramiseks

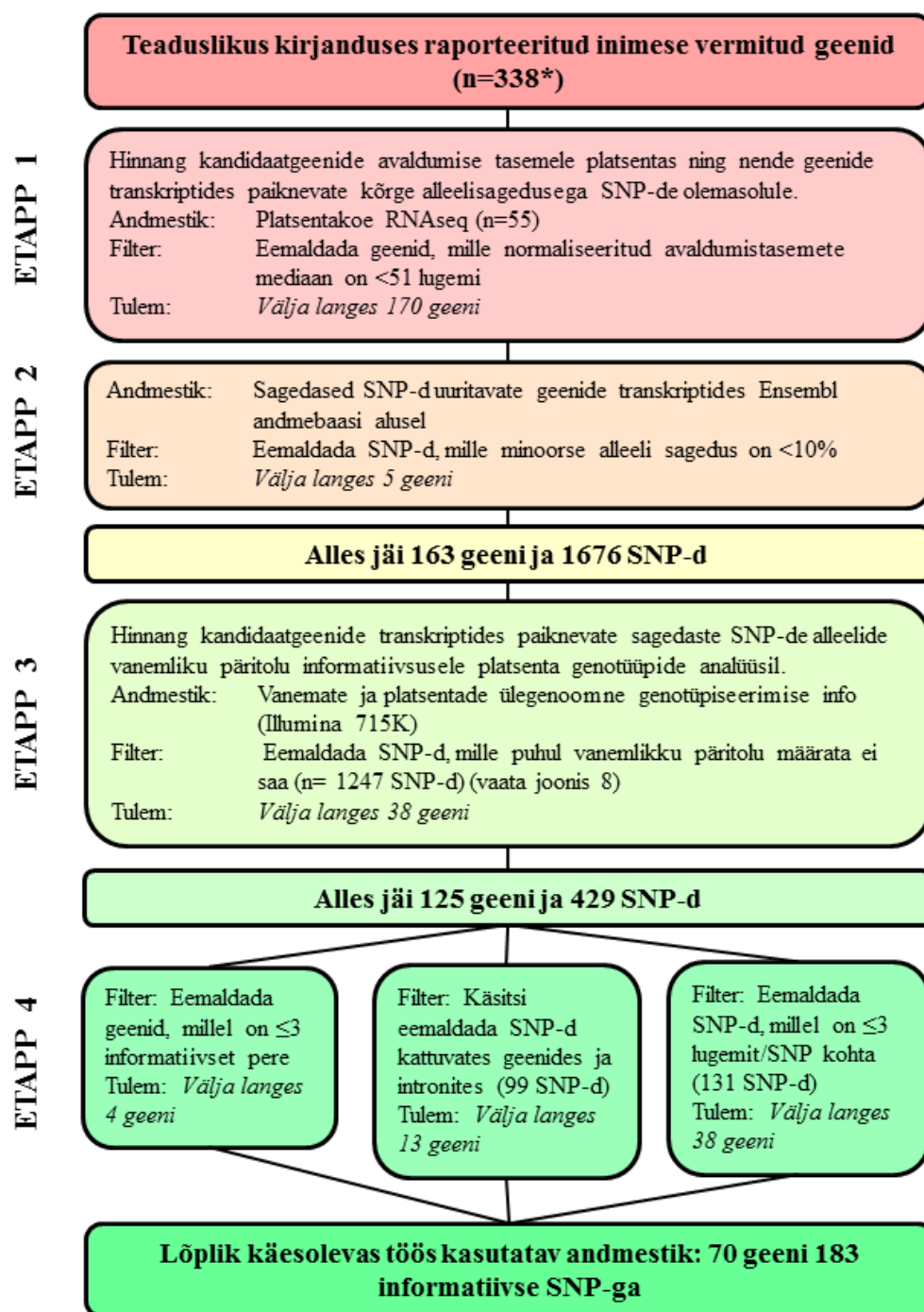
Bioinformaatilise andmeanalüüsi kriteeriumid antud uurimistöös:

- Geeni lugemin vermituks siis, kui $\geq 80\%$ -l platsentadest olid informatiivsete SNP-de positsioonid monoalleelsed ja pärit ühelt vanemalt.
- Geeni lugemin bialleelseks siis, kui $\geq 40\%$ -l platsentadest olid informatiivsete SNP-de positsioonid bialleelsed.
- Geeni lugemin RMAE-ks siis, kui $\leq 40\%$ platsentadest olid bialleelsed, ülejäänud olid monoalleelsed, aga mõlema vanema alleelid olid esindatud.
- Ülejäänud geenid klassifitseeriti konfliktseteks platsentas vermitud staatuse kohta.

2.3. Tulemused

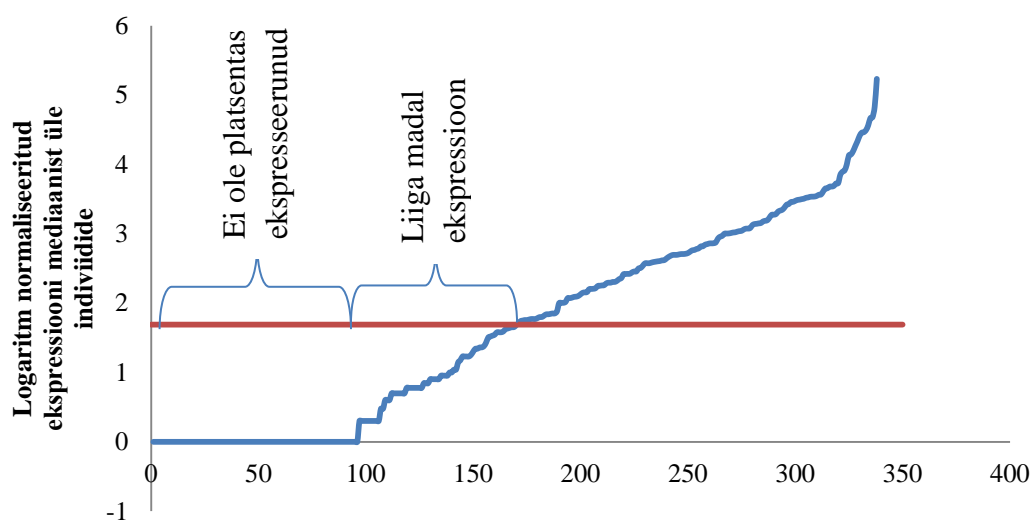
2.3.1. Ainult pooled kirjanduses raporteeritud vermitud geenidest on avaldunud platsentas

Rangete andmete kvaliteedi- ja piisavuse kriteeriumite järgi filtreerimise tulemusena sai vanemlikku päritolu usaldusväärselt kinnitada või ümber lükata 338 geenist 70-l. Allpool toodud skeem võtab kokku neli etappi (joonis 9).



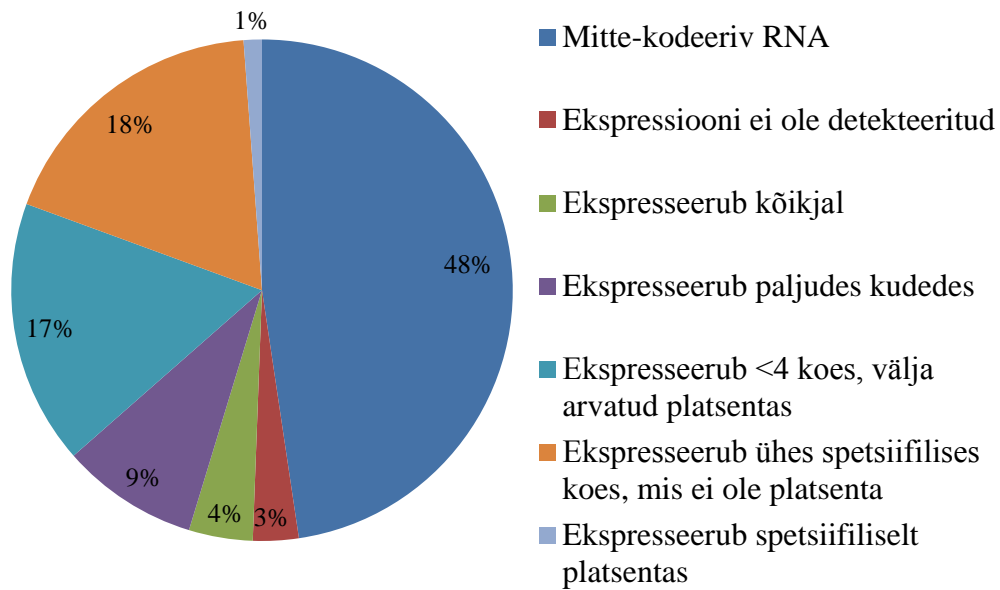
Joonis 9. Andmestiku filtreerimise etapid koos kriteeriumite ja tulemustega.

Esmalt hinnati andmete piisavust geeniekspressiooni taseme järgi, piisavaks loeti normaliseeritud geeniekspressiooni mediaani üle 51. Selles etapis filtreeriti välja kõige rohkem gene, ligikaudu 50% ehk 170 geeni (joonis 10). Välja jäänud geenidest üle poole (87 geeni) ei olnud minu andmestikus platsentas üldse ekspresseerunud. Ülejäänutel oli platsentas väga madal ekspressioon. Nende geenide ekspressioon ei olnud piisav platsentas verimise taseme määramiseks antud uuritavas materjalis ning vaja oleks suuremat valimit ja parema kattuvusega sekveneerimisandmeid.



Joonis 10. Graafikul on välja toodud 338 geeni logaritmiline normaliseeritud ekspressiooni mediaanide jaotus üle indiviide. Punaselt on tähistatud 51 lugemi piirjoon.

Filtreeringu tulemusi toetab ka Proteiinatlase andmebaas (joonis 11). Enamik gene ei ekspresseeru või on madala ekspressioonitasemega platsentas ning suure tõenäosusega selle tõttu ei suutnud ma nende geenide transkripte antud valimis tuvastada. Välja langes 81 mittekodeerivat RNA-d ja 60 geeni, mis ei ekspresseeru platsentas. Ekspressiooni ei ole tuvastatud 5 geenil. 24 geeni ekspresseerub platsentas ning 22 neist veel mitmes koes.



Joonis 11. Proteinaatlase andmebaasi põhjal välja filtreeritud geenide ekspressiooni asukoht.

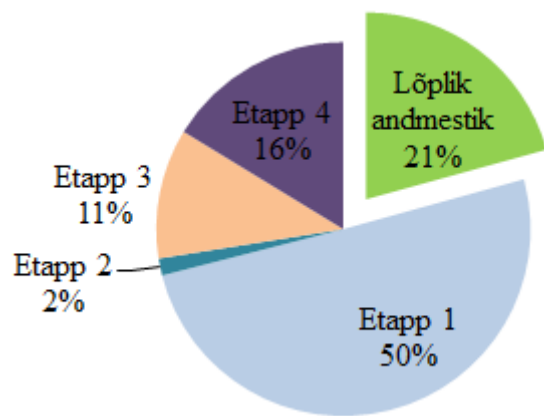
Teises etapis filtreeriti välja haruldasemad SNP-d, mille minoorse alleeli sagedus oli $\leq 10\%$. Viiel geenil (*COPG2*, *IFITM1*, *MCTS2P*, *NNAT*, *ERLIN2*) polnud kodeerivas alas piisavalt sagedasi SNP-sid, et selle põhjal saaks määrata vanemaspetsiifilist ekspressiooni.

Kolmandas andmetötluse etapis eemaldati SNP-d, mis puudusid DNA kiibilt ($n = 38$ geeni, 11% SNP-dest). Suure tõenäosusega puudusid need SNP-d DNA kiibilt, kuna olid LD-s mõne teise sagedase SNP-ga.

Neljandas etapis filtreeriti välja:

- geenid, millel oli ≤ 3 informatiivset pere ($n=4$ geeni);
- SNP-d, mis asusid intronites või kattuvates geenides ($n=99$ SNP-d; $n=13$ geeni);
- SNP-d, millel oli ≤ 3 toorlugemist/SNP kohta (131 SNP-d; $n = 38$ geeni).

Selles etapis filtreeriti välja 16% SNP-dest. Käesolevas magistritöös kasutatav lõplik andmestik koosnes 70 geenist ja 183 informatiivsest SNP-st, mis moodustas algandmestikust 21%. Igas etapis välja filtreeritud geenide protsentuaalset jaotust iseloomustab joonis 12.

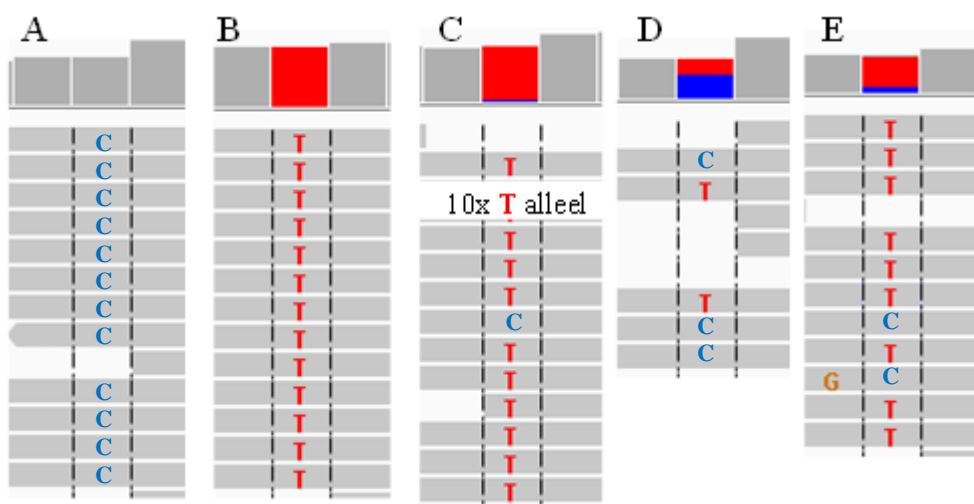


Joonis 12. Skeemil on graafiliselt kujutatud, kui suure osa geenidest iga etapp välja filtreeris.

2.3.2. Vermitud geenid

Käesolevas töös oli vermitud geenil $\geq 80\%$ lugemitest informatiivsete SNP positsioonides monoalleelsed ja pärit ühelt vanemalt. Kokku tuvastasin kindlalt ja usaldusväärselt platsentas 11 vermitud geeni 70-st (15,7%): *AIM1*, *DLK1*, *H19*, *IGF2*, *MEG3*, *MEST*, *PEG10*, *PHLDA2*, *PLAGL1*, *RTL1* ja *ZFAT* (tabel 6).

7 geeni 11-st on platsentaspetsiifilised, 13 SNP-d 25-st olid täieliku vermitusega. 12 SNP puhul esines taustaekspressiooni. Käesolevas magistritöös tuvastatud 10 vermitud geeni vastasid eelnevalt kirjanduses publitseeritule, välja arvatud *RTL1*, mis näitas vastupidist alleelset vanemlikku päritolu (tabel 6). *RTL1* geeni SNP rs6575805 puhul ekspresseerus 7 platsental emapoolne alleel ning ainult ühel isapoolne, 4-l platsental oli bialleelne ekspressioon. Näide erinevate indiviidide geeniekspressioonist samas SNP positsioonis kasutades Integrative Genomics Viewerit (joonis 13) (Robinson *et al.*, 2011).



Joonis 13. RNA-seq tulemus esitatud graafiliselt Integrative Genomics Vieweriga. Pildil *PHLDA2* SNP rs1056819 5-l informatiivsel indiviidil. A) Monoalleelne referentsalleeli ekspressioon (C nukleotiid); B) Monoalleelne alternatiivse alleeli ekspressioon (T nukleotiid); C) Monoalleelne ekspressioon väikese taustaekspressiooniga; D) Bialleelne ekspressioon E) Bialleelne ekspressioon alleelse kallutatusega.

11-st vermitud geenist 8-t on seostatud loote kasvuga (tabel 7). 11 vermitud geeni asusid 7-s genoomipiirkonnas ning kahes neist on mitu geeni (tabel X). *IGF2*, *PHLDA2* ja *H19* asuvad samas klastris 11. kromosoomil (joonis 13.A). *IGF2* on isapoolse ekspressiooniga, *PHLDA2* ja *H19* emapoolse avaldumisega. Samas klastris asub palju teisi kirjanduses raporteeritud vermitud gene, mis kahjuks jäid minu andmestikus andmeanalüüsist välja. Varasemalt publitseeritud andmete põhjal on isapoolse ekspressiooniga *IGF2AS*, *INS* ja *KCNQ1OT1*;

emapoolse ekspressiooniga *KCNQ1*, *KCNQ1DN*, *CDKN1C* ja *SLC22A18AS*; *TH* ja *NAP1L4* pole inimesel vermitud, aga hiirel on; *TSPAN32* on nii hiires kui inimeses bialleelne (Lisa 3).

DLK1, *RTL1* ja *MEG3* asuvad 14. kromosoomil ühes klastris (Joonis 13.B). Isapoolse ekspressiooniga on *DLK1* ja emapoolse ekspressiooniga *MEG8*. *RTL1* ekspresseeris meie andmestikus emapoolset alleeli, kirjanduses põhjal ekspresseerib isapoolset (tabel 6 ja 7). *RTL1* kõrval asetsev *MEG8* on emapoolse ekspressiooniga.

Tabel 6. Käesolevas magistritöös tuvastatud vermitud geenid platsentas RNA-seqi andmete põhjal.

Geen	SNP ID	chr	SNP posit- sioon	Inform. indiviidid	NGEM selles pos.-s	Lugemite päritolu (indiviidide arv)			Ekspr alleel	Proteinatlas RNA koe grupp
						Ainult emapoolsed	Ainult isa- poolsed	Bialleelsed		
<i>AIM1</i> ^A	rs1159148	6	Ekson	20	13,5	0 (0%)	18 (90%)	2 (10%)	Isa	Segagrupp
	rs3747787		Ekson	11	29	0 (0%)	8 (73%)	3 (27%)	Isa	
	rs4946759		5'UTR	15	10	0 (0%)	12 (80%)	3 (20%)	Isa	
<i>PLAGL1</i> ^A	rs9373409	6	UTR	18	25	0 (0%)	17 (94%)	1 (6%)	Isa	Kõikjal
<i>MEST</i> ^A	rs10863	7	3'UTR	11	46	0 (0%)	11 (100%)	0 (0%)	Isa	Platsenta
<i>PEG10</i> ^A	rs13073	7	3'UTR	16	117	0 (0%)	16 (100%)	0 (0%)	Isa	Neerupealised, platsenta
	rs13226637		3'UTR	12	136	1 (8%)	11 (92%)	0 (0%)	Isa	
	rs3750105		Ekson	12	166,5	1 (8%)	11 (92%)	0 (0%)	Isa	
<i>ZFAT</i> ^A	rs11538239	8	3'UTR	13	55,5	0 (0%)	13 (100%)	0 (0%)	Isa	Neerud, platsenta
	rs12541381		Ekson	18	138,5	0 (0%)	18 (100%)	0 (0%)	Isa	
	rs3739422		Ekson	12	165,5	0 (0%)	12 (100%)	0 (0%)	Isa	
	rs3739429		Ekson	7	108,5	0 (0%)	7 (100%)	0 (0%)	Isa	
<i>H19</i> ^A	rs217727	11	3'UTR	14	886,5	14 (100%)	0 (0%)	0 (0%)	Ema	A/P
<i>IGF2</i> ^A	rs2585	11	3'UTR	18	156,5	0 (0%)	18 (100%)	0 (0%)	Isa	Platsenta
	rs3802971		3'UTR	4	38,5	0 (0%)	4 (100%)	0 (0%)	Isa	
	rs7873		3'UTR	5	476,5	0 (0%)	5 (100%)	0 (0%)	Isa	
<i>PHLDA2</i> ^A	rs1056819	11	3' UTR	14	4	12 (86%)	0 (0%)	2 (14%)	Ema	Platsenta
<i>DLK1</i> ^A	rs1802710	14	Ekson	16	56,5	0 (0%)	16 (100%)	0 (0%)	Isa	Neerupealised, platsenta

Geen	SNP ID	chr	SNP posit-sioon	Inform. individid	NGEM selles pos.-s	Lugemite päritolu (indiviidide arv)			Ekspr alleel	Proteinatlas RNA koe grupp
						Ainult emapoolsed	Ainult isa-poolsed	Bialleelse d		
<i>MEG3</i> ^A	rs10132552	14	UTR	16	38	16 (100%)	0 (0%)	0%	Ema	A/P
	rs10147988		5'UTR	19	108,5	18 (95%)	0 (0%)	1 (5%)	Ema	
	rs1054013		3'UTR	18	12	17 (94%)	1 (6%)	0 (0%)	Ema	
	rs11160608		3'UTR	18	10	9 (50%)	9 (50%)	0 (0%)	RMAE	
	rs3742395		3'UTR	20	8,5	16 (80%)	4 (20%)	0 (0%)	Ema	
<i>RTL1</i> ^A	rs3825569	14	Ekson	13	13	13 (100%)	0 (0%)	0 (0%)	Ema*	Neerupealised, platsenta
	rs6575805		Ekson	12	8	7 (58%)	1 (8%)	4 (33%)	RMAE	

*– Varasemalt raporteeritud kui isapoolne ekspresiooni, A/P – Andmed puuduvad, ^A – Geneimprint andmebaasi järgi vermitud geenid, NGEM-normaliseeritud geeniekspresiooni mediaan, chr- kromosoom, ekspr- ekspresseerunu

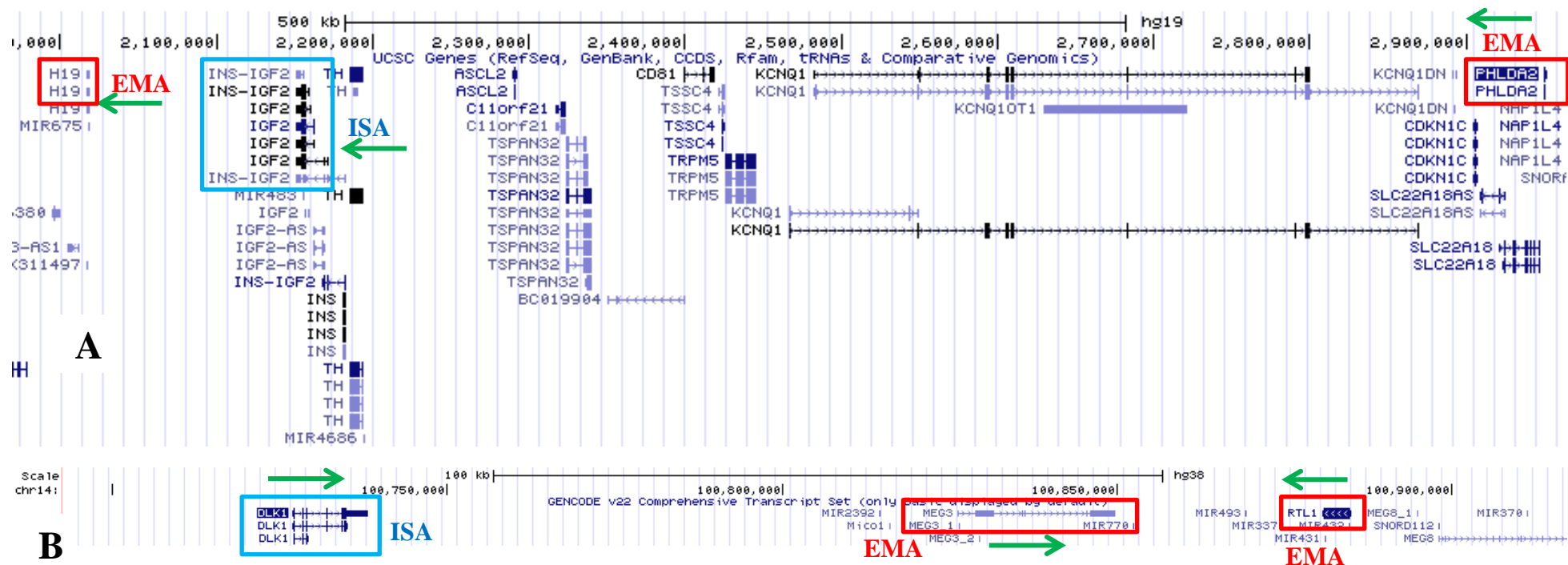
AIM1 geenil on kaks transkripti ning varasemalt on leitud, et vermitud on ainult pikem transkript ning lühem on bialleelne (Das et al., 2013). Käesolevas töös kasutatud SNP-d kuulusid pikemasse transkripti. Vermitud geenide üldiseloomustus on toodud tabelis 7.

Tabel 7. Vermitud geenide iseloomustus varasemalt publitseeritud teadusartiklite põhjal.

Geen	Positsioon	Ingliskeelne nimetus	Funktsioon	Seos raseduspatoloogiatega	Viide	Ekspr. alleel	Ekspr. alleel hiies
<i>AIM1</i>	6q21	<i>absent in melanoma 1</i>	raku morfoloogia	A/P	Vainio <i>et al.</i> , 2012	Isa	Bialleelne
<i>PLAGL1</i>	6q24-q25	<i>PLAG1 like zinc finger 1</i>	rakutsükli reguleerimine, tuumorsupressor	TNDM1 IUGR	Spengler <i>et al.</i> , 1997; Mackay ja Temple, 2010; Iglesias-Platas <i>et al.</i> , 2014; Huang <i>et al.</i> , 2001	Isa	Isa
<i>MEST</i>	7q32	<i>mesoderm specific transcript paternally expressed 10</i>	platsenta ja detsiidua angiogenees	GD IUGR	Mayer <i>et al.</i> , 2000; El Hajj <i>et al.</i> , 2013; McMinn <i>et al.</i> , 2006	Isa	Isa
<i>PEG10</i>	7q21	<i>paternally expressed 10</i>	antiapoptootiline faktor	IUGR preeklampsia spontaanabort, loote surm	Lux <i>et al.</i> , 2010; Diplas <i>et al.</i> , 2009; Liang <i>et al.</i> , 2014; Dória <i>et al.</i> , 2010	Isa	Isa
<i>ZFAT</i>	8q24.22	<i>zinc finger and AT-hook domain containing imprinted maternally expressed transcript</i>	transkriptsiooni regulaator	preeklampsia angiogenees, apoptoos	Koyanagi <i>et al.</i> , 2008; Barboux <i>et al.</i> , 2012; Yoshida <i>et al.</i> , 2010; Fujimoto <i>et al.</i> , 2009	Isa	Bialleelne
<i>H19</i>	11p15.5	<i>imprinted maternally expressed transcript</i>	glutatiooni regulatsioon vähi korral, kartsinogenees, trofoblasti proliferatsiooni regulatsioon	BWS PE SGA	Zheng <i>et al.</i> , 2016; Li <i>et al.</i> , 2014a; Jiang <i>et al.</i> , 2016; Weksberg <i>et al.</i> , 2010; Yu <i>et al.</i> , 2009; Koukoura <i>et al.</i> , 2011; Gao <i>et al.</i> , 2012	Ema	Ema

Geen	Positsioon	Inglis-keelne nimetus	Funktsioon	Seos rasedus- patoloogiatega	Viide	Ekspr. alleel	Ekspr. alleel hires
<i>IGF2</i>	11p15.5	<i>insulin like growth factor 2</i>	platsenta ja loote kasvufaktor	IUGR spontaanabort, loote surm BWS SRS	Constância <i>et al.</i> , 2002; McMinn <i>et al.</i> , 2006; Dória <i>et al.</i> , 2010; Weksberg <i>et al.</i> , 2010; Azzi <i>et al.</i> , 2014	Isa	Isa
<i>PHLDA2</i>	11p15.5	<i>pleckstrin homology like domain family A member 2</i>	trofoblasti proliferatsiooni regulatsioon	preeklampsia IUGR spontaanabort, loote surm BWS	Jin <i>et al.</i> , 2015; McMinn <i>et al.</i> , 2006; Dória <i>et al.</i> , 2010; De Crescenzo <i>et al.</i> , 2013	Ema	Ema
<i>DLK1</i>	14q32	<i>delta-like 1 homolog (Drosophila)</i>	adipogeneesi inhibiitor, Notch signaaliraja regulaator	GD SGA	Sul, 2009; Falix <i>et al.</i> , 2012; Li <i>et al.</i> , 2014b; Díaz <i>et al.</i> , 2013	Isa	Isa
<i>MEG3</i>	14q32	<i>maternally expressed 3</i>	tuumorsupressor, apoptoosi reguleerimine, p53 aktiveerimine	IUGR PE	Zhang <i>et al.</i> , 2003; McMinn <i>et al.</i> , 2006; Zhang <i>et al.</i> , 2015; Zhou <i>et al.</i> , 2007	Ema	Ema
<i>RTL1</i>	14q32	<i>retrotransposon-like 1</i>	loote kapillaaride alalhoidmine [#] , kartsinogenees [#]	loote surm [#] , neonataalne letaalsus [#]	Sekita <i>et al.</i> , 2008; Riordan <i>et al.</i> , 2013	Isa	Isa

ER- endoplasmaatiline retiikulum; A/P- andmed puuduvad; IUGR- üsasisene kasvupeetus; TNDM1- mööduv neonataalne tüüp 1 diabeet; SGA- gestatsiooniaja kohta väike laps; BWS - Beckwith-Wiedemanni sündroom; PWS- Prader-Willi sündroom ; PE- preeklampsia; [#] - näidatud ainult hiire mudelil; GD- gestatsioonidiabeet



Joonis 13. Vermitud geenide klastrid. A) 11. Kromosoomil paiknevad H19 ja IGF2 asukoht üksteise suhtes. IGF2 kõrval on näha INS, TH ja IGF2-AS, mis filtreerimise käigus välja langesid. B) 14. Kromosoomil DLK1, MEG3 ja RTL1 paiknemine. Paremäl ääres paistab samas klastris ka MEG8, mis kahjuks filtreerimisel välja langes. Rohelise noolega märgitud transkriptsiooni suund. Värviliselt märgitud vanemaspetsiifiline ekspressioon meie andmetes.

2.3.2.1. Kõrvalekalded vermisest

Nelja vermitud geeni puhul on erandite seas üks indiviid II trimestri raseduste seas. Kõrvalekallete seas on lisaks üks gestatsioonidiabeediga platsenta kolmel korral ning kaks makrosoomiaga platsentat kahel juhul (tabel 8).

Tabel 8. Indiviidid vastupidise vermitusega

Geen	Rsnumber	Erandid Arv	Esindatud alagrupid
<i>AIM1</i>	rs1159148	2	GD, MNA
	rs3747787	3	PE, PE, GD
	rs4946759	3	NN, SGA, SGA
<i>PLAGL1</i>	rs9373409	1	LGA
<i>PEG10</i>	rs13226637	1	MNA ¹
	rs3750105	1	MNA ¹
<i>PHLDA2</i>	rs1056819	2	GD ² , MNA ¹
<i>MEG3</i>	rs10147988	1	PE
	rs1054013	1	LGA ³
	rs11160608	9	SGA, LGA ³ , LGA ⁴ , LGA, GD ² , MNA, MNA, MNA ¹ , OSA
	rs3742395	4	SGA, LGA ⁴ , LGA, GD ²

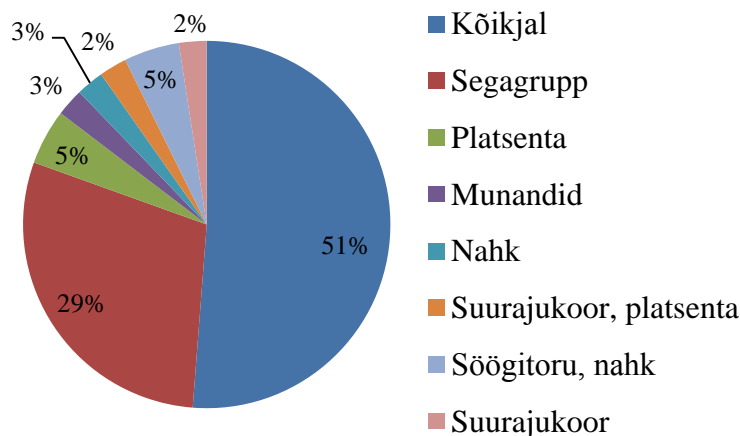
^{1/2/3/4} - sarnaselt on tähistatud sama indiviid

Huvitavaks leiuks oli makrosoomia grupi platsenta, millel *PLAGL1* ekspresseerus bialleelselt, mis kirjanduse põhjal põhjustab TNDM1 haigust. Kuid emal ei tuvastatud raseduse jooksul gestatsioonidiabeeti ning lapsel ei esinenud neonataalselt diabeeti ning seda ei ole tekkinud lapse 2 aasta vanuseni. Käesolevas magistritöös oli emapoolse päritoluga kõrge kvaliteediga alleele 2 ja isapoolseid 7, mis viitab alleelsele kallutatusele bialleelse ekspressiooni puhul.

2.3.3. Geenid, mis ei ole platsentas vermitud

2.3.3.1. Bialleelse ekspressiooniga geenid

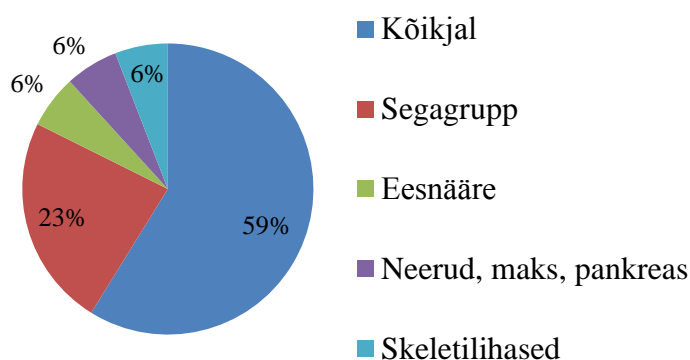
Käesolevas magistritöös oli bialleelsel geenil $\geq 40\%$ -l platsentadest informatiivsete SNP-de positsioonid bialleelsed. Sellele kriteeriumile vastas 41 geeni 70-st (59%) (Lisa 2). Bialleelsete geenide grupi normaliseeritud geeniekspressiooni mediaan oli 18 lugemit SNP kohta. 21 (51%) bialleelset geeni ekspresseerus kõigis kudedes, platsentas ekspresseerus 3 geeni (joonis 14). Varasemalt on näidatud, et *SLC22A3* on I trimestri spetsiifiline vermitud geen (Monk et al., 2006). Kahjuks ei olnud minu valimi I trimestris informatiivseid indiviide ning ma ei saanud seda kinnitada.



Joonis 14. Käesolevas magistritöös tuvastatud bialleelsete geenide ekspressiooni asukoht Proteinatlase andmebaasi järgi.

2.3.3.2. RMAE

Geen loeti RMAE-ks, kui tal oli $\leq 40\%$ platsentadest olid bialleelsed, ülejäänud olid monoalleelsed, aga mõlema vanema alleelid olid esindatud. RMAE gene oli 3 70-st (4%), RMAE või bialleelse ekspressiooni vahepealseid gene oli 14 70-st (20%). *PAPPA2* jäi grupeerimata, kuna tulemused olid konfliktised (Lisa 2). RMAE ja bialleelne/ RMAE gruppi kuuluvatel geenidel oli üldiselt madalam ekspressioon ning NGEM SNP positsioonis jäi enamusel geenidest alla 10 lugemi (mediaan 4). *RHOBTB3* ja *PAPPA2*, millel oli kõrge ekspressioon, on esindatud mõlemad alleelid tugeva kallutatusega. RMAE geenid ei ekspresseerunud spetsiifiliselt platsentas, 10 geeni (59%) ekspresseerus kõikjal (joonis 15).



Joonis 15. Käesolevas magistritöös tuvastatud RMAE geenide ekspressiooni asukoht Proteinatlase andmebaasi järgi.

2.4. Arutelu

2.4.1. Käesoleva magistritöö kattuvus kirjanduses raporteeritud andmetega

Käesolevas magistritöös lugesin geeni vermituks siis, kui $\geq 80\%$ -l platsentadest olid informatiivsete SNP-de positsioonid monoalleelselt ekspresseeritud ja kõigil indiviididel oli avaldunud alleeli päritolu samalt vanemalt. Kahjuks ei eksisteeri teaduskirjanduses ühest definitsiooni transkriptomide andmete põhjal geenide vermituse staatuse määramise kohta ning loetud teadusartiklites oli kasutatud erinevat definitsiooni ja filtreerimise kriteeriume. Käesolevas töös sain anda hinnangu ainult platsenta vermitusele, mitte loote üsasisesse perioodi või neonataalse vermituse kohta.

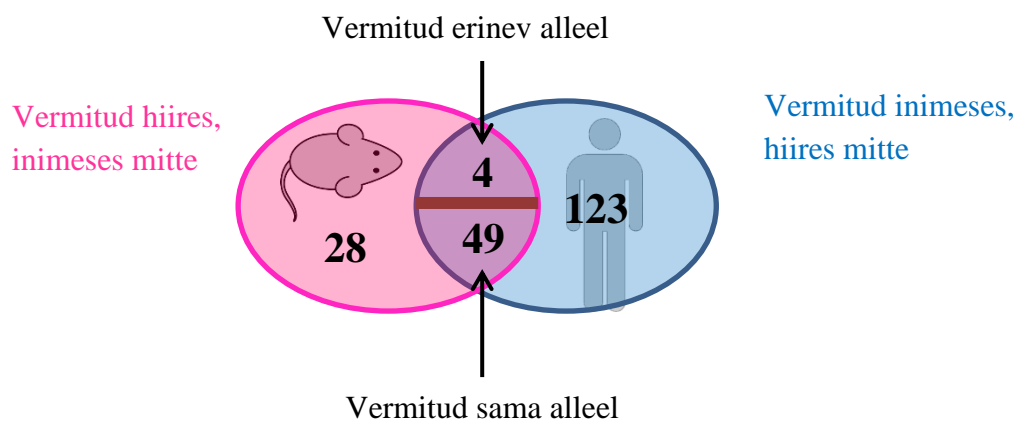
2.4.1.1. Andmete võrdlus Geneimprint andmebaasiga

Antud töös kinnitati 11 vermitud geeni platsentas, sealhulgas 10 vermitud geeni puhul langes avaldunud alleeli vanemlik päritolu kokku kirjanduses avaldatud andmetega. Erandina tuvastasime meie andmetes geeni *RTL1*, mille puhul tuvastasime ühe analüüsitud SNP positsiooni kohta emapoolse alleeli ekspressiooni, kuigi kirjanduses on andmed isapoolse alleeli avaldumise kohta. *RTL1* geeni teise SNP alleelse ekspressiooni analüüsimisel saadud andmed näitasid pigem RMAE stsenaariumit. *RTL1* vermituse staatust ja avaldunud alleelide vanemlikku päritolu on plaanis uurida edasi lookuspõhise lähenemisega.

Lisaks langes kokku Geneimprint andmebaasi andmetega *AMPD3*, *ASCL2* ja *DHCR7* bialleelse ekspressiooni tuvastamine.

Minu uurimistöö andmed 25 geeni kohta ei vastanud andmebaasis olevale vermituse staatusele ning 245 geeni jäid põhiliselt olematu või madala ekspressiooni tõttu analüüsist välja. Vähene kattuvus võib olla põhjustatud organispetsiifilisest vermitusest, kuna platsentas ei ekspresseerugi kõik kirjanduses raporteeritud vermitud geenid (joonis 11).

Kuna hiire mudelil on vermitust palju uuritud, on hiire vermitud geenid hästi teada. Inimese ja hiire geenide vermitus on sarnase mustriga (Wilkins *et al.*, 2016), kuid Geneimprint andmebaasis järgi on 204 inimese ja hiire ortoloogist ainult 49 mõlemal vermitud ja sama vanema spetsiifilise alleelse ekspressiooniga (Joonis 16; Lisa 3).



Joonis 16. Geneimprint andmebaasi põhjal inimese vermitud geenide ülekate hiire ortoloogidega.

2.4.1.2. Andmete võrdlus hiljuti avaldatud teaduskirjandusega

Metsalu *et al.* teadustöös oli vermitud geeni definitsioon erinev, vaadati alleelspetsiifilise ekspressiooni osakaalu geeni kohta heterosügootsetest SNP-dest. Tuvastati RNA sekveneerimise meetodil neli vermitud geeni platsentas: *AIM1*, *PEG10*, *RHOBTB3* ja *ZFAT-ASI* (Metsalu *et al.*, 2014). Käesolevas magistritöös kinnitasin platsentas *AIM1* ja *PEG10* vermitust. *ZFAT-ASI* langes kahjuks filtreerimise käigus välja ning *RHOBTB3*-l ekspresseeruvad siiski mõlemad alleelid ning esineb bialleelset ekspressiooni. Oma töös ei kinnitanud ma Metsalu *et al.* potentsiaalseid uusi vermitud gene (*LGALS8*, *LGALS14*, *PAPPA2* ja *SPTLC3*) ega teisi vermitud kandidaatgene alleelspetsiifilise ekspressiooni kallutatusega.

Hiljuti avaldatud Court *et al.* teadustöös määrati vermitud geenid platsentas DNA metülatsiooni analüüsiga. Kahjuks ei kinnitanud ma oma magistritöös Court *et al.* uusi potentsiaalselt vermitud gene, küll aga kinnitusid varasemalt teadaolevad vermitud geenid (*AIM1*, *PLAGL1*, *MEST*, *PEG10*, *ZFAT*, *H19* ja *IGF2*). (Court *et al.*, 2014)

2.4.2. RNA sekveneerimise tugevad ja nõrgad küljed

RNA sekveneerimine on suure statistilise võimsusega meetod vermituse tuvastamiseks, kuid ka sellel meetodil on omad puudused. Süstemaatiline viga võib tekkida praimerite seondumisel, fragmentatsioonil, PCR-i amplifikatsiooniprotsessil, raamatukogude tegemisel, sekveneerimisel ja joondamisel (DeVeale *et al.*, 2012).

Käesoleva magistritöö üheks piiranguks oli sekveneerimise sügavus, eriti madala ekspressiooniga geenide puhul, et otsustada vermitust positsiooni kohta. Liiga väheste lugemite korral on juhuslikkusel suur roll nähtavas jaotuses.

RNA sekveneerimisandmeid kasutades on võimalik eristada RMAE-d ning tõelist vermitust, samas vajab see informatiivsete positsioonide olemasolu. Madala sagedusega SNP-de puhul on vaja suuremat valimit, et leida piisavalt selliseid indiviide. Käesolevas töös ei vaadatud transkriptoomi ühe raku tasemel, seega ei saa eristada RMAE-d bialleelsest ekspressioonist. Monoalleelne ekspressioon RMAE geenide puhul võib olla tingitud madalast ekspressioonitasemest, kus teise alleeli lugemid on jäänud detekteerimata.

Samuti ei saa informatsiooni geenide metüleerituse kohta, mida loetakse oluliseks komponendiks geenide vermituse määramisel.

Tulevikus on plaanis valideerida *RTL1* vermitust, kuna meie valimis on tal vastupidine tulemus kirjanduses publitseeritule, ning *MEG3*, kuna üks SNP on juhusliku monoalleelse ekspressiooniga (RMAE) ning ei ühti teiste sama geeni SNP-de vermitusega.

KOKKUVÕTE

Käesolevas magistritöös anti kirjanduse alusel ülevaade geenide monoalleelsest ekspressioonist, dünaamilise ja fikseeritud RMAE erinevustest, vermituse häiretest ning inimese vermitud geenide seosest rasedustüsistustega.

Antud töö eesmärgiks oli määrata platsentas avaldunud ja kirjanduses raporteeritud potentsiaalselt vermitud geenide transkriptide vanemaspetsiifiline päritolu ja tuvastada vermitud geenide ekspressiooniolekut läbi raseduse kolme trimestri ning rasedustüsistuste puhul.

Magistritöö lähtematerjalideks oli kaks ülegenoomset analüüsi, mis kombineerides moodustasid tervikandmestiku ema-isa-platsenta triod/duo kohta:

1. 55 inimese platsenta koeproovide transkriptoomid (RNA-seq)
2. Ülegenoomsed genotüübid koos uuritavate perede (isa ja ema) veredest eraldatud DNA variantide andmetega.

Rangete andmete kvaliteedi- ja piisavuse kriteeriumite järgi filtreerimise tulemusena sai vanemlikku päritolu usaldusväärset kinnitada või ümber lükata 338 geenist 70-l, kuna ainult pooled ainult pooled kirjanduses raporteeritud vermitud geenidest on avaldunud platsentas.

Kasutades rangeid kriteeriume ja mitut filtreerimise etappi, tuvastasin platsentas 11 vermitud geeni: *AIM1*, *DLK1*, *H19*, *IGF2*, *MEG3*, *MEST*, *PEG10*, *PHLDA2*, *PLAGL1*, *RTL1* ja *ZFAT*. Käesolevas magistritöös tuvastatud 10 vermitud geeni vastasid eelnevalt kirjanduses publitseeritule, välja arvatud *RTL1*, mis näitas vastupidist alleelset vanemlikku päritolu

Käesolevas töös tuvastasin 41 bialleelset geeni, 3 RMAE geeni ning RMAE või bialleelse ekspressiooni vahepealseid gene oli 14 70-st. *PAPPA2* jäi grupeerimata, kuna tulemused olid konfliktised.

Tulevikus on plaanis valideerida *RTL1* vermitust, kuna meie valimis on tal vastupidine tulemus kirjanduses publitseeritule, ning *MEG3*, kuna üks SNP on juhusliku monoalleelse ekspressiooniga (RMAE) ning ei ühti teiste sama geeni SNP-de vermitusega.

Monoallelic parent-of-origin gene expression in human placenta

Diana Nõmmemees

SUMMARY

Diploid organisms inherit one gene copy from each parent and genes can be expressed from both alleles (biallelic expression) or only one allele (monoallelic expression). Imprinted genes are exclusively expressed from either the paternally or maternally derived allele and the other allele is silenced via epigenetic mechanisms.

A number of imprinted genes express in placenta, where they regulate placental and fetal growth. Disrupted imprinting can lead to pregnancy complications.

The aims of this study were:

1. To determine imprinted gene transcript allele specific expression in placenta based on previously published articles and Geneimprint database
2. To identify imprinted gene expression in three pregnancy trimesters and in pregnancy complications

In this study we identified 11 imprinted genes: *AIM1*, *DLK1*, *H19*, *IGF2*, *MEG3*, *MEST*, *PEG10*, *PHLDA2*, *PLAGL1*, *RTL1* ja *ZFAT*. 10 genes validated previously reported parent specific allelic expression, but contrary to previously described paternal expression *RTL1* exhibited maternal expression in our study. We also detected 41 genes with biallelic expression and 3 with random monoallelic expression. 14 genes showed characteristics similar to random monoallelic expression and biallelic expression.

In future we plan to validate maternal allele specific expression of *RTL1* and *MEG3* because one of *MEG3* SNP-s showed controversial results.

Kasutatud kirjanduse loetelu

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169.

Azzi, S., Abi Habib, W., and Netchine, I. (2014). Beckwith-Wiedemann and Russell-Silver Syndromes: from new molecular insights to the comprehension of imprinting regulation. *Curr. Opin. Endocrinol. Diabetes. Obes.* *21*, 30–38.

Bamberg, C., and Kalache, K.D. (2004). Prenatal diagnosis of fetal growth restriction. *Semin. Fetal Neonatal Med.* *9*, 387–394.

Barboux, S., Gascoïn-Lachambre, G., Buffat, C., Monnier, P., Mondon, F., Tonanny, M.-B., Pinard, A., Auer, J., Bessières, B., Barlier, A., et al. (2012). A genome-wide approach reveals novel imprinted genes expressed in the human placenta. *Epigenetics* *7*, 1079–1090.

Constância, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., Stewart, F., Kelsey, G., Fowden, A., Sibley, C., et al. (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* *417*, 945–948.

Court, F., Tayama, C., Romanelli, V., Martin-Trujillo, A., Iglesias-Platas, I., Okamura, K., Sugahara, N., Simón, C., Moore, H., Harness, J. V., et al. (2014). 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.

De Crescenzo, A., Sparago, A., Cerrato, F., Palumbo, O., Carella, M., Miceli, M., Bronshtein, M., Riccio, A., and Yaron, Y. (2013). Paternal deletion of the 11p15.5 centromeric-imprinting control region is associated with alteration of imprinted gene expression and recurrent severe intrauterine growth restriction. *J. Med. Genet.* *50*, 99–103.

Das, R., Lee, Y.K., Strogantsev, R., Jin, S., Lim, Y.C., Ng, P.Y., Lin, X.M., Chng, K., Yeo, G.S., Ferguson-Smith, A.C., et al. (2013). DNMT1 and AIM1 Imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. *BMC Genomics* *14*, 685.

DeVeale, B., van der Kooy, D., and Babak, T. (2012). Critical evaluation of imprinted gene expression by RNA-Seq: a new perspective. *PLoS Genet.* *8*, e1002600.

- Díaz, M., Bassols, J., Aragonés, G., Mazarico, E., López-Bermejo, A., and Ibáñez, L. (2013). Decreased placental expression of pre-adipocyte factor-1 in children born small-for-gestational-age: association to early postnatal weight gain. *Placenta* 34, 331–334.
- Diplas, A.I., Lambertini, L., Lee, M.-J., Sperling, R., Lee, Y.L., Wetmur, J., and Chen, J. (2009). Differential expression of imprinted genes in normal and IUGR human placentas. *Epigenetics* 4, 235–240.
- Dória, S., Sousa, M., Fernandes, S., Ramalho, C., Brandão, O., Matias, A., Barros, A., and Carvalho, F. (2010). Gene expression pattern of IGF2, PHLDA2, PEG10 and CDKN1C imprinted genes in spontaneous miscarriages or fetal deaths. *Epigenetics* 5, 444–450.
- English, F.A., Kenny, L.C., and McCarthy, F.P. (2015). Risk factors and effective management of preeclampsia. *Integr. Blood Press. Control* 8, 7–12.
- Falix, F.A., Aronson, D.C., Lamers, W.H., and Gaemers, I.C. (2012). Possible roles of DLK1 in the Notch pathway during development and disease. *Biochim. Biophys. Acta* 1822, 988–995.
- Flanagan, S.E., Patch, A.-M., Mackay, D.J.G., Edghill, E.L., Gloyn, A.L., Robinson, D., Shield, J.P.H., Temple, K., Ellard, S., and Hattersley, A.T. (2007). Mutations in ATP-sensitive K⁺ channel genes cause transient neonatal diabetes and permanent diabetes in childhood or adulthood. *Diabetes* 56, 1930–1937.
- Fujimoto, T., Doi, K., Koyanagi, M., Tsunoda, T., Takashima, Y., Yoshida, Y., Sasazuki, T., and Shirasawa, S. (2009). ZFAT is an antiapoptotic molecule and critical for cell survival in MOLT-4 cells. *FEBS Lett.* 583, 568–572.
- Gao, W.-L., Liu, M., Yang, Y., Yang, H., Liao, Q., Bai, Y., Li, Y.-X., Li, D., Peng, C., and Wang, Y.-L. (2012). The imprinted H19 gene regulates human placental trophoblast cell proliferation via encoding miR-675 that targets Nodal Modulator 1 (NOMO1). *RNA Biol.* 9, 1002–1010.
- Gregg, C., Zhang, J., Butler, J.E., Haig, D., and Dulac, C. (2010a). Sex-specific parent-of-origin allelic expression in the mouse brain. *Science* 329, 682–685.
- Gregg, C., Zhang, J., Weissbourd, B., Luo, S., Schroth, G.P., Haig, D., and Dulac, C. (2010b). High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* 329, 643–648.

- El Hajj, N., Pliushch, G., Schneider, E., Dittrich, M., Müller, T., Korenkov, M., Aretz, M., Zechner, U., Lehnen, H., and Haaf, T. (2013). Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes* 62, 1320–1328.
- Huang, S.M., Schönthal, A.H., and Stallcup, M.R. (2001). Enhancement of p53-dependent gene activation by the transcriptional coactivator Zac1. *Oncogene* 20, 2134–2143.
- Iglesias-Platas, I., Martin-Trujillo, A., Petazzi, P., Guillaumet-Adkins, A., Esteller, M., and Monk, D. (2014). Altered expression of the imprinted transcription factor PLAGL1 deregulates a network of genes in the human IUGR placenta. *Hum. Mol. Genet.* 23, 6275–6285.
- Jiang, X., Yan, Y., Hu, M., Chen, X., Wang, Y., Dai, Y., Wu, D., Wang, Y., Zhuang, Z., and Xia, H. (2016). Increased level of H19 long noncoding RNA promotes invasion, angiogenesis, and stemness of glioblastoma cells. *J. Neurosurg.* 124, 129–136.
- Jin, F., Qiao, C., Luan, N., and Shang, T. (2015). The expression of the imprinted gene pleckstrin homology-like domain family A member 2 in placental tissues of preeclampsia and its effects on the proliferation, migration and invasion of trophoblast cells JEG-3. *Clin. Exp. Pharmacol. Physiol.* 42, 1142–1151.
- 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.
- Kasak, L., Rull, K., Vaas, P., Teesalu, P., and Laan, M. (2015). Extensive load of somatic CNVs in the human placenta. *Sci. Rep.* 5, 8342.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- Kobayashi, H. (2014). Imprinting genes associated with endometriosis. *EXCLI J.* 13, 252–264.
- Koukoura, O., Sifakis, S., Zaravinos, A., Apostolidou, S., Jones, A., Hajiioannou, J., Widschwendter, M., and Spandidos, D.A. (2011). Hypomethylation along with increased H19

expression in placentas from pregnancies complicated with fetal growth restriction. *Placenta* 32, 51–57.

Koyanagi, M., Nakabayashi, K., Fujimoto, T., Gu, N., Baba, I., Takashima, Y., Doi, K., Harada, H., Kato, N., Sasazuki, T., et al. (2008). ZFAT expression in B and T lymphocytes and identification of ZFAT-regulated genes. *Genomics* 91, 451–457.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Li, H., Yu, B., Li, J., Su, L., Yan, M., Zhu, Z., and Liu, B. (2014a). Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget* 5, 2318–2329.

Li, J., Zhu, Q., Wang, H., Han, C., Zhou, Q., Huang, H., and Dong, M. (2014b). Decreased fetal pre-adipocyte factor-1 in pregnancies complicated by gestational diabetes mellitus. *Clin. Chim. Acta.* 431, 93–95.

Liang, X.Y., Chen, X., Jin, Y.Z., Chen, X.O., and Chen, Q.Z. (2014). Expression and significance of the imprinted gene PEG10 in placenta of patients with preeclampsia. *Genet. Mol. Res.* 13, 10607–10614.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.

Lux, H., Flammann, H., Hafner, M., and Lux, A. (2010). Genetic and molecular analyses of PEG10 reveal new aspects of genomic organization, transcription and translation. *PLoS One* 5, e8686.

Mackay, D.J.G., and Temple, I.K. (2010). Transient neonatal diabetes mellitus type 1. *Am. J. Med. Genet. C. Semin. Med. Genet.* 154C, 335–342.

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.

Mayer, W., Hemberger, M., Frank, H.G., Grümmer, R., Winterhager, E., Kaufmann, P., and Fundele, R. (2000). Expression of the imprinted genes MEST/Mest in human and murine placenta suggests a role in angiogenesis. *Dev. Dyn.* 217, 1–10.

- McMinn, J., Wei, M., Schupf, N., Cusmai, J., Johnson, E.B., Smith, A.C., Weksberg, R., Thaker, H.M., and Tycko, B. (2006). Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta* 27, 540–549.
- Metsalu, T., Viltrop, T., Tiirats, A., Rajashekar, B., Reimann, E., Kõks, S., Rull, K., Milani, L., Acharya, G., Basnet, P., et al. (2014). Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics* 9, 1397–1409.
- 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. U. S. A.* 103, 6623–6628.
- Peters, J. (2014). The role of genomic imprinting in biology and disease: an expanding view. *Nat. Rev. Genet.* 15, 517–530.
- Pinter, S.F., Colognori, D., Beliveau, B.J., Sadreyev, R.I., Payer, B., Yildirim, E., Wu, C.-T., and Lee, J.T. (2015). Allelic Imbalance Is a Prevalent and Tissue-Specific Feature of the Mouse Transcriptome. *Genetics* 200, 537–549.
- R Development Core Team (2011), R: A Language and Environment for Statistical Computing. Vienna, Austria : the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at <http://www.R-project.org/>.
- Reinius, B., and Sandberg, R. (2015). Random monoallelic expression of autosomal genes: stochastic transcription and allele-level regulation. *Nat. Rev. Genet.* 16, 653–664.
- Riordan, J.D., Keng, V.W., Tschida, B.R., Scheetz, T.E., Bell, J.B., Podetz-Pedersen, K.M., Moser, C.D., Copeland, N.G., Jenkins, N.A., Roberts, L.R., et al. (2013). Identification of *rtl1*, a retrotransposon-derived imprinted gene, as a novel driver of hepatocarcinogenesis. *PLoS Genet.* 9, e1003441.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26.
- Rull, K. (2009). Human chorionic gonadotropin beta genes and recurrent miscarriage: expression and variation study. Tartu University Press.
- Sekita, Y., Wagatsuma, H., Nakamura, K., Ono, R., Kagami, M., Wakisaka, N., Hino, T., Suzuki-Migishima, R., Kohda, T., Ogura, A., et al. (2008). Role of retrotransposon-derived imprinted gene, *Rtl1*, in the feto-maternal interface of mouse placenta. *Nat. Genet.* 40, 243–

248.

Shield, J.P.H., Temple, I.K., Sabin, M., Mackay, D., Robinson, D.O., Betts, P.R., Carson, D.J., Cavé, H., Chevenne, D., and Polak, M. (2004). An assessment of pancreatic endocrine function and insulin sensitivity in patients with transient neonatal diabetes in remission. *Arch. Dis. Child. Fetal Neonatal Ed.* 89, F341–F343.

Sildver, K., Veerus, P., and Lang, K. (2015). Sünnikaalukõverad Eestis ja sünnikaalu mõjutavad tegurid : registripõhine uuring. *Eesti Arst* 94, 465–470.

Spengler, D., Villalba, M., Hoffmann, A., Pantaloni, C., Houssami, S., Bockaert, J., and Journot, L. (1997). Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc finger protein expressed in the pituitary gland and the brain. *EMBO J.* 16, 2814–2825.

Sul, H.S. (2009). Minireview: Pref-1: role in adipogenesis and mesenchymal cell fate. *Mol. Endocrinol.* 23, 1717–1725.

Söber, S., Reiman, M., Kikas, T., Rull, K., Inno, R., Vaas, P., Teesalu, P., Marti, J.M.L., Mattila, P., and Laan, M. (2015). Extensive shift in placental transcriptome profile in preeclampsia and placental origin of adverse pregnancy outcomes. *Sci. Rep.* 5, 13336.

Zhang, X., Zhou, Y., Mehta, K.R., Danila, D.C., Scolavino, S., Johnson, S.R., and Klibanski, A. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J. Clin. Endocrinol. Metab.* 88, 5119–5126.

Zhang, Y., Zou, Y., Wang, W., Zuo, Q., Jiang, Z., Sun, M., De, W., and Sun, L. (2015). Down-regulated long non-coding RNA MEG3 and its effect on promoting apoptosis and suppressing migration of trophoblast cells. *J. Cell. Biochem.* 116, 542–550.

Zheng, Z.-G., Xu, H., Suo, S.-S., Xu, X.-L., Ni, M.-W., Gu, L.-H., Chen, W., Wang, L.-Y., Zhao, Y., Tian, B., et al. (2016). The Essential Role of H19 Contributing to Cisplatin Resistance by Regulating Glutathione Metabolism in High-Grade Serous Ovarian Cancer. *Sci. Rep.* 6, 26093.

Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D.L., Gejman, R., Ansell, P.J., Zhao, J., Weng, C., and Klibanski, A. (2007). Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem.* 282, 24731–24742.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq

reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.

Vainio, P., Mpindi, J.-P., Kohonen, P., Fey, V., Mirtti, T., Alanen, K.A., Perälä, M., Kallioniemi, O., and Iljin, K. (2012). High-throughput transcriptomic and RNAi analysis identifies AIM1, ERGIC1, TMED3 and TPX2 as potential drug targets in prostate cancer. *PLoS One* 7, e39801.

Weksberg, R., Shuman, C., and Beckwith, J.B. (2010). Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* 18, 8–14.

Wilkins, J.F., Úbeda, F., and Van Cleve, J. (2016). The evolving landscape of imprinted genes in humans and mice: Conflict among alleles, genes, tissues, and kin. *Bioessays* 38, 482–489.

Yoshida, Y., Tsunoda, T., Takashima, Y., Fujimoto, T., Doi, K., Sasazuki, T., Kuroki, M., Iwasaki, A., and Shirasawa, S. (2010). ZFAT is essential for endothelial cell assembly and the branch point formation of capillary-like structures in an angiogenesis model. *Cell. Mol. Biol. Lett.* 15, 541–550.

Yu, L., Chen, M., Zhao, D., Yi, P., Lu, L., Han, J., Zheng, X., Zhou, Y., and Li, L. (2009). The H19 gene imprinting in normal pregnancy and pre-eclampsia. *Placenta* 30, 443–447.

Kasutatud veebiaadressid

1. FIMM <https://www.fimm.fi/>
2. <http://www.geenivaramu.ee/et>
3. <http://www.geneimprint.com/site/genes-by-species>
4. <https://www.python.org/>

LISAD

Lisa 1.

Tabel 9. Käesolevas magistritöös uuritud potentsiaalselt vermitud geenide loetelu.

Plantsentas potentsiaalselt vermitud geenide nimekiri					
Geneimprint andebaasi potentsiaalselt vermitud geenide nimekiri*					Allikad**
ABCC9	GFI1	NDN	SNORD115-3	TH	AGBL3
ABCG8	GLI3	NKAIN3	SNORD115-30	TIGD1	CXORF56
ACD	GLIS3	NKX6-2	SNORD115-31	TMEM255B	DCAF10
ADAMTS16	GNAS	NLRP2	SNORD115-32	TMEM52	DOCK1
AIM1	GNAS-AS1	NNAT	SNORD115-33	TMEM60	ERLIN2
ALDH1L1	GPR1	NPAP1	SNORD115-35	TMEM88	FAM196A
AMPD3	GPT	NTM	SNORD115-37	TP73	FAM20A
ANO1	GRB10	OBSCN	SNORD115-38	TRPM5	GPR1-AS
APBA1	H19	OR11L1	SNORD115-39	TSHZ3	HTR5A
ASB4	HES1	OSBPL5	SNORD115-4	TSIX	IGF1R
ASCL2	HIST3H2BB	OTX1	SNORD115-40	TSPAN32	MCCC1
ATP10A	HM13	PAOX	SNORD115-41	TSSC4	MIR4508
B4GALNT4	HOXA11	PEG10	SNORD115-44	TTC40	MIR512-1
BLCAP	HOXA2	PEG3	SNORD115-45	UBE3A	N4BP2L1
BMP8B	HOXA3	PEX10	SNORD115-47	USB1	NHP2L1
BTNL2	HOXA4	PHLDA2	SNORD115-48	USP29	PDE4D
C10orf91	HOXA5	PHPT1	SNORD115-5	UVSSA	PPIEL
C9orf116	HOXB2	PKP3	SNORD115-6	VAX2	RGMA
C9orf85	HOXB3	PLAGL1	SNORD115-8	VENTX	ZNF331
CALCR	HOXC4	PON1	SNORD115-9	WRAP73	ZNF396
CCDC85A	HOXC9	PON2	SNORD116-1	WT1	TRAPPC9
CD81	HSPA6	PON3	SNORD116-10	XIST	WDR27
CDH18	HTR2A	PPAP2C	SNORD116-11	AKAP7	WRB
CDK4	HYMAI	PPP1R9A	SNORD116-12	ANLN	AKAP7
CDKN1C	IFITM1	PRDM16	SNORD116-13	AOC1	ANLN
CELF4	IGF2	PRIM2	SNORD116-15	BCLAF1	AOC1
CHMP2A	IGF2AS	PTPN14	SNORD116-16	CASC5	BCLAF1
CHST8	IGF2R	PURG	SNORD116-18	CEP170	CASC5
COL9A3	IMPACT	PYY2	SNORD116-19	CEP63	CEP170
COMMD1	INPP5F	RAB1B	SNORD116-2	CHD1	CEP63
COPG2	INS	RASGRF1	SNORD116-20	CLASP1	CHD1
CPA4	ISM1	RB1	SNORD116-23	DLG5	CLASP1
CSF2	KBTBD3	RBP5	SNORD116-24	ELMO2	DLG5
CTNNA3	KCNK9	RNU5D-1	SNORD116-25	GRHL1	ELMO2
CYP1B1	KCNQ1	RPL22	SNORD116-26	HEG1	GRHL1
DCN	KCNQ1DN	RTL1	SNORD116-29	HLA-DRB1	HEG1
DDC	KCNQ1OT1	SALL1	SNORD116-3	IFI30	HLA-DRB1
DGCR6	KLF14	SDHD	SNORD116-5	KIAA1191	IFI30
DGCR6L	L3MBTL1	SFMBT2	SNORD116-8	KIAA1551	KIAA1191
DHCR7	LDB1	SGCE	SNORD116-4	KIAA1919	KIAA1551

DIO3	LILRB4	SGK2	SNORD116-28	LGALS14	KIAA1919
DIRAS3	LIN28B	SIM2	SNORD116-22	LGALS8	LGALS14
DLGAP2	LMX1B	SLC22A18	SNORD115-7	OR51B5	LGALS8
DLK1	LRRTM1	SLC22A18AS	SNORD115-26	OVCH2	OR51B5
DLX5	LY6D	SLC22A2	SNORD115-13	PAPPA2	OVCH2
DNMT1	MAGEL2	SLC22A3	SNORD64	PTGFRN	PAPPA2
DVL1	MAGI2	SLC26A10	SNRPN	RHOBTB3	PTGFRN
E2F7	MCTS2P	SLC38A4	SNURF	SEC16A	RHOBTB3
EGFL7	MEG3	SLC4A2	SOX8	SPTLC3	SEC16A
EVX1	MEG8	SNORD107	SPATA31D1	SYNE2	SPTLC3
FAM50B	MEST	SNORD108	SPON2	TLR3	SYNE2
FASTK	MESTIT1	SNORD109A	ZC3H12C		TLR3
FBRSL1	MIMT1	SNORD109B	ZDBF2		
FERMT2	MIR184	SNORD115-1	ZFAT		
FGFRL1	MIR296	SNORD115-10	ZFAT-AS1		
FOXF1	MIR298	SNORD115-11	ZFP36L2		
FOXG1	MIR371A	SNORD115-14	ZIC1		
FUCA1	MKRN3	SNORD115-15	ZIM2		
GAB1	MPC1	SNORD115-16	ZIM3		
GABRA5	MRAP2	SNORD115-17	ZNF215		
GABRB3	MRGBP	SNORD115-2	ZNF225		
GABRG3	MZF1	SNORD115-22	ZNF229		
GAREM	MYEOV2	SNORD115-24	ZNF264		
GATA3	NAA60	SNORD115-25	ZNF597		
GATM	NAP1L4	SNORD115-27	TCEB3C		
GDAP1L1	NAP1L5	SNORD115-28	TFPI2		

* Mustaga märgitud geenid on võetud Geneimprint kodulehelt (veebiaadress 3).

** Rohelises kirjas on märgitud geenid, mis on võetud Court *et al.*, 2014 teadusartiklist. Tumesinises kirjas on märgitud geenid Metsalu *et al.*, 2014 artiklist.

Lisa 2.

Tabel 10. Käesolevas magistritöös tuvastatud mittevermitud geenid platsentas RNA-seqi andmete põhjal.

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Ekspr kohta	Kirjanduses raporteeritud ekspr alleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
AMPD3 ^A	rs16907852	Ekson	6	4,00	1 (17%)	2 (33%)	3 (50%)	Bi	Bi	Kõikjal
	rs3741041	Ekson	23	5,00	9 (39%)	3 (13%)	11 (48%)			
ANO1 ^A	rs10898112	Ekson	7	9,00	2 (29%)	1 (14%)	4 (57%)	Bi	Ema	Segagrupp
	rs2276066	Ekson	13	8,00	3 (23%)	4 (31%)	6 (46%)			
AOC1 ^C	rs10156191	Ekson	15	178,50	2 (13%)	2 (13%)	11 (73%)	Bi	RMAE	Platsenta
	rs1049748	Ekson	18	83,00	7 (39%)	3 (17%)	8 (44%)			
	rs10893	Ekson	13	91,50	1 (8%)	2 (15%)	10 (77%)			
ASCL2 ^A	rs7126721	5'UTR	5	7,00	3 (60%)	0 (0%)	2 (40%)	Bi	Bi	Segagrupp
CASC5 ^C	rs8040502	Ekson	13	3,00	1 (8%)	7 (54%)	5 (38%)	Bi	Isa#	Munandid
CLASP1 ^C	rs2304560	Ekson	6	12,00	1 (17%)	0 (0%)	5 (83%)	Bi	Isa#	Kõikjal
DHCR7 ^A	rs4316537	Ekson	5	4,00	1 (20%)	0 (0%)	4 (80%)	Bi	Bi	Kõikjal
DLG5 ^C	rs1058198	Ekson	19	123,00	0 (0%)	0 (0%)	19 (100%)	Bi	RMAE#	Kõikjal
	rs1058202	3'UTR	25	97,00	1 (4%)	0 (0%)	24 (96%)			
	rs1058203	3'UTR	19	92,00	0 (0%)	0 (0%)	19 (100%)			
	rs2165046	3'UTR	21	93,00	2 (10%)	0 (0%)	19 (90%)			
	rs2165047	3'UTR	19	41,00	0 (0%)	0 (0%)	19 (100%)			
DNMT1 ^A	rs2228611	Ekson	16	16,50	0 (0%)	4 (25%)	12 (75%)	Bi	Isa	Kõikjal
DOCK1 ^B	rs1060561	Ekson	14	22,50	1 (7%)	0 (0%)	13 (93%)	Bi	Isa	Segagrupp
	rs2229597	Ekson	11	59,50	0 (0%)	0 (0%)	11 (100%)			
	rs2296635	Ekson	16	9,00	3 (19%)	1 (6%)	12 (75%)			
	rs3740543	3' UTR	27	5,00	6 (22%)	4 (15%)	17 (63%)			

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Ekspri kohta	Kirjanduses raporteeritud ekspri alleel	Proteinaatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
EGFL7 ^A	rs2297538	Ekson	9	136,50	0 (0%)	0 (0%)	9 (100%)	Bi	Isa	Kõikjal
FERMT2 ^A	rs2357947	Ekson	6	29,00	0 (0%)	0 (0%)	6 (100%)	Bi	Isa	Segagrupp
	rs7143400	3'UTR	7	17,00	1 (14%)	0 (0%)	6 (86%)			
GAREM ^A	rs3786308	Ekson	11	9,50	4 (36%)	1 (9%)	6 (55%)	Bi	Isa	Segagrupp
GATA3 ^A	rs1058240	3'UTR	8	7,50	2 (25%)	0 (0%)	6 (75%)	Bi	Isa	Nahk
	rs10905277	UTR	19	11,00	1 (5%)	0 (0%)	18 (95%)			
GNAS ^A	rs7121	Ekson	21	85,00	0 (0%)	0 (0%)	21 (100%)	Bi	Isovorm sõltuv	Kõikjal
GRB10 ^A	rs1800504	Ekson	15	14,00	4 (27%)	1 (7%)	10 (67%)	Bi	Isovorm sõltuv	Kõikjal
	rs4947710	Ekson	11	5,50	4 (36%)	1 (9%)	6 (55%)			
GRHL1 ^C	rs1465638	3'UTR	13	24,50	0 (0%)	2 (15%)	11 (85%)	Bi	RMAE#	Söögitoru, nahk
	rs16867251	Ekson	14	95,00	0 (0%)	0 (0%)	14 (100%)			
	rs16867256	Ekson	14	74,00	0 (0%)	0 (0%)	14 (100%)			
	rs2303920	Ekson	16	56,00	0 (0%)	1 (6%)	15 (94%)			
	rs934303	3'UTR	13	22,50	1 (8%)	2 (15%)	10 (77%)			
HEG1 ^C	rs10804567	Ekson	16	13,00	1 (6%)	2 (13%)	13 (81%)	Bi	RMAE#	Kõikjal
	rs16836316	3'UTR	12	13,00	1 (8%)	2 (17%)	9 (75%)			
	rs1909586	3'UTR	20	14,00	4 (20%)	2 (10%)	14 (70%)			
	rs2860440	Ekson	19	20,50	0 (0%)	0 (0%)	19 (100%)			
	rs2981546	Ekson	19	18,00	0 (0%)	0 (0%)	19 (100%)			
	rs3184121	3'UTR	16	11,50	5 (31%)	5 (31%)	6 (38%)			
	rs4404487	Ekson	16	14,00	1 (6%)	0 (0%)	15 (94%)			
	rs6438869	Ekson	18	10,00	0 (0%)	2 (11%)	16 (89%)			

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Ekspri kohta	Kirjanduses raporteeritud ekspri alleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
HEG1 ^C	rs6438874	Ekson	19	23,00	1 (5%)	1 (5%)	17 (89%)	Bi	RMAE#	Kõikjal
	rs6790837	Ekson	19	7,00	5 (26%)	5 (26%)	9 (47%)			
IGF1R ^B	rs2684788	3'UTR	23	7,00	3 (13%)	2 (9%)	18 (78%)	Bi	Isa	Kõikjal
	rs3743249	3'UTR	11	12,00	2 (18%)	2 (18%)	7 (64%)			
IGF2R ^A	rs3743262	Ekson	5	17,00	1 (20%)	0 (0%)	4 (80%)			
	rs1803989	Ekson	6	22,50	0 (0%)	0 (0%)	6 (100%)	Bi	Bi	Kõikjal
	rs1805075	Ekson	7	23,00	0 (0%)	0 (0%)	7 (100%)			
KIAA1551 ^C	rs998075	Ekson	17	35,50	0 (0%)	0 (0%)	17 (100%)			
	rs10771894	Ekson	20	10,50	6 (30%)	2 (10%)	12 (60%)	Bi	RMAE#	Kõikjal
	rs16919122	Ekson	11	16,00	0 (0%)	0 (0%)	11 (100%)			
	rs2388982	Ekson	17	23,00	0 (0%)	0 (0%)	17 (100%)			
KIAA1919 ^C	rs3759300	Ekson	7	25,00	0 (0%)	0 (0%)	7 (100%)			
	rs3759302	Ekson	11	15,00	1 (9%)	1 (9%)	9 (82%)			
	rs240987	3'UTR	9	18,00	0 (0%)	0 (0%)	9 (100%)	Bi	RMAE#	Segagrupp
	rs240988	3'UTR	13	17,00	0 (0%)	0 (0%)	13 (100%)			
LGALS14 ^C	rs4830	Ekson	15	56,00	0 (0%)	3 (20%)	12 (80%)	Bi	Isa	Platsenta
LGALS8 ^C	rs1041934	Ekson	15	49,50	2 (13%)	0 (0%)	13 (87%)	Bi	Ema	Kõikjal
	rs1041939	3'UTR	1	33,50	0 (0%)	0 (0%)	1 (100%)			
	rs15701	Ekson	15	38,00	1 (7%)	0 (0%)	14 (93%)			
	rs2273863	Ekson	1	33,00	0 (0%)	0 (0%)	1 (100%)			
	rs2472126	Ekson	15	35,50	1 (7%)	1 (7%)	13 (87%)			
	rs2758997	3'UTR	14	36,00	0 (0%)	0 (0%)	14 (100%)			
MKRN3 ^A	rs2239669	Ekson	14	8,00	3 (21%)	5 (36%)	6 (43%)	Bi	Isa	Suurajukoor, platsenta
NLRP2 ^A	rs10412915	Ekson	14	23,50	6 (43%)	0 (0%)	8 (57%)	Bi	Ema	Segagrupp

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Eksprikohta	Kirjanduses raporteeritud ekspir alleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
PKP3 ^A	rs1134578	Ekson	21	17,00	0 (0%)	1 (5%)	20 (95%)	Bi	Ema	Söögitoru, nahk
	rs11748	Ekson	22	18,50	0 (0%)	0 (0%)	22 (100%)			
PTGFRN ^C	rs10801923	3'UTR	19	12,00	4 (21%)	5 (26%)	10 (53%)	Bi	RMAE#	Segagrupp
	rs10923191	3'UTR	19	6,00	3 (16%)	3 (16%)	13 (68%)			
	rs1127655	3'UTR	19	12,00	7 (37%)	4 (21%)	8 (42%)			
	rs13951	3'UTR	20	7,00	5 (25%)	2 (10%)	13 (65%)			
	rs3829881	3'UTR	19	9,50	0 (0%)	2 (11%)	17 (89%)			
PTPN14 ^A	rs1135352	Ekson	11	16,50	1 (9%)	0 (0%)	10 (91%)	Bi	Ema	Segagrupp
	rs1472441	3'UTR	15	18,00	0 (0%)	0 (0%)	15 (100%)			
	rs2291831	3'UTR	10	10,00	1 (10%)	1 (10%)	8 (80%)			
	rs7543972	3'UTR	5	10,00	1 (20%)	0 (0%)	4 (80%)			
SEC16A ^C	rs10781497	Ekson	9	18,00	0 (0%)	0 (0%)	9 (100%)	Bi	RMAE#	Kõikjal
	rs1127162	3'UTR	9	8,00	3 (33%)	1 (11%)	5 (56%)			
	rs3812589	3'UTR	9	7,00	4 (44%)	1 (11%)	4 (44%)			
	rs3812595	Ekson	15	10,00	6 (40%)	1 (7%)	8 (53%)			
SLC22A3 ^A	rs1810126	3'UTR	10	6,50	1 (10%)	2 (20%)	7 (70%)	Bi	Ema ^s	Segagrupp
	rs2076828	3'UTR	12	3,00	5 (42%)	2 (17%)	5 (42%)			
	rs2292334	Ekson	10	6,00	3 (30%)	1 (10%)	6 (60%)			
SLC4A2 ^A	rs2303937	Ekson	17	48,00	0 (0%)	0 (0%)	17 (100%)	Bi	Ema	Kõikjal
SPTLC3 ^C	rs243887	Ekson	20	256,50	1 (5%)	2 (10%)	17 (85%)	Bi	Isa	Segagrupp
	rs243888	Ekson	21	245,50	0 (0%)	0 (0%)	21 (100%)			
	rs3761896	5'UTR	17	145,00	1 (6%)	1 (6%)	15 (88%)			
	rs4510290	3'UTR	20	80,50	0 (0%)	0 (0%)	20 (100%)			
	rs6078938	Ekson	6	122,00	0 (0%)	0 (0%)	6 (100%)			

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Ekspri kohta	Kirjanduses raporteeritud ekspri alleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
SYNE2 ^C	rs1890908	Ekson	4	44,00	0 (0%)	0 (0%)	4 (100%)	Bi	RMAE#	Kõikjal
	rs3829767	Ekson	11	62,00	0 (0%)	0 (0%)	11 (100%)			
	rs4027402	Ekson	9	57,50	0 (0%)	0 (0%)	9 (100%)			
	rs4027404	Ekson	9	44,50	0 (0%)	0 (0%)	9 (100%)			
	rs4027405	Ekson	6	50,50	0 (0%)	0 (0%)	6 (100%)			
	rs4902264	Ekson	9	65,50	0 (0%)	0 (0%)	9 (100%)			
	rs8010699	Ekson	11	75,50	0 (0%)	0 (0%)	11 (100%)			
	rs9944035	Ekson	5	90,50	0 (0%)	0 (0%)	5 (100%)			
TLR3 ^C	rs3775291	Ekson	18	14,00	1 (6%)	2 (11%)	15 (83%)	Bi	RMAE#	Segagrupp
	rs3775296	UTR	15	11,50	0 (0%)	1 (7%)	14 (93%)			
TRAPPC9 ^B	rs3735801	Ekson	21	43,00	0 (0%)	0 (0%)	21 (100%)	Bi	Isa	Kõikjal
	rs3735803	Ekson	19	41,50	0 (0%)	0 (0%)	19 (100%)			
WDR27 ^B	rs1137523	5'UTR	14	6,00	4 (29%)	4 (29%)	6 (43%)	Bi	Isa	Kõikjal
	rs4236176	Ekson	14	6,00	4 (29%)	2 (14%)	8 (57%)			
ZNF264 ^A	rs917340	Ekson	12	12,50	1 (8%)	0 (0%)	11 (92%)	Bi	Teadmata	Kõikjal
ZNF331 ^B	rs8109631	Ekson	14	18,50	1 (7%)	3 (21%)	10 (71%)	Bi	Isa	Kõikjal
	rs8110538	3'UTR	14	6,50	2 (14%)	6 (43%)	6 (43%)			
CCDC85A ^A	rs960489	5'UTR	9	8,00	4 (44%)	1 (11%)	4 (44%)	Bi	Isa	Suurajukoor
INPP5F ^A	rs3188055	Ekson	11	6,00	1 (9%)	2 (18%)	8 (73%)	Bi	Isa	Kõikjal
ZDBF2 ^A	rs10932150	Ekson	18	19,00	1 (6%)	13 (72%)	4 (22%)	RMAE/Bi	Isa	Kõikjal
	rs1448902	3'UTR	14	3,00	3 (21%)	7 (50%)	4 (29%)			
	rs3732084	Ekson	14	11,00	1 (7%)	7 (50%)	6 (43%)			
	rs7582864	Ekson	16	13,00	2 (13%)	6 (38%)	8 (50%)			
PPP1R9A ^A	rs12536183	Ekson	10	5,00	7 (70%)	1 (10%)	2 (20%)	RMAE/Bi	Ema	Segagrupp

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Eksp kohta	Kirjanduses raporteeritud ekspr alleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
DCAF10 ^B	rs10115457	3'UTR	13	4,00	4 (31%)	6 (46%)	3 (23%)	RMAE/Bi	Isa	Kõikjal
GATM ^A	rs12437887	5'UTR	18	12,00	11 (61%)	4 (22%)	3 (17%)	RMAE/Bi	Teadmata	Neerud, maks, pankreas. Segagrupp
L3MBTL1 ^A	rs2071969	5'UTR	6	3,00	3 (50%)	2 (33%)	1 (17%)	RMAE/Bi	Isa	Kõikjal
CEP63 ^C	rs1874883	5'UTR	18	3,00	6 (33%)	12 (67%)	0 (0%)	RMAE/Bi	RMAE#	
	rs9821505	3'UTR	11	3,00	6 (55%)	5 (45%)	0 (0%)			
KIAA1191 ^C	rs9827878	Ekson	13	3,50	3 (23%)	7 (54%)	3 (23%)	RMAE/Bi	Ema#	Kõikjal
	rs409744	3'UTR	19	3,00	9 (47%)	8 (42%)	2 (11%)			
NAA60 ^A	rs13740	3'UTR	13	4,00	8 (62%)	0 (0%)	5 (38%)	RMAE/Bi	Ema	Kõikjal
	rs2074524	3'UTR	4	5,50	0 (0%)	1 (25%)	3 (75%)			
	rs2285815	3'UTR	14	5,00	6 (43%)	3 (21%)	5 (36%)			
RHOBTB3 ^C	rs34896	Ekson	18	342,00	7 (39%)	8 (44%)	3 (17%)	RMAE/Bi	Isa	Kõikjal
MCCC1 ^B	rs2270968	Ekson	16	5,00	2 (13%)	10 (63%)	4 (25%)	RMAE/Bi	Isa	Kõikjal
PDE4D ^B	rs10071088	3' UTR	12	3,00	5 (42%)	5 (42%)	2 (17%)	RMAE/Bi	Isa	Kõikjal
	rs702531	3' UTR	13	3,00	3 (23%)	8 (62%)	2 (15%)			
TMEM60 ^A	rs7953	Ekson	16	3,00	9 (56%)	3 (19%)	4 (25%)	RMAE/Bi	Isa	Kõikjal
OBSCN ^A	rs12035900	Ekson	14	4,00	5 (36%)	4 (29%)	5 (36%)	RMAE/Bi	Isa	Skeletilihased.
SPON2 ^A	rs2242278	Intron/ platsenta spetsiifiline transkript	19	4,00	11 (58%)	8 (42%)	0 (0%)	RMAE	Isa	Eesnääre.
CHMP2A ^A	rs873104	5'UTR	19	6,00	10 (53%)	9 (47%)	0 (0%)	RMAE	Ema	Kõikjal
FAM20A ^B	rs2286562	Ekson	15	3,00	5 (33%)	10 (67%)	0 (0%)	RMAE	Isa	Segagrupp
	rs2302234	Ekson	17	4,00	5 (29%)	11 (65%)	1 (6%)			
	rs2907373	Ekson	9	4,00	2 (22%)	4 (44%)	3 (33%)			

Geen	SNP ID	SNP positsioon	Inform. Ind.	NGEM selles pos.	Lugemite päritolu (indiviidide arv)			Eksprikohta	Kirjanduses raporteeritud eksprialleel	Proteinatlas RNA koe grupp
					Ainult emapoolsed	Ainult isapoolsed	Bialleelsed			
GLI3 ^A	rs2051935	3'UTR	17	3,00	9 (53%)	6 (35%)	2 (12%)	RMAE/Bi	Ema	Segagrupp
	rs3735361	3'UTR	21	4,00	10 (48%)	6 (29%)	5 (24%)			
	rs846266	Ekson	15	4,00	4 (27%)	6 (40%)	5 (33%)			
PAPPA2 ^C	rs1044299	3'UTR	11	260,00	6 (55%)	2 (18%)	3 (27%)	Konfliktne	Ema	Platsenta
	rs4492574	3'UTR	10	708,50	1 (10%)	0 (0%)	9 (90%)			

A -Geneimprinti andmebaasist; B - Court et al., 2014; C- Metsalu et al., 2014; \$- I trimestri spetsiifiline vermitus; #- vermituse kandidaatgeen, millel on alleelse kallutatusega ekspressioon

Lisa 3.

Tabel 11. Geneimprint andmebaasi põhjal inimese ja hiire vermitud ortoloogide ülekate.

1. Vermitud sama alleel

GPRI, ZDBF2, NAP1L5, PLAGL1, HYMAI, DDC, PEG10, SGCE, CALCR, PPP1R9A, TFPI2, MEST, KLF14, KCNK9, PEG13, INPP5F, KCNQ1OT1, IGF2AS, PHLDA2, H19, KCNQ1, IGF2, SLC22A18, INS, CDKN1C, ANO1, HTR2A, MEG3, DLK1, RTL1, MEG8, MAGEL2, MKRN3, UBE3A, ATP10A, PWCR1, SNRPN, NDN, SNURF, PEG3, PSIMCT-1, BLCAP, NNAT, MCTS2, GNAS, MIR298, MIR296, SLC22A2, SLC22A3

2. Vermitud erinev alleel

GRB10, MAGI2, COPG2, ZIM2

B. Vermitud kas hiires või inimeses

1. Vermitud hiires, inimeses mitte

Commd1, Gab1, Igf2r, Asb4, Pon3, Pon2, Sfmbt2, Ampd3, Nap1l4, Th, Tssc4, Ascl2, Cd81, Dher7, Slc38a4, Dcn, Dio3, Znf127as, Gatm, Rasgrf1, Mir184, Impact, Znf264, Usp29, Zim3, Hm13, Tsix, Xist

2. Vermitud inimeses, hiires mitte

GFII, FUCAI, BMP8B, DVLI, TP73, WRAP73, RPL22, PRDM16, PEX10, TMEM52, PTPN14, OBSCN, OR11L1, LRRTM1, OTX1, VAX2, CCDC85A, CYP11B1, ABCG8, ZFP36L2, MYEOV2, ALDH1L1, ZIC1, HES1, FGFRL1, UVSSA, SPON2, ADAMTS16, CDH18, CSF2, BTNL2, FAM50B, MRAP2, AIM1, TMEM60, LIN28B, EVX1, HOXA4, HOXA2, HOXA11, HOXA3, MPC1, GLI3, HOXA5, PONI, DLX5, CPA4, FASTK, SLC4A2, PURG, DLGAP2, NKAIN3, LY6D, GPT, ZFAT, GLIS3, APBA1, C9orf85, SPATA31D1, LMX1B, C9orf116, EGFL7, PHPT1, GATA3, CTNNA3, LDB1, NKX6-2, CFAP46, PAOX, WT1, PKP3, OSBPL5, B4GALNT4, TRPM5, IFITM1, RAB1B, ZC3H12C, KBTBD3, SDHD, ABCC9, SLC26A10, HOXC4, HOXC9, CDK4, E2F7, FBRSL1, RB1, FAM70B, FOXG1, FERMT2, SNORD116@, SNORD115@, GABRA5, GABRB3, GABRG3, SNORD64, NAA60, ZNF597, SOX8, SALL1, USB1, ACD, FOXF1, TMEM88, HOXB2, HOXB3, CELF4, GAREM, PLPP2, DNMT1, CHMP2A, TSHZ3, CHST8, MZF1, NLRP2, ISM1, GDAP1L1, SGK2, COL9A3, L3MBTL, MRGBP, SIM2

@- geeniklaster, sinine - inimesel isapoolne, punane - inimesel emapoolne, must – sõltub isovormist

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina Diana Nõmmemees (sünnikuupäev: 28.11.1991)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Vanemaspetsiifiline geeniekspressioon inimese platsentas“, mille juhendaja on prof Maris Laan,
 - 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates **01.01.2020** kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 27. mai 2016